DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 1989-2019
DOI: 10.1515/acgeo-2016-0084

Various Approaches to Forward and Inverse Wide-Angle Seismic Modelling Tested on Data from DOBRE-4 Experiment

Tomasz JANIK[1], Piotr ŚRODA[1], Wojciech CZUBA[1], and Dmytro LYSYNCHUK[2]

[1]Institute of Geophysics, Polish Academy of Sciences, Warsaw, Poland; e-mail: janik@igf.edu.pl

[2]Institute of Geophysics, National Academy of Sciences of Ukraine, Kiev, Ukraine

Abstract

The interpretation of seismic refraction and wide angle reflection data usually involves the creation of a velocity model based on an inverse or forward modelling of the travel times of crustal and mantle phases using the ray theory approach. The modelling codes differ in terms of model parameterization, data used for modelling, regularization of the result, *etc*. It is helpful to know the capabilities, advantages and limitations of the code used compared to others.

This work compares some popular 2D seismic modelling codes using the dataset collected along the seismic wide-angle profile DOBRE-4, where quite peculiar/uncommon reflected phases were observed in the wavefield.

The ~505 km long profile was realized in southern Ukraine in 2009, using 13 shot points and 230 recording stations. Double P_MP phases with a different reduced time (7.5-11 s) and a different apparent velocity, intersecting each other, are observed in the seismic wavefield. This is the most striking feature of the data. They are interpreted as reflections from strongly dipping Moho segments with an opposite dip. Two steps were used for the modelling. In the previous work by Starostenko *et al.* (2013), the trial-and-error forward model based on refracted and reflected phases (SEIS83 code) was published. The interesting feature is the high-amplitude (8-17 km) variability of the Moho depth in the form of downward and upward bends. This model is compared with results from other seismic inversion methods: the first arrivals tomography package FAST based on first arrivals; the JIVE3D code, which can also use later refracted arrivals and reflections; and the forward and inversion code RAYINVR using both refracted and reflected phases. Modelling with all the codes tested showed substantial variability of the Moho depth along the DOBRE-4 profile. However, SEIS83 and RAYINVR packages seem to give the most coincident results.

Key words: seismic modelling, Moho boundary, ray tracing, tomography inversion.

1. INTRODUCTION

The interpretation of data from seismic refraction and wide angle reflection profiles usually involves the creation of the P-wave velocity model based on forward or automated inverse modelling of the travel times of observed crustal and mantle phases. Such a model is sometimes considered as a first step, followed for example by synthetic seismogram calculations using the full-waveform method, or as a final result. The codes used for velocity modelling are usually based on ray theory and allow for tracing of seismic rays through the medium with a given velocity distribution and for calculation of corresponding travel times from sources to receivers. The velocity model is sought by minimizing the difference between observed and calculated travel times. Two approaches for such a modelling exist. In the first one, the forward problem is solved for a given initial model by calculation of ray paths and travel times; then, the inverse problem is solved by manual trial-and-error modifications of the model, seeking to improve the fit of the calculated and observed travel times. Both steps are applied repetitively. It is possible to verify and, if needed, correct the phase identification during this procedure. The second approach is a mostly automated procedure. An initial model is updated in a number of iterative steps to minimize the misfit of observed and calculated travel times, *e.g.*, in the simplest case according to the regularized least squares method.

During the last few decades, an inverse (tomographic) modelling mostly replaced a forward approach as a faster and more effective tool for the determination of the seismic velocity distribution, due to several drawbacks of the latter method. A manual, trial-and-error process of attempting to fit the data for several shotpoints simultaneously can be very tedious, difficult and time-consuming, compared to inverse methods. Also, as the sequence of model modifications depends on the interpreter's decisions, it introduces some subjectivity into the final solution. The result of the inverse modelling is also subjective to some extent − as in any type of travel time modelling, it depends on a subjective process of picking the seismic phases. Moreover, it depends on arbitrary settings of the inversion procedure (*e.g.*, choice of the grid spacing, values of the smoothing/regularization weights, *etc.*) but once these parameters are set, given the input data, the calculations are easy to repeat, *e.g.*, in order to independently verify the result, which is not the case for the forward approach. Also, a systematic evaluation of the model resolution and uncertainty is problematic for the forward approach.

However, the advantage of the forward modelling is the possibility to continuously control the correctness of the interpretation of seismic data (identification of the travel times of given seismic phase as a refraction or reflection from a particular layer/discontinuity). At each stage, the interpreter can verify the identification of picked seismic phases. For example, difficulties in fitting some particular seismic phase by a model consistently fitting all the remaining data can help to decide if a "suspect" phase actually comes from an "in-line", 2D structure (*e.g.*, from a local high-velocity body) and should be incorporated into a 2D model, or if it originates as a side reflection or diffraction from some local structure, and thus should not be used for the calculation of a 2D model. This helps to avoid significant model errors, bias or artefacts that could be introduced by the modelling of some (initially) wrongly identified phases, whose nature at first glance may not be obvious. The proper identification of seismic phases used for modelling was stressed by Zelt (1999) who wrote that "A model developed by the analysis of wide-angle travel time data is only as good as the picks".

The DOBRE-4 WARR profile (Starostenko *et al.* 2013) was realized in southern Ukraine in 2009. The experiment was aimed at investigating the structure of the crust and uppermost mantle at the southwestern corner of the East European Craton (the East European Platform and Ukrainian Shield), at its transition into the Trans-European Suture Zone (TESZ). The location of the DOBRE-4 profile is shown in Fig. 1. It extends (Fig. 2) from the Lower Prut High in the SW through the South Ukrainian Homocline to the Ukrainian Shield. The profile crosses the Precambrian East European Craton and its younger SW part, the Scythian Platform. Its margin is a major tectonic boundary, the Teisseyre–Tornquist Zone (Guterch *et al.* 1986), representing contact with younger, Palaeozoic and Alpine units.

The seismic model along the profile and its geological and tectonic interpretation is presented by Starostenko *et al.* (2013). The unusual features observed in the seismic wavefield are triplications of the Moho reflections. They are visible in several record sections as superimposed branches of the P_MP phase with a different apparent velocity, intersecting each other (Figs. 3a and b). Such recordings, interpreted as reflections from a steeply undulating Moho discontinuity, are rather uncommon and present a valuable

Fig. 1. Location of the DOBRE-4 profile. The yellow stars represent shot points, grey dots show the recording stations of the experiment. The inset map indicates the location of the study area in Europe. EEC – East European Craton, STZ – Sorgenfrei-Tornquist Zone, TTZ – Teisseyre–Tornquist Zone.

material to test the various approaches of 2D wide-angle modelling (this unusual phenomenon was previously seen on other DSS profile in Ukrainian Shield and interpreted as P_MP triplication in paper by Grad and Tripolsky (1995)). Therefore, this paper is an attempt to test how such a complex Moho structure can be resolved, depending on the phases used for modelling (P_MP only, P_n only, $P_MP + P_n$) but also depending on the modelling approach and model parameterization.

To compare the published model obtained by modelling using the SEIS83 package (Červený and Pšenčík 1984), we have chosen, from the many codes available (*e.g.*, also Korenaga *et al.* 2000, Koulakov 2009, Rawlinson and Urvoy 2006), several well known and popular codes: FAST by Zelt and Barton (1998) – a code for tomographic inversion of travel times corresponding to refracted waves, JIVE3D by Hobro (1999) and Hobro *et al.*

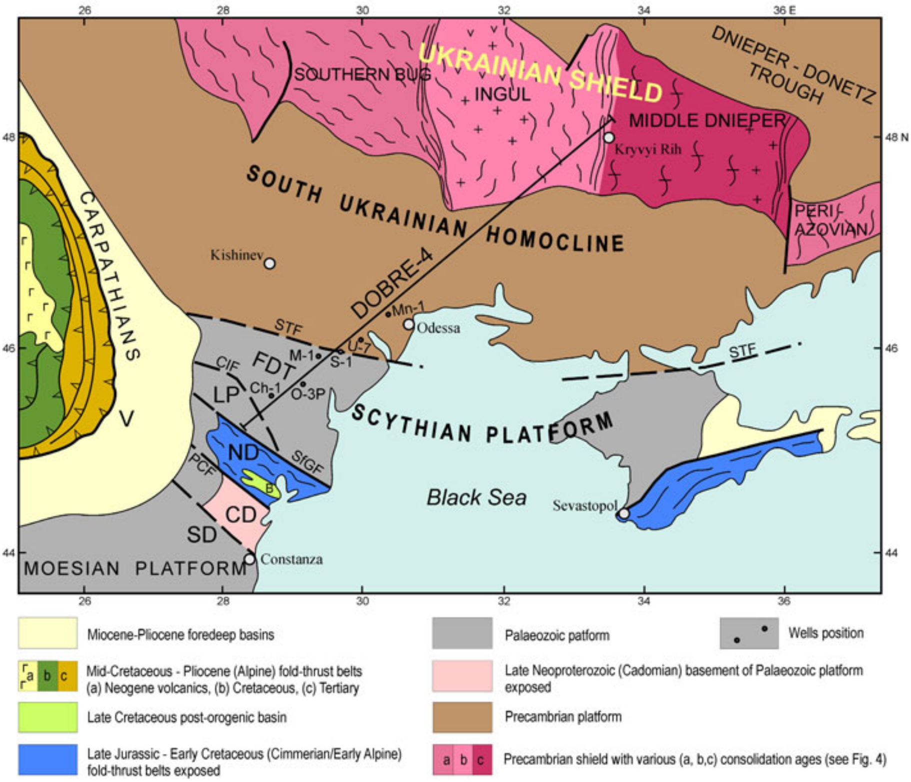


Fig. 2. Tectonic setting of profile DOBRE-4. Abbreviations: B – Babadag Basin, CD – Central Dobrudja, CIF – Cahul-Ismail Fault, FDT – Fore-Dobrudja Trough, LP – Lower Prut High, ND – North Dobrudja, PCF – Peceneaga-Camena Fault, SD – Southern Dobrudja, SfGF – Sfantu Gheorghe Fault, STF – Scytho-Turanian Fault, V – Vrancea. The deep boreholes: Ch1 – Chervonoarmeyskaya-1, O-3P – Orekhovskaya-3P, M-1 – Mirnopolskaya-1, S-1 – Saratskaya-1, U-7 – Uspenovskaya-7, Mn-1 – Mirnenskaya-1.

(2003) – a joint refraction and reflection travel time tomography code, and RAYINVR by Zelt and Smith (1992) – code for travel time inversion with a small number of model parameters related by *a priori* functionals for relatively simple final models. In general, available codes differ in terms of velocity model parameterization (equidistant grid or user-defined arbitrary grid, adaptive grid, single grid or separate grids for predefined layers separated by velocity discontinuities), the travel time data used for modelling (first arrivals only or all observed arrivals, including reflections), smoothing/regularization of the result. This is by no means a representative selection of various modelling methods, but anyway it allows us to check the result of

(a)

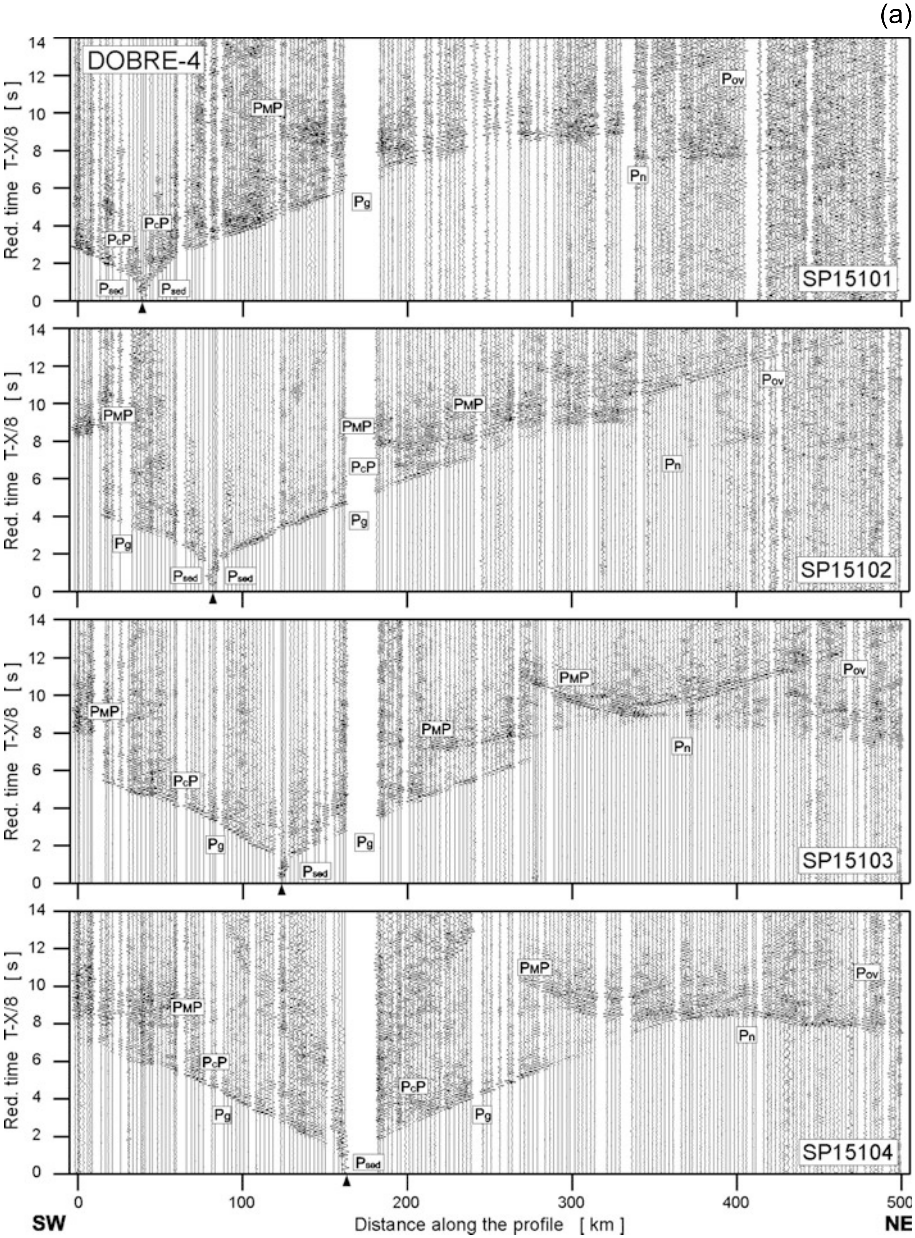


Fig. 3. Caption on next page.

modelling for a few basically different approaches – forward or inverse modelling, usage of reflected (P_MP) or refracted (P_n) waves, representation of the velocity model as a single layer or as multiple layers, parameterization

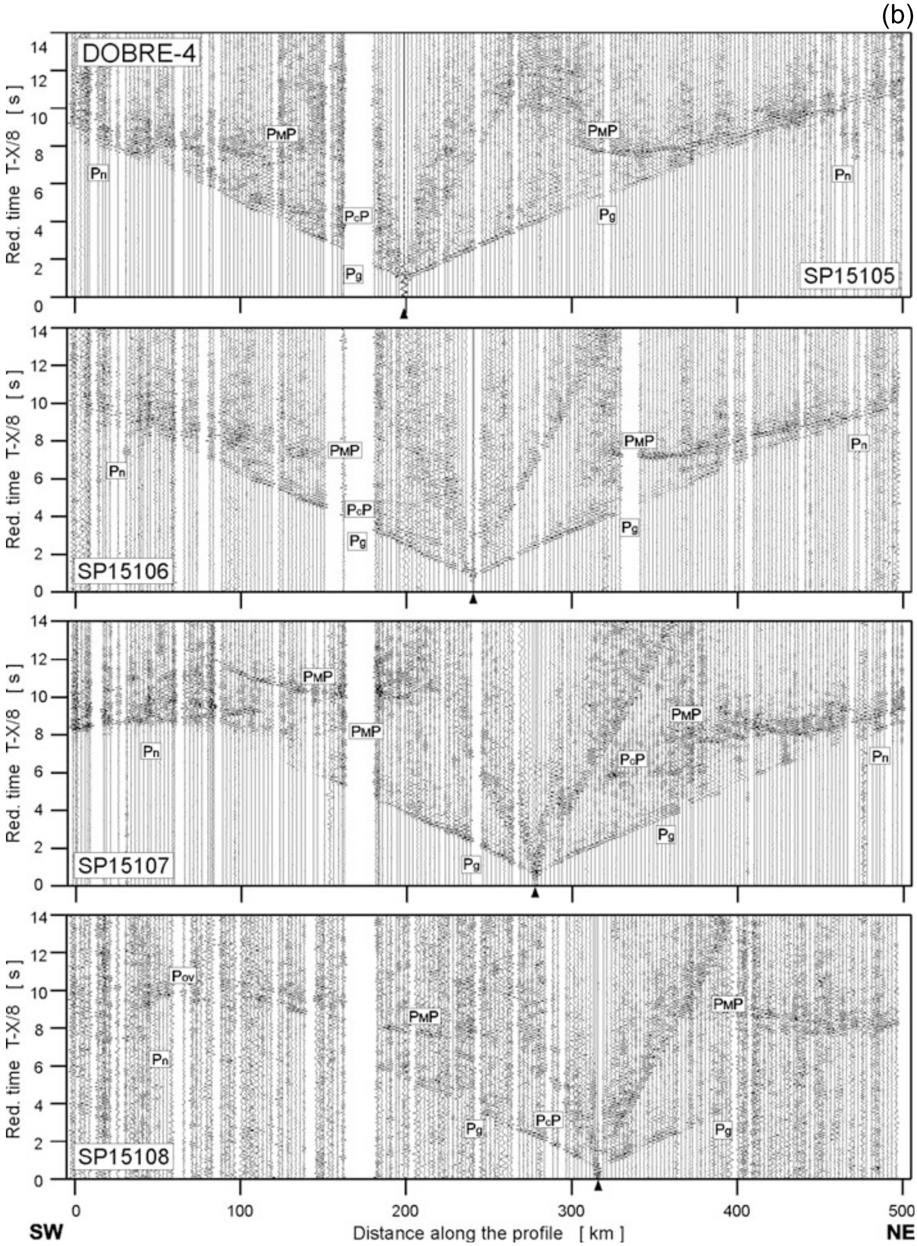


Fig. 3. Caption on next page.

on dense, regular velocity grid *versus* a sparse, irregular grid. Recently, a comparison of computational possibilities and limitations of above mentioned codes was published by Malinowski (2013).

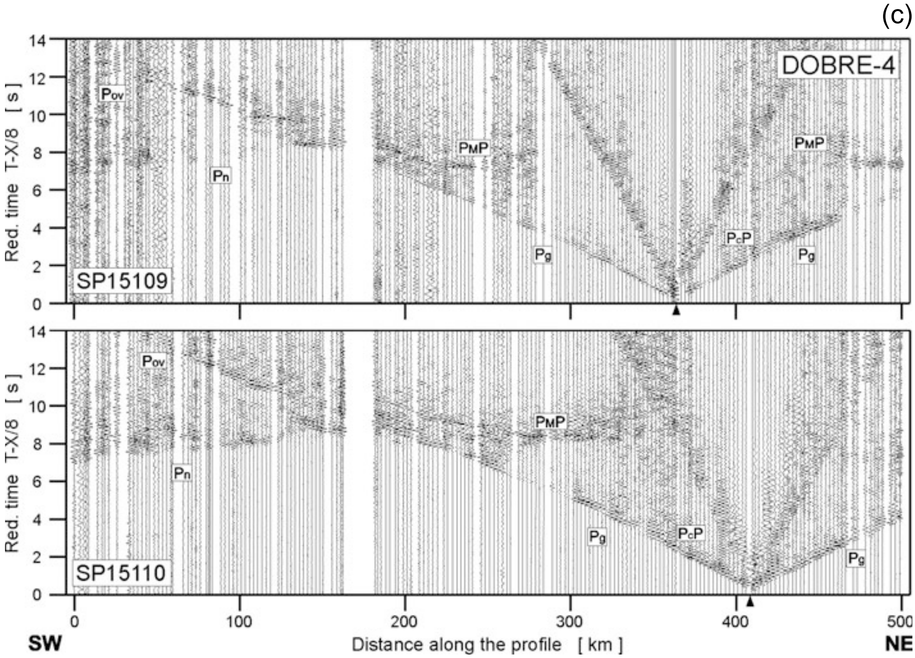


Fig. 3: (a)-(c) Example of trace-normalized, vertical-component seismic record sections for the P-wave (SP15101-SP15110), filtered by the band-pass filter 2-12 Hz. Abbreviations: P_g – seismic refractions from the upper and middle crystalline crust, P_{ov} – overcritical crustal phases, P_cP – reflections from the middle crust discontinuities, P_MP – reflected waves from the Moho boundary, P_n – refractions from the sub-Moho upper mantle. The reduction velocity is 8.0 km/s.

2. DATA

The DOBRE-4 profile is ~505 km long. The field acquisition included 13 shot points (SP), every 30-50 km and 230 recording stations, every 2.5 km.

The seismic sections recorded along the profile are of good quality. A detailed description of the recorded wavefield was presented by Starostenko *et al.* (2013). The seismic record sections are presented in Fig. 3. The record sections show a complex character of the wavefield, particularly the phases reflected from the Moho discontinuity, reflecting its complex topography along the profile.

2.1 P-wave refracted and reflected arrivals

The first P-wave arrivals represent the refracted waves from the upper crustal sedimentary layers (P_{sed}), the upper/middle crystalline crust (P_g), and the phases from the upper mantle (P_n). The P_{sed} phases, representing the refractions in the sedimentary sequences, are observed in the southern part of the profile. At larger distances (10-200 km), the P_g phase is observed. The waves propagating in the mantle (P_n) are observed on several shot points for the offset range of about 200-400 km. They represent refractions below the Moho and, possibly, reflections from the mantle discontinuities and scattered/diffracted waves. The latter phases may appear to be in "first arrivals" if the real first arrival phase – the Moho refraction – is not visible in the seismic section due to its low amplitude.

Reflected phases at 7.5-11 s reduced travel time and 100-250 km offset range are the most striking feature observed in the DOBRE-4 dataset. They are very coherent and with extremely high amplitudes. In the sections at both ends of the profile, they represent a typical P_MP phase (reflection from the Moho boundary). On the other sections, we often observe two phases – with a slow and fast apparent velocity, intersecting each other (*e.g.*, SP15103, Fig. 3) or arriving at a different reduced time (*e.g.*, SP15107 with a shallower Moho reflection at ~8 s reduced time and a deeper one at ~10 s reduced time).

Travel time picking uncertainties, estimated based on the pulse width and on the signal-to-noise ratio and used for modelling with all tested codes, are 0.1 s for the refracted waves (P_{sed}, P_g, P_{ov}, P_n) and 0.2 s for Moho reflections (P_MP). P_{ov} are phases of later refracted crustal arrivals ("overcritical").

3. SEISMIC MODELLING

3.1 Forward modelling with SEIS83

The trial-and-error forward modelling was done (Starostenko *et al.* 2013) using the SEIS83 package (Červený and Pšenčík 1984) with the graphical interface MODEL (Komminaho 1998) and ZPLOT (Zelt 1994). The code uses the shooting method (tracing a ray starting at the source for given initial direction by solving the ray (eikonal) equation) for calculation of the ray paths, travel times, and synthetic seismograms in the high-frequency approximation. The model consists of layers with seismic velocities parameterized on a non-equidistant, user-defined rectangular grid and interpolated by bicubic splines. Velocity discontinuities are allowed between the individual layers. An initial model of the velocity distribution in the uppermost crust presented in Fig. 4 was used in the modelling.

Resolving the velocity distribution of the upper crustal sedimentary layers along the profile is important for the further modelling of the deeper structure of the crust. The geological and geophysical data, boreholes and nearby CDP studies were used to prepare a cross section and the starting model of the upper crust (Fig. 4a).

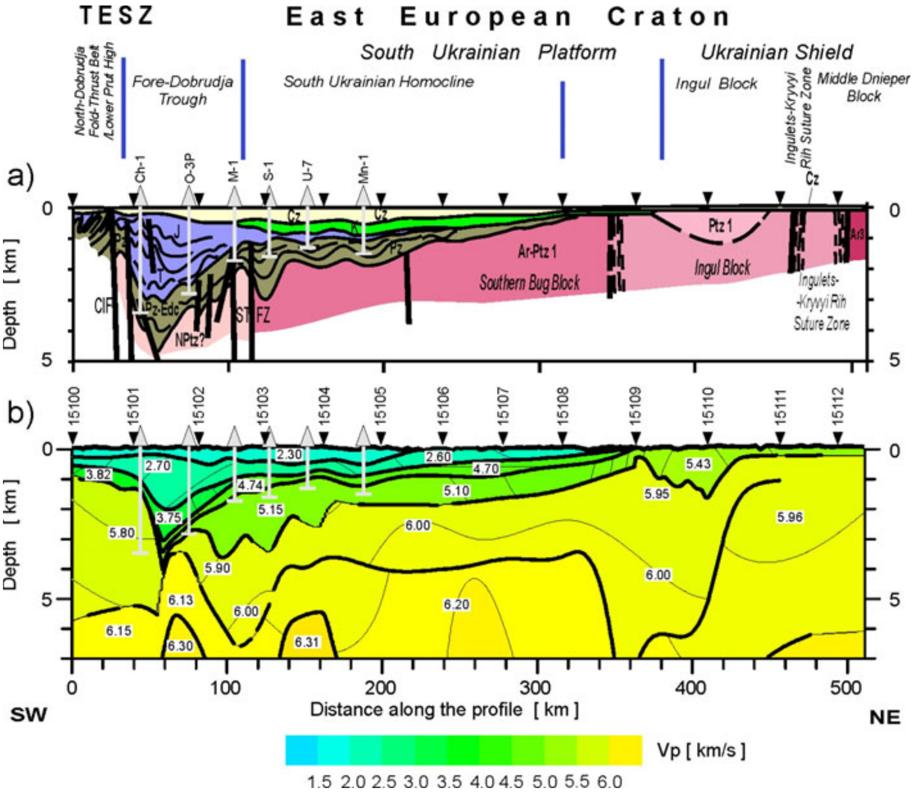
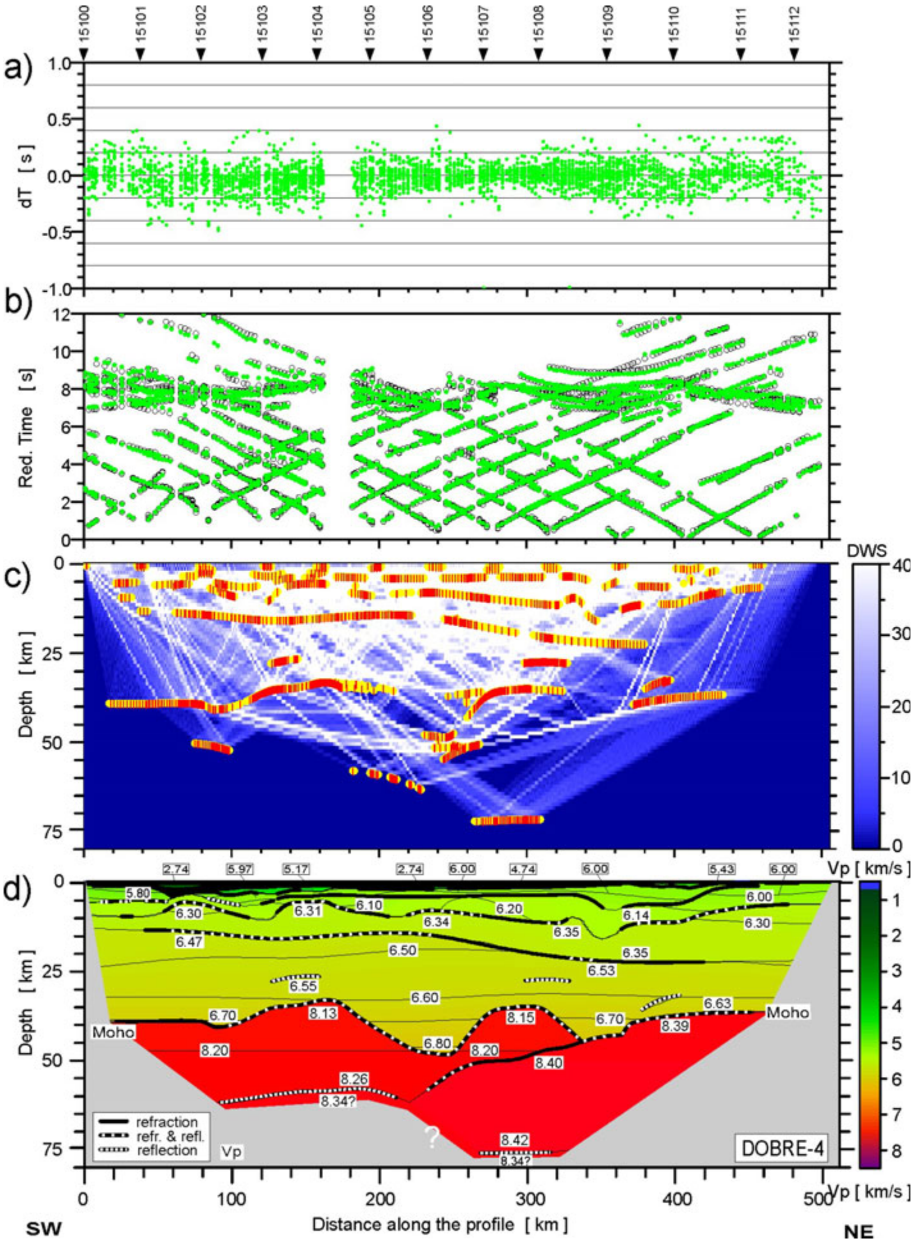


Fig. 4: (a) A simplified geological cross section based on seismic investigations, outcrop geology data, and six 1.5 km to 3.5 km deep boreholes located within 5.6 km of the profile, in the southwestern part (abbreviations are as in Fig. 2). Age assignments: Ar – Archaean, Ptz – Proterozoic, Edc – Ediacaran, Pz – Palaeozoic, T – Triassic, J – Jurassic, K – Cretaceous, Cz – Cenozoic. Major faults: CIF – Cahul-Ismail Fault, STFZ – Scytho-Turanian Fault Zone. These data were used as a basis for building a starting model (for details see text); (b) The final 2D model (SEIS83) of the seismic P-wave velocity in the sedimentary cover; vertical exaggeration is ~19.3:1. The position of large-scale crustal blocks is indicated. The arrows show positions of shot points. The grey triangles mark location of boreholes.

Fig. 5. 2D model of seismic P-wave velocity in the crust and upper mantle derived by forward ray tracing modelling using the SEIS83 package: (a) Travel time residuals; (b) Diagrams showing theoretical (black) and observed travel times (green); and (c) Ray coverage. Yellow lines – fragments of discontinuities constrained by reflected phases. The red points plotted along the interfaces mark the bottoming points of the modelled reflected phases (every third point is plotted) and their density is a measure of the positioning accuracy of the reflectors. DWS – derivative weight sum. The reduction velocity is 8 km/s; (d) The model is the same as that published by Starostenko *et al.* (2013). Respective calculated travel times differ by no more than



0.1 s from the previous one. Those parts of the first order discontinuities that have been constrained by reflected or/and refracted arrivals of P waves are marked by thick lines. The thin lines represent velocity isolines with values in km/s shown in white boxes. The vertical exaggeration is ~2.4:1.

The solution was sought in a succession of attempts to improve the data fit by modifying model parameters. Based on travel time misfit, the model was manually modified in order to enhance the data fit. The modelling also involved the calculation of synthetic seismograms. Amplitudes of synthetic and observed data were then qualitatively compared. This provided better constraints on the velocity gradients and contrasts at the discontinuities.

The high quality of the analysed data allowed for constructing a velocity model (Fig. 5) that fits the observed (experimental) travel times for both refracted and reflected waves with good accuracy. An example is shown in Fig. 13 in the paper by Starostenko *et al.* (2013). Diagrams showing theoretical and observed travel times for all the phases along the profile, ray coverage and travel time residuals from forward modelling are shown in Fig. 5. The RMS (and χ^2) values are 0.36 s ($\chi^2 = 13$) for P_{sed}, 0.16 s ($\chi^2 = 2.6$) for the P_g and P_{ov} phases (waves refracted in the crust), 0.17 s ($\chi^2 = 0.72$) for P_MP, and 0.31 s ($\chi^2 = 9.6$) for Moho refractions (P_n phases). The overall RMS value for 3880 picks is 0.37 s ($\chi^2 = 3.9$). The abilities of this modelling package are summarized in Table 1.

Table 1

Comparison of the abilities of modelling software packages used in this study

Package (modelling method)	Phases possible to use	Phases used	P-wave velocity distribution	Sedimentary layer taken into account	Determination of the crustal and Moho discontinuities	Synthetic seismograms
FAST (tomographic inversion)	first arrivals	P_g, P_n	+ (+)	–	– –	–
JIVE3D (tomographic inversion)	all	P_g, P_MP, P_n, P_{ov}	+ (+)	+ (–)	+ (–) + (+)	–
RayInvr (forward and inverse modelling)	all	P_g, P_MP, P_n	+ (+)	+ (+)	+ (–) + (+)	+ (–)
SEIS83 (forward modelling)	all	P_g, P_cP, P_MP, P_n, P_{ov}	+ (+)	+ (+)	+ (+) + (+)	+ (+)

Notice: Marks in parentheses concern features actually used in this study.

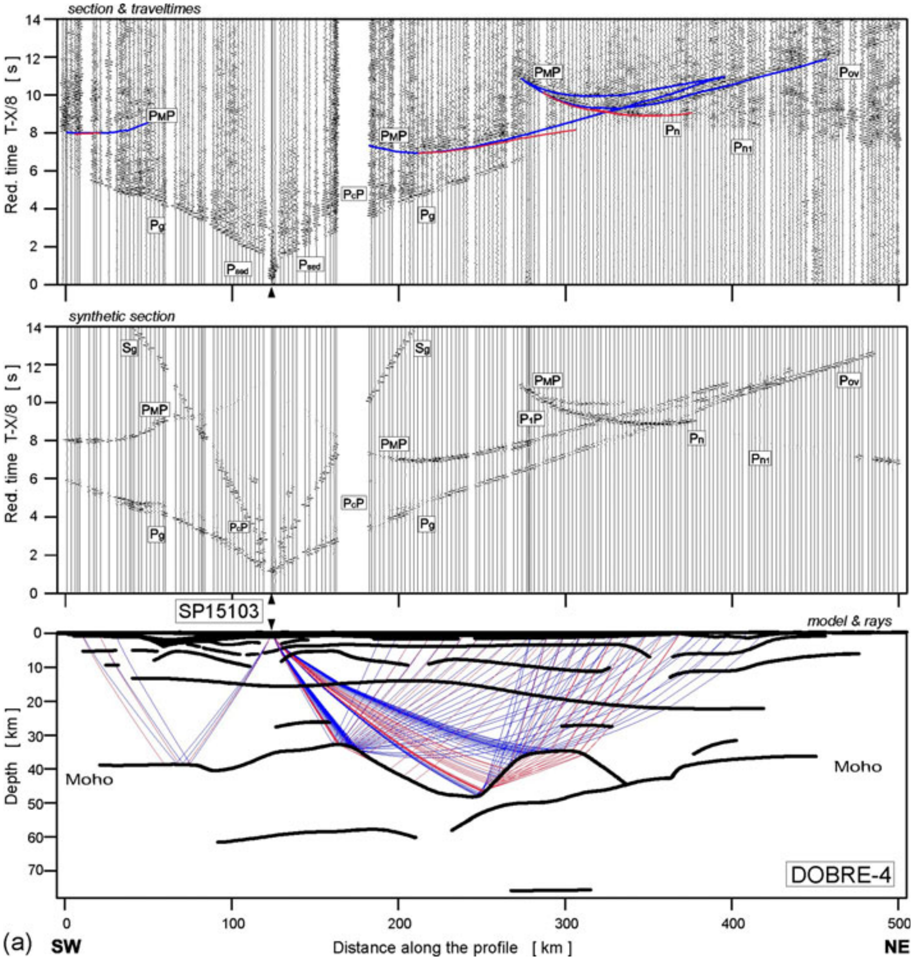


Fig. 6. Caption on next page.

The forward modelling showed that double arrivals observed on the record sections (Figs. 3a and b) represent the reflections from the opposite-dipping segments of a strongly undulating Moho boundary (Starostenko *et al.* 2013). The fragments of the travel time curve with high apparent velocity were reflected by the Moho fragments dipping up with increasing offset, while those with slow velocity were reflected from the Moho segments dipping down. Together, this produces characteristic triplication of the travel time curve. Usually, these phases are of very good quality; therefore, in spite of their unusual character, they could be identified with a high confidence. Examples of modelling P_MP waves for shots 15103, 15105, 15107, and 15110 are presented in Fig. 6.

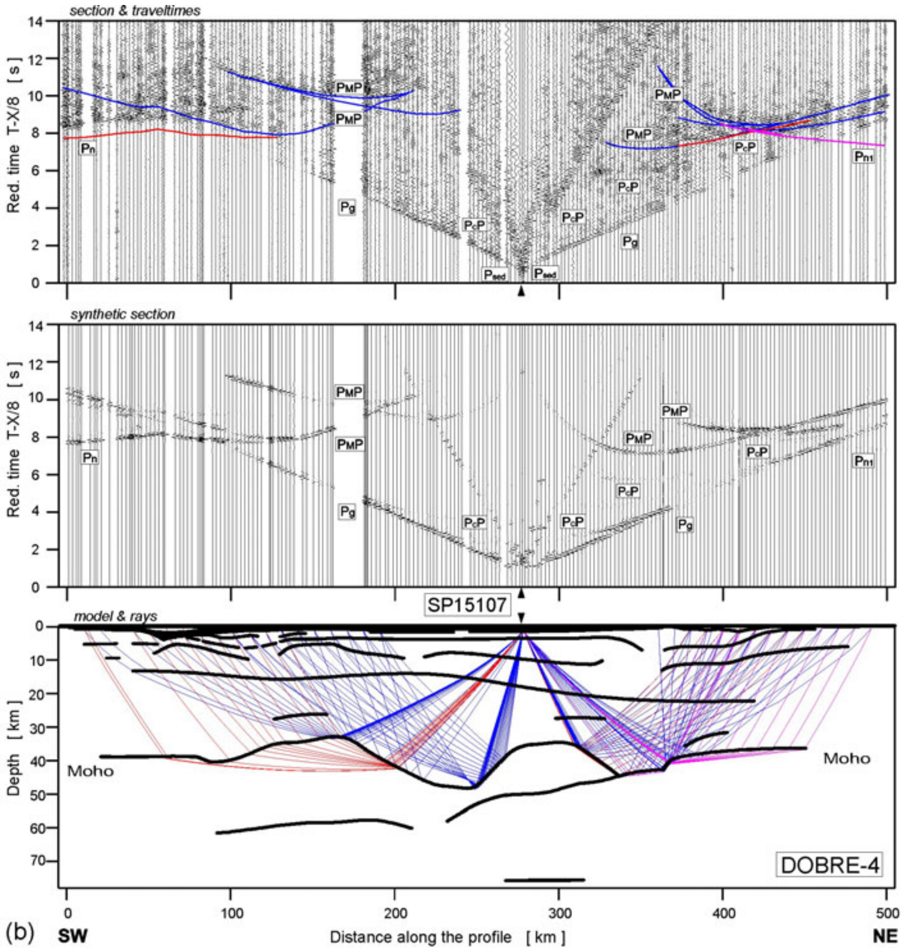


Fig. 6: (a)-(b). Examples of seismic modelling of the P-wave with selected theoretical travel times (P_MP and P_n – blue and red travel times, respectively) calculated using the SEIS83 ray tracing technique for SP15103 and SP15107, respectively. Seismic record sections (amplitude-normalized vertical component, the band-pass filter of 2-12 Hz, reduction velocity of 8.0 km/s). Synthetic seismograms (middle diagram) and the ray diagram of selected rays using the SEIS83 (bottom diagram). All examples were calculated for the model presented in Fig. 5d. P_{n1} – refraction in the higher velocity upper mantle. Other abbreviations are as in Fig. 3.

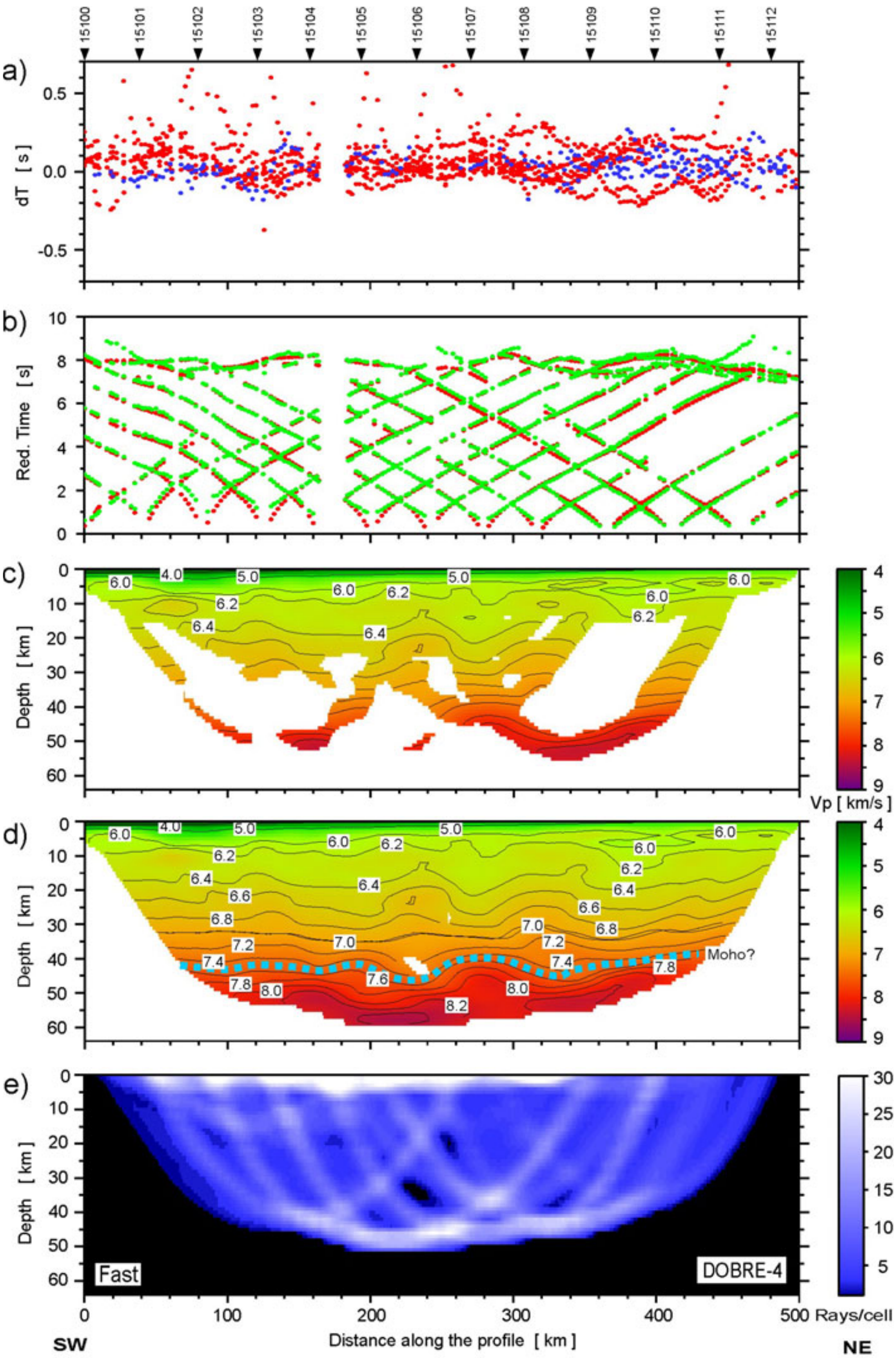
3.2 Refraction travel time tomography with FAST

We have used the First Arrival Seismic Tomography (FAST) program (Zelt and Barton 1998) to invert first arrivals of refracted P_{sed}, P_g, and P_n phases. The velocity model is parameterized on a rectangular equidistant grid. FAST uses an eikonal solver in the forward step to produce travel times and ray paths for an initial velocity model. In the inverse step, regularized inversion based on LSQR variant of conjugate gradient technique is used to obtain velocity perturbations from the travel time residuals. The method allows the user to constrain the flatness and smoothness of the slowness perturbations. Velocity calculations are organized in iterative form in order to minimize the root mean square residual difference between the calculated and observed travel times. This procedure allows the nonlinear problem to be linearized and to solve the linear system in several iterations and is a common way to get the Earth's interior velocity based on travel times of the first seismic arrivals on some general information to build an initial model. Similarly to all travel time tomographic codes, the drawback of FAST is a smoothing of (potentially existing) velocity contrasts in the resulting model which is characteristic of all tomographic inversion methods. Any iterative inversion involves for the smoothing of results to ensure the stability of the whole process and also FAST uses the velocity parameterization on a continuous grid without the existence of velocity discontinuities representing geological boundaries or fault zones.

Another issue in case of most inversion methods is the influence of the initial model on the solution. The final model is strongly dependent on the initial one, and we were faced with this problem too. We have got very different resulting models for the number of initial ones, which may suggest that the data (P_g and P_n phase alone) do not have good resolving capability. A formal criterion to stop the iterative inversion process is reaching $\chi^2 \sim 1$ (meaning that differences between the observed and estimated travel times are comparable to data uncertainty, thus the data fit is satisfactory). In order to obtain a model with an acceptable fit and simplest velocity structure possible, a smoothing of the velocity field is usually applied.

However, in a number of cases, applying several inversion iterations, besides decreasing χ^2, leads also to the appearance of small-scale velocity anomalies, which are most likely artefacts due to uneven and anisotropic ray coverage (directions of rays propagating trough given model cell are not distributed uniformly).

When working on this model, it was found that the occurrence of these artefacts for some acceptable χ^2 (for instance less than 2) depends on the proximity of the initial and final model. In other words, it is easier to get the final model without obvious artefacts and satisfactory fit when the initial model already shows a reasonable travel time fit to the data. Thus, it is important to properly select the initial model. Figure 7 represents the velocity distribution calculated using the FAST package. The model size was 500×72 km and grid spacing was 1×1 km for the forward calculations of the rays and 5×2 km for inverse computations. The code allows for taking


15100
15101
15102
15103
15104
15105
15106
15107
15108
15109
15110
15111
15112
a)
dT [s]
b)
Red. Time [s]
c)
Depth [km]
Vp [km/s]
d)
Moho?
e)
Fast
DOBRE-4
Rays/cell
Distance along the profile [km]
SW
NE

the surface topography into account by positioning the source and receiver at their true depth. However, as the elevation along the DOBRE-4 profile did not exceed 160 m, a constant elevation (0 m a.s.l.) of sources and receiver was assumed for the modelling. The data set included 1139 P_{sed} + P_g picks and 310 P_n picks determined for 13 shot points along the profile. To build the initial model, first a one-dimensional velocity curve was flattened and smoothed to provide the smallest RMS travel time misfit of the starting model. As the geological/geophysical cross section (Fig. 4) shows varying velocities and thickness of the sediments along the profile, next, the initial (2D) model was prepared by a linear interpolation between previously prepared one-dimensional curves modified in the upper part (0-5 km) and located at the model edges, to reflect these velocity/depth variations (~3.5-5.5 km/s in the left edge and 5.5-5.9 km/s in the right, at 0-3 km depth) and to build a more realistic initial model. This initial model provided an acceptable RMS of 0.22 s ($\chi^2 = 5$) at the first iteration and ensured the successful determination of the final one.

The final model was calculated in 9 iterations, using initial regularization strength (lambda 0) of 30, and vertical to horizontal model roughness (*Sz*) value 0.2. The χ^2 for final model was 1.3 (RMS $dt = 0.11$ s). The abilities of this modelling package is summarized in Table 1.

In Fig. 7a, the diagram of residuals show few positive residuals (mainly in the area with substantial thickness of the sediments – distances 0-250 km along the profile) reaching 0.5-0.7 s, more than the overall RMS residual. It means that modeled velocities in the uppermost part of the sedimentary layer are too high compared to real values. It is most likely due to a relatively large vertical spacing of the inverse grid (2 km) and regularization (smoothing) applied to the model, which prevented the precise recovering of low (< 5 km/s) velocities in a thin (< 2 km) layer of subsurface sediments.

Tomographic inversion, using a single velocity grid and smoothing, will image existing velocity discontinuities as zones of an elevated velocity gradient, rather than as sharp velocity contrasts. This is also the case of the Moho discontinuity in the presented model. In order to approximately locate the Moho, we assume that the gradient zone at the lower crustal/upper man-

Fig. 7. The P-wave velocity model obtained from tomographic inversion of the first arrivals (P_g and P_n phases) using the FAST package: (a) Travel time residuals (red – P_g phase, blue – P_n phase); (b) Observed (green) and calculated (red) travel times; (c) *Vp* velocity model based on the P_g phase; (d) *Vp* velocity model based on the P_g and P_n phases. The 7.5 km/s velocity isoline, considered to approximately represent the location of the Moho discontinuity in a model with smooth velocity distribution, is marked by a blue dotted line; (e) Ray density.

tle depth in the model represents a smoothed image of the velocity increase from typical lower crustal (< 7 km/s) to typical upper mantle (> 8 km/s) values. Thus we tentatively locate the Moho boundary at the 7.5 km/s velocity isoline, representing the average of these values (Fig. 6), not at the isoline of velocity expected for the uppermost mantle (8.15-8.4 km/s).

3.3 Joint refraction and reflection travel time tomography with JIVE3D

The tomographic code JIVE3D was used to test the ability of inversion in case of a complex Moho shape. Methods based just on first arrivals (*e.g.*, Hole 1992, Zelt and Barton 1998) use a part of the available travel time data only, without other (reflected) arrivals. The JIVE3D 3D tomographic software package (Hobro 1999, Hobro *et al.* 2003) is a suitable solution as it allows for the use of later refracted phases as well as reflected arrivals, apart from refracted first arrivals, to build a layered model.

The JIVE3D code is based on the regularized least squares inversion approach. The model is defined as a stack of layers separated by interfaces which can represent velocity discontinuities. Both refracted and reflected arrivals can be used. Models are parameterized by regular grids of velocity and depth nodes, from which the interface surfaces and velocity fields for individual layers are B-spline interpolated. The ray theory and ray perturbation method are used for travel time calculations. The inversion is performed using the iterative regularized least-squares method.

The JIVE3D package allows every layer and interface to be modelled separately or jointly. The usual approach is to model the upper parts of the crust first, and then successively resolve the deeper layers. The shallower, previously modelled layers can be fixed during the following inversion steps, in order to focus the algorithm on the deeper layer only. This reduces the number of model parameters and stabilizes the inversion procedure. Several runs of inversion (loops) with different regularization strengths are iteratively calculated. Thus, initially, a smooth solution is obtained and subsequently, the smoothness of the solution is gradually decreased in order to recover the smaller-scale details of the structure. Based on initial tests, appropriate regularization strengths (according to the terminology of Hobro (1999) and Hobro *et al.* (2003)) and a number of iterations in each loop (inversion run) were determined. The output model from one step is used as an input model for the next step of the inversion. Calculations are usually stopped when the value of χ^2 stabilizes. The 1D initial input model was defined based on the 2D model from Starostenko *et al.* (2013) and some inversion tests. The code allows taking the surface topography into account. However, in this case a constant elevation (0 m a.s.l.) of sources and receivers was assumed in the model. The grids cell size was defined as 5×0.5 km for the crust and 10×0.5 km for the upper mantle, node spacing for the Moho interface was set to 5 km.

The sequence of modelling consisted of the following steps. First, refracted waves in the crust (P_{sed}, P_g, P_{ov}) were modelled. P_{sed} branches were incorporated into the crustal layer. Modelling of a separate sedimentary layer increased the calculation time about twice, because of the relatively low density of rays compared to the vertical size of the layer. At the 6th loop with 6th iteration, the RMS value of 0.19 s ($\chi^2 = 3.42$) was reached and stabilized. The hit rate (percentage of receiver locations where a ray was traced successfully from the source) was 98%. This model was used for the next step of the modelling as the starting model. In this step, a joint inversion of crustal velocities and Moho boundary based on waves refracted in the crust and reflected at the Moho interface was done (P_{sed}, P_g, P_{ov}, P_MP). We have limited Moho reflections to the first Moho arrivals. Unfortunately, JIVE3D software does not allow for the inversion of two phases from the same interface. At the 6th loop and 8th iteration the RMS of 0.18 s ($\chi^2 = 2.66$) was reached and stabilized. The hit rate was 94%. This model was used as the starting model for the last step of the modelling. We used waves refracted in the upper mantle (P_n) in this step of inversion. At the 4th loop and 12th iteration RMS of 0.16 s ($\chi^2 = 2.42$) was reached and stabilized. The hit rate was 89%.

All the models are quite well determined. The tomographic model, ray density together with a comparison of experimental and calculated travel times and residuals are presented in Fig. 8. The residuals are largely limited to ±0.15 s. The abilities of this modelling package are summarized in Table 1.

The best ray coverage is obtained in the upper crust down to 20 km depth (Fig. 8). The best ray coverage in the lower crust is obtained at distances of 140-225 km and 260-330 km. The Moho boundary is better determined at distances of 140-400 km except for the segment in the vicinity of 210 km. The best information about the upper mantle is obtained at distances of 150-370 km, apart from the segment in the vicinity of 210 km.

Comparing the tomographic inversion model to the forward ray tracing model described in Starostenko *et al.* (2013), we can see a similarity of the main features of the velocity field. Instead of some crustal discontinuities there are undulating velocity isolines at distances of about 100, 230, and 400 km. They coincide with a similar shape of the upper crustal layer in the model by Starostenko *et al.* (2013), but they reach the lower crust, probably due to smearing of the solution to the lower crust, not constrained by refracted waves.

The Moho shape is similar to this modelled by Starostenko *et al.* (2013), but the amplitude of undulations is lower. The undulations of the crustal velocity field roughly follow the Moho shape. There is an area of lower mantle velocity in the vicinity of 220 km of the model, which could indicate that the

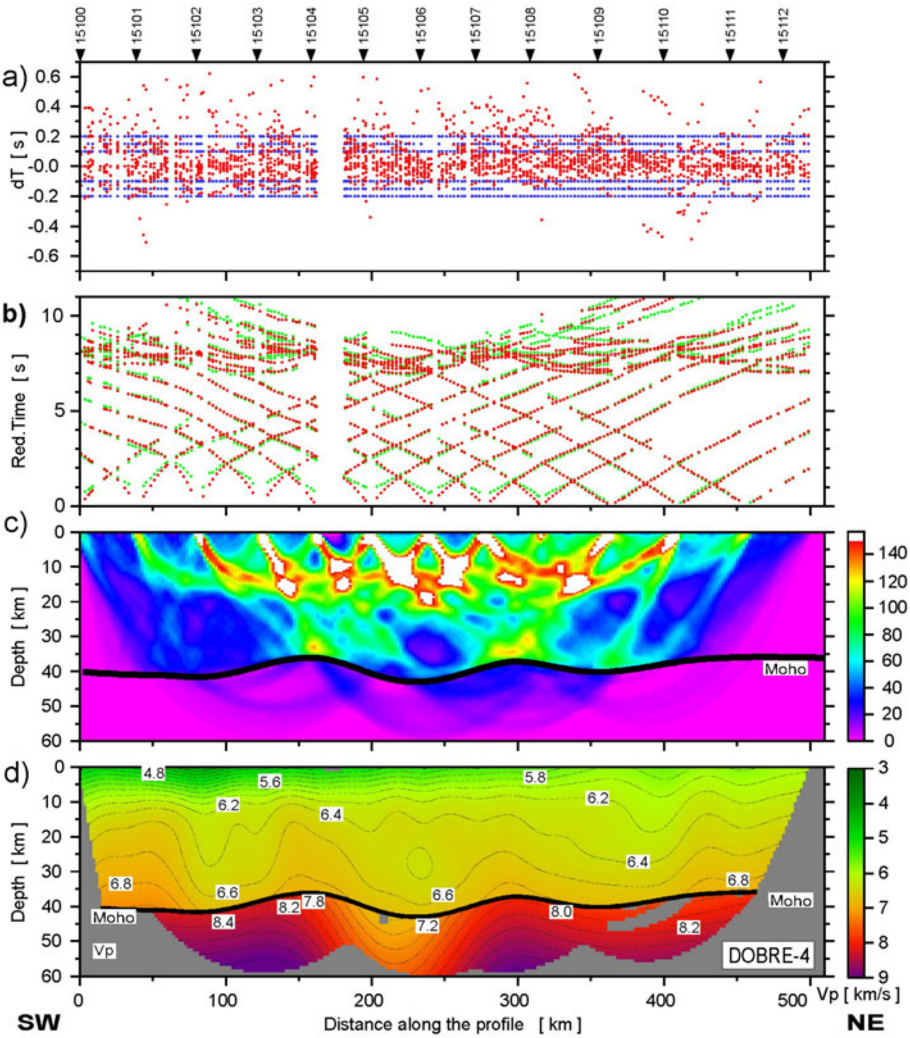


Fig. 8. The result of the JIVE3D modelling (P_{sed}, P_g, P_{ov}, P_MP and P_n phases): (a) Residuals, blue dots – error limits (picking errors: 0.1 s for P_g, 0.15 s for P_n, and 0.2 s for P_MP), red dots – residuals; (b) Travel time fit, red dots – calculated travel times, green dots – experimental travel times; (c) Ray density. Colours represent the number of rays crossing a cell; (d) The final model. Model limited to the ray coverage area. Colours represent the P-wave velocity distribution, black thin lines – velocity isolines, black thick line – Moho seismic boundary.

inversion algorithm attempted to compensate for the insufficient deepening of the Moho at this distance. Thus, it could indicate that the true Moho undulation amplitude could be higher than that obtained by the modelling. Such a velocity undulation in the mantle was not observed by Starostenko *et al.* (2013) because the Moho interface modelled there is determined using all the Moho reflected branches, which is not the case for JIVE3D.

3.4 Travel time inversion with RAYINVR

Another variant of the solution was produced using the forward and inversion code RAYINVR (Zelt and Smith 1992). This code calculates ray paths, travel times and amplitudes of P- and S-waves using ray tracing by a numerical solution of the eikonal equation for 2D media with an inhomogeneous velocity distribution. The model can consist of layers separated by velocity discontinuities. The velocity is parameterized in each layer at irregularly located nodes and linearly interpolated between the nodes to define the velocity in each point of the model. Calculated travel times can be used as a basis for trial-and-error forward modelling or for regularized inversion to find the velocity distribution that fits the seismic data. The code allows for taking the surface topography into account. However, as the elevation in the study area did not exceed 160 m, a constant elevation (0 m a.s.l.) of sources and receivers was assumed in the model.

In this study, we attempted to use the regularized inversion with the DMPLSTSQR code, available within the RAYINVR package, in order to obtain a layered model that could be compared with results of the trial-and-error forward modelling. The model (Fig. 9) is composed of three layers – sediments, consolidated crust, and uppermost mantle. The velocities in the sediments were obtained from previous geophysical data and from the SEIS83 model by simplifying the detailed and complex velocity distribution into one gradient layer. The velocity in the sedimentary layer has been kept constant during the inversion.

As the input data, the first arrivals, P_{sed} (refraction from the sediments with apparent velocity < 5.7 km/s) and P_g (refraction from the consolidated crust with apparent velocity ≥ 5.7 km/s) were used to constrain the depth of the sediments/consolidated crust boundary and the velocity distribution in the consolidated crust. The reflected phases from the Moho discontinuity (P_MP) as well as sub-Moho refractions (P_n) served for modelling the depth of the Moho and the velocity in the uppermost mantle. The initial, 1D model of the velocity in the consolidated crust was obtained by a one-dimensional inversion of the averaged curve of all observed P_g phase travel times. The shape of the bottom boundary of the sediments and the distribution of the *Vp* velocity in the crust was modelled by an inversion of the travel times of the P_g phase. After three iterations, the P_g RMS travel time residual was 0.09 s (χ^2 = 0.8) (with 939 rays traced out of 1070 total travel times, 88% hit rate). The P_{sed} phase was not used for an inversion of the velocity distribution in

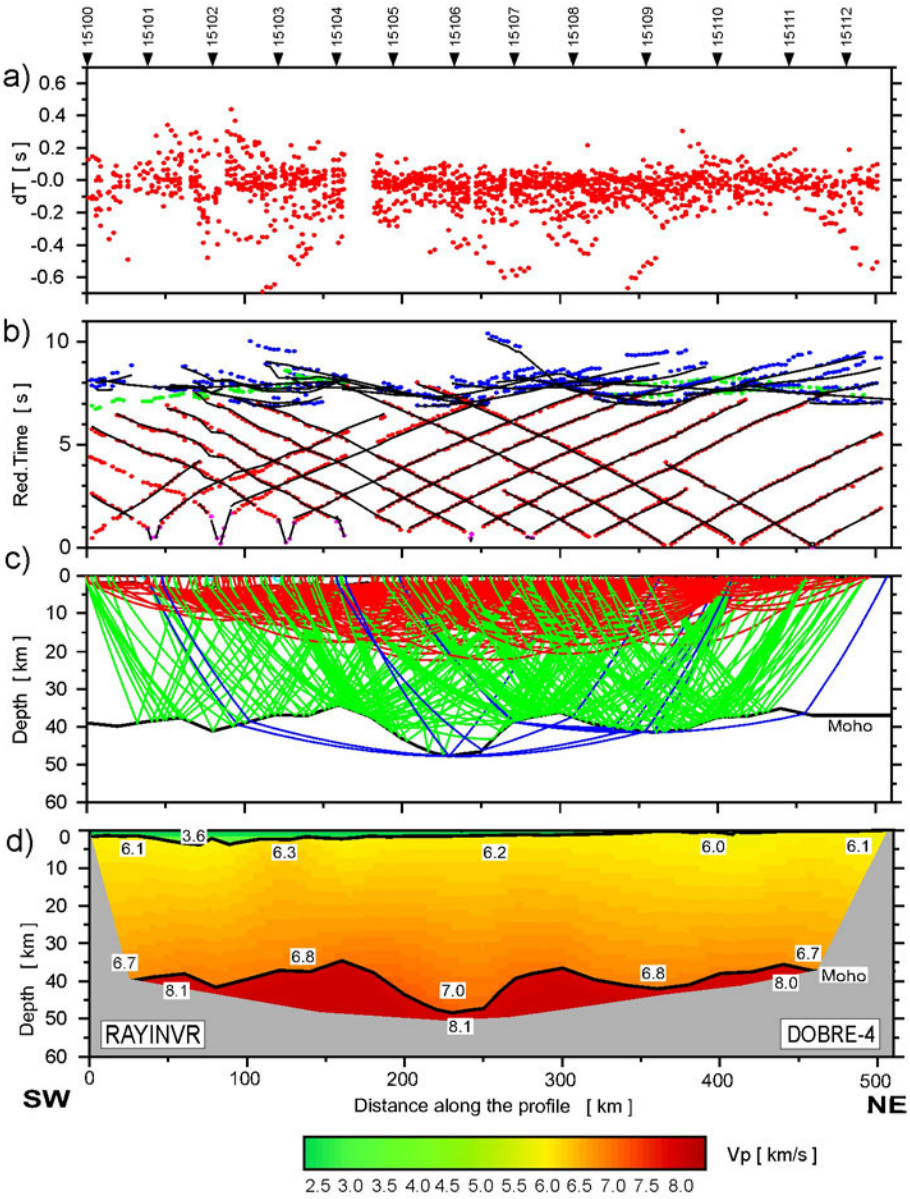


Fig. 9. P-wave velocity model obtained from inversion of the P_g, P_MP, and P_n phases using the RAYINVR package: (a) Residuals; (b) Travel time fit, black lines – calculated travel times, dots – observed travel times for P_g (red), P_MP (blue), and P_n (green) phase; (c) Ray diagram; (d) The final model limited to the ray coverage area. Colours represent the P-wave velocity distribution.

the sediments, but P_{sed} RMS residuals calculated for *Vp* assumed (and kept fixed) in this layer are low enough (0.11 s, $\chi^2 = 1.2$) to prove that the assumed velocity is realistic and does not bias the results from the deeper layers. The resulting crustal model shows an increasing thickness of the sedimentary layer – from 0 km in the north to 2-4 km in the south, consistently with the SEIS83 model and with tomographic models. In deeper parts, a slight lateral differentiation of the crustal velocity is observed, especially in the lower crust where high (~7.1 km/s) velocities are observed in the central part, compared to 6.7 km/s in the north and south.

This crustal model served as a basis for modelling the Moho topography and the velocities in the uppermost mantle. For this, a simultaneous inversion of the P_MP and P_n travel times was performed (keeping the crustal velocities constant). The P_MP phase in some seismic sections was observed as double (overlapping) phases with different apparent velocities (Figs. 3 and 6) (most likely resulting from abrupt changes of the Moho topography – reflections from Moho fragments with an opposite dip). This caused technical difficulties in using all available data for the inversion. The RAYINVR code is able to trace rays belonging to both overlapping phases in the forward step; however, during the inversion in locations where double arrivals of the phase are recorded, only the earliest calculated arrival time is taken for calculating the travel time residual, used subsequently in the inversion procedure to calculate the corrections/update of the velocity model. Therefore, when double P_MP phases were observed, only the fragments with first-arriving travel times were used (Fig. 10). The five steps of the inversion of the P_MP and P_n phase resulted in a model with a RMS travel time residual of 0.13 s ($\chi^2 = 0.42$) for the P_MP phase (with 550 rays traced for total 588 travel times, 93% hit rate) and 0.20 s ($\chi^2 = 4$) for the P_n phase (with 41 rays traced for 160 observed travel times). The poor hit rate for the P_n phase is due to "shadow zones" formed due to strong variations of the modelled Moho topography, which is inherent for most modelling methods based on ray approximation.

During this step, the crustal velocity was not included in the inversion. Nevertheless, as the velocities in this layer were interpolated between *Vp* values defined at nodes located along the layer boundaries, modification of the Moho interface resulted in a relocation of the corresponding velocity nodes and effectively affected the interpolated crustal velocity distribution. A modified crustal velocity field resulted in an increase of the P_g RMS residuals to 0.11 s ($\chi^2 = 1.2$). Therefore, 4 iterations of the inversion of the P_g travel times were applied, which allowed for a decrease in the residuals to0.09 s but increased the P_MP residuals to 0.16 s ($\chi^2 = 0.64$). Finally, an inversion of the P_MP/P_n travel times decreased the corresponding residuals to 0.13/0.20 s ($\chi^2 = 0.42/\chi^2 = 4$), respectively. The value of the P_g residuals for

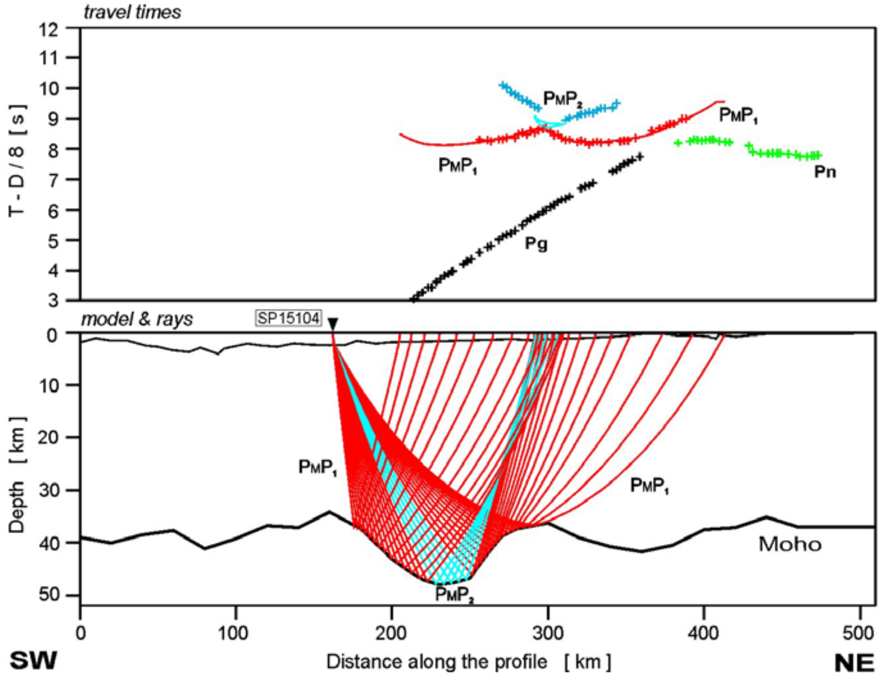


Fig. 10. Example of RAYINVR modelling of the overlapping P_MP phase as a reflection from opposite dipping slopes of a Moho trough. Top – observed (crosses) and calculated (lines) travel times. Bottom – a model with calculated rays reflected from Moho. Red – rays and travel times of the P_MP phase in earlier arrivals (used for inversion), Light blue – rays and travel times of P_MP in later arrivals (not used for inversion).

the final model was 0.1 s ($\chi^2 = 1$). The abilities of this modelling package are summarized in Table 1.

The resulting model shows the shape of the crust/mantle boundary very similar to the SEIS83 model, although with a smaller amplitude of the Moho trough at ~300 km distance and with some differences at the edges of the model, not constrained by the data. This shows that using only first-arriving fragments of the overlapping P_MP phases for inversion gives a satisfactory result. However, it is likely that using all available travel times of the double P_MP phase would result in a more precise delineation of the Moho depth.

4. DISCUSSION

The unique Moho structure found along the DOBRE-4 profile provides valuable material to test different approaches of 2D wide-angle modelling. For these tests, SEIS83, FAST, JIVE3D and RAYINVR codes were used.

The main objective of the comparison was to check to what extent the complex geometry of the Moho boundary will be recovered by tested software packages. The largest Moho depth change detected in previous study is from 32 km at a distance of 165 km, down to 48 km at a distance of 250 km (Starostenko *et al*. 2013).

Although all the software packages allow the Earth's surface topography to be included, we did not take this into account, exception the SEIS83 model. The maximum terrain elevation is less than 160 m, so the topography could be ignored.

The sedimentary sequences show a substantially lower velocity than the underlying basement; therefore, it is advisable to model them as a separate layer. This was possible in the case of the SEIS83, JIVE3D, and RAYINVR models. In the case of the FAST package, the model is parameterized as a single layer and therefore sediments are represented by a low velocity/high gradient area in the first few kilometers depth. In the JIVE3D model, after some tests, the sediments were also incorporated into the crustal layer due to a huge increase of the calculation time in case of a separate sedimentary layer. The time consumption necessary for the whole modelling procedure is also an important matter. The FAST package is the fastest way to have the first velocity model, although seismic boundaries cannot be represented by first order velocity discontinuities. Also, only refracted phases can be used.

The RAYINVR is also very fast in terms of computation time. However, the number of model parameters is typically orders of magnitude smaller than in the case of other tomographic codes. On the one hand, it allows for interactive, arbitrary (manual, if needed) selection of model parameters (boundary and velocity nodes) used for inversion, what makes it very flexible. On the other hand, sparse model parameterization results in a relatively low resolution. Moreover, selection of parameters for inversion requires interactive user input and testing of several variants of the solution, which makes the procedure time consuming. RAYINVR allows for using all refracted and reflected phases. The JIVE3D package is more time consuming but it can use P_g, P_{ov} (refractions in later arrivals) and reflected phases.

The modelling with SEIS83 package is the most time-consuming. The code can use all refracted and reflected phases, including multiples or converted waves. The SEIS83 code solves the forward problem only (ray path and travel time calculation for a given velocity model). Unlike other discussed codes, it does not solve the inverse problem because it does not calculate the Frechet matrix to derive the model perturbations.

The calculations with SEIS83 are fairly fast, similarly to the RAYINVR code. However, manual, trial-and-error process of seeking the model fitting the data can be very time-consuming and usually takes large part of total modelling time. Such a procedure allows for better control of the modelling process, which is an advantage, but in the same time introduces some subjectivity into the final solution. Another advantage is that during modelling, it allows to modify the identification of picked phases. For example, problems with fitting some particular phase by a model fitting all the remaining data may indicate that such a phase actually originates as, *e.g.*, side reflection or diffraction from some local structure, and thus should not be included in 2D modelling. This helps to eliminate artefacts that could be introduced by the modelling of some incorrectly interpreted phases, whose nature may not be obvious at first glance.

Forward modelling is a good tool for checking the phase identification, which is a key for obtaining a correct model. Also, it allows to build a detailed model using most of the information contained in the record sections.

Packages SEIS83 and RAYINVR allow to calculate synthetic seismograms and compare them with the experimental data.

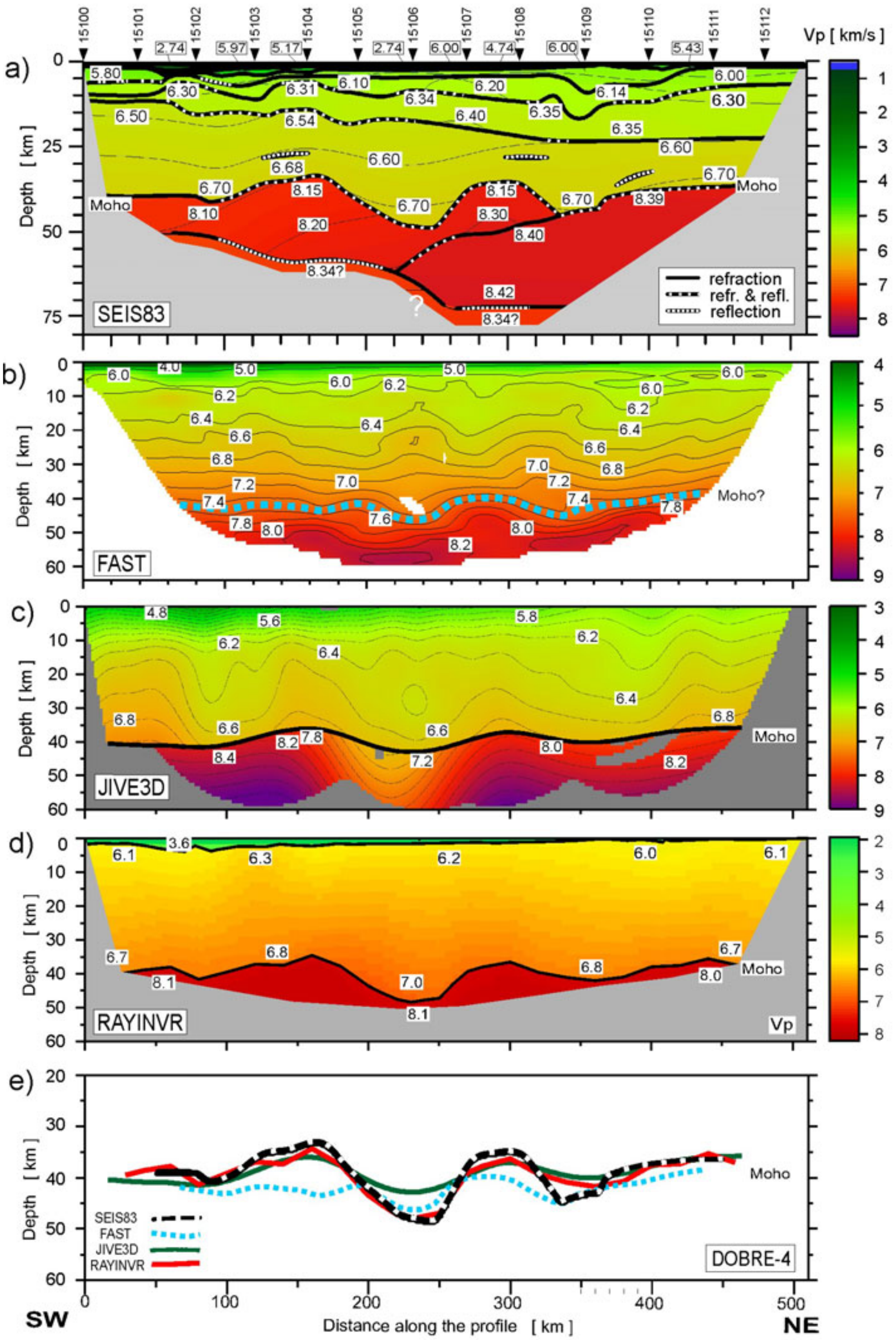
Modelling of double (overlapping) phases in seismic sections, as observed in DOBRE-4 data, is problematic. SEIS83 can do it, RAYINVR can calculate rays and travel times for overlapping reflections in forward step, but in inverse step, if two arrivals of the same phase at the same source-receiver offset are calculated, only the earlier one is used for computing the residual and for inversion (see Fig. 10). JIVE3D cannot use overlapping reflections for inversion. To use whole branches of overlapping reflections for inversion, modifications of the RAYINVR or JIVE3D would be necessary.

2D tomographic codes are fast in terms of calculation time. However, finding a "realistic" solution requires several tests to find optimum parameters for inversion, including parameterization of the model (*e.g.*, grid spacing, amount of regularization/smoothing, number of iterations) and the choice of the initial model. This can substantially increase the total modelling time. Nevertheless, inversion is still much faster than forward trial-and-error methods, as repetitive manual modification to fit the data for several shotpoints simultaneously is very complex and time-consuming. Moreover, the final result is to some extent subjective, as it depends on the experience, knowledge, and preferences of the interpreter. Also, this subjectivity makes it hard to quantitatively estimate the uncertainty or resolution of the model.

Figure 11 presents a comparison of the Moho shape derived by different approaches. All lines show a substantial variation of the Moho depth. The

---

Fig. 11. Comparison of the 2D seismic velocity models counted along the DOBRE-4 profile: (a) SEIS83 – the trial-and-error, ray tracing modelling; (b) FAST inversion; (c) JIVE3D – tomographic modelling; (d) RAYINVR inversion; (e) Comparison of the Moho boundaries obtained by using different approaches. Vertical exaggeration is ~3.6:1 instead of ~2.4:1, as used in other diagrams.


a)
SEIS83
Vp [km/s]
Depth [km]
Moho
refraction
refr. & refl.
reflection
b)
FAST
Moho?
c)
JIVE3D
Moho
d)
RAYINVR
Vp
e)
SEIS83
FAST
JIVE3D
RAYINVR
DOBRE-4
SW
NE
Distance along the profile [km]

largest change is from 32 km at a distance of 165 km down to 48 km at a distance of 250 km in the SEIS83 model. The line obtained by RAYINVR has a very similar shape (deviation from the SEIS83 Moho < 3 km), with the exception of a much less pronounced depression at km 340, where the Moho depth is ~40 km instead of 46 km in the SEIS83 model. The JIVE3D modelled the Moho depth similarly to SEIS83 and RAYINVR, but all depressions (km 250 and km 340) are smoothed (depth smaller by 7-8 km). Also the FAST line, which is based only on P_g and P_n phases, displays differences at a depth of the approximately determined Moho. In the points of maximum depth, the FAST Moho is deeper by 12 and 7 km, respectively.

The Moho depth in the presented models fits with the 2.5-5 km accuracy the methods using P_MP phases. The biggest differences are observed at Moho uplifts and depressions. Better fit occurs for the Moho uplifts, with differences not exceeding 2.5 km. The JIVE3D Moho is shallower in the depressions, with difference up to 5 km. In the FAST model, the Moho boundary (approximately delineated along 7.5 km/s velocity isoline) differs from other models by up to 10 km in depth. The conclusion is that using three methods (SEIS83, RAYINVR, and JIVE3D) to similar phases, we can obtain models with differences of 2.5-5 km in the Moho depth. Inversion of first arrivals only (FAST) gives much bigger differences, in our case *ca*. two times.

It is difficult to compare the quality and accuracy of the discussed models. The modelling procedure was different and different phases were used. For the FAST and RAYINVR, P_g waves were used in one inversion step, while for the JIVE3D, P_g (even P_{ov}) and P_MP waves were modelled simultaneously. Table 1 presents seismic phases used by respective methods. For all the resulting models we can compare the fit of the solution to the data used for modelling (Figs. 5, 7, 8 and 9). It should be noted that this is only a measure of how far the synthetic data are from the observed data (travel time residual), not how far the final model is from the real structure (model uncertainty). Tomographic inversion methods, like FAST, RAYINVR, and JIVE3D, give a better fit, especially for P_n waves. On the other hand, with these methods, it was not possible, or difficult, to use all the phases that were used for forward modelling with SEIS83 (Table 1).

In the case of a dataset with a complex wavefield, as presented here, some uncommon phases are hard to identify reliably at first glance, *e.g.*, the overlapping Moho reflections modelled in this work. This poses a problem for any seismic inversion modelling, as it requires correct identification of the observed seismic phases (wave type and layer). In this case, the procedure of phase identification, ray tracing and inversion requires few "iterations", involving, if needed, re-identification of phase and subsequent modelling. For this, forward modelling is helpful, as it allows for verification of the correspondence between calculated rays/phases and observed phases.

After a reliable identification confirmed by ray tracing, the data can be used for inversion modelling.

Another problem was encountered during modelling of the velocities in the upper mantle. For the DOBRE-4 profile, the topography of the Moho discontinuity produced the shadow zones for mantle phases due to the nature of the ray method. Therefore, for the modelling of the upper-mantle velocity distribution, two concurrent approaches were applied. The first one was based only on the ray theoretical algorithm, and the other approach used the full-waveform FD calculation to overcome the limitations of the ray method. This was described in the previous paper of Starostenko *et al.* (2013). Similarly, shadow zones for the P_n were also formed during inversion modelling, which substantially decreased the hit rate for this phase. The effect of shadow zones, inherent for the ray-tracing algorithms used here, could be eliminated by application of the codes based on the graph (shortest path) method, described by Moser (1991) and successfully implemented for tomographic inversion by, amongst others, Korenaga *et al.* (2000) and Meléndez *et al.* (2015).

5. CONCLUSIONS

Modelling with all the codes tested showed substantial variability of the Moho depth along the DOBRE-4 profile. However, SEIS83 and RAYINVR packages seem to give the most coincident results. In the case of the FAST package, modelling of the Moho topography is problematic for two reasons. First, the model parameterization does not allow us to define velocity discontinuities which are represented in the model as zones of increased velocity gradient. Second – information about Moho is based on P_n modelling only, as the reflected phases (most important for the modelling of Moho) are not used for inversion The JIVE3D model seems to be an intermediate model.

Not all the software packages can calculate double (overlapping) P_MP phases observed in seismic sections. This is possible with SEIS83; also RAYINVR could calculate rays and travel times for overlapping reflections in a forward step, but in the inverse step, if two arrivals of the same phase are calculated, only the earlier one is used for computing the residual and for inversion. JIVE3D cannot use overlapping reflections for inversion. Results of modelling of the DOBRE-4 dataset show that it would be interesting to introduce modifications to the RAYINVR or JIVE3D code to include all the travel times of overlapping reflected phases in inversion.

A c k n o w l e d g e m e n t s. The public domain GMT software (Wessel and Smith 1991, 1998) was used to produce most of the figures. This work was partially supported within statutory activities No. 3841/E-41/S/2015 of the

Ministry of Science and Higher Education of Poland and partially been financed from the funds of the Leading National Research Centre (KNOW) received by the Centre for Polar Studies for the period 2014-2018. The authors are grateful to two anonymous reviewers for helpful comments.

References

Červený, V., and I. Pšenčík (1984), SEIS83 – Numerical modeling of seismic wave fields in 2-D laterally varying layered structures by the ray method. **In:** E.R. Engdal (ed.), *Documentation of Earthquake Algorithms*, Rep. SE-35, World Data Center A for Solid Earth Geophysics, Boulder, USA, 36-40.

Grad, M., and A.A. Tripolsky (1995), Crustal structure from P and S seismic waves and petrological models of the Ukrainian shield, *Tectonophysics* **250**, 89-112.

Guterch, A., M. Grad, R. Materzok, and E. Perchuć (1986), Deep structure of the Earth's crust in the contact zone of the palaeozoic and precambrian platforms in Poland (Tornquist–Teisseyre zone), *Tectonophysics* **128**, 251-279.

Hobro, J.W.D. (1999), Three-dimensional tomographic inversion of combined reflection and refraction seismic travel-time data, Ph.D. Thesis, Department of Earth Sciences, University of Cambridge, Cambridge.

Hobro, J.W.D., S.C. Singh, and T.A. Minshull (2003), Three-dimensional tomographic inversion of combined reflection and refraction seismic travel time data, *Geophys. J. Int.* **152**, 1, 79-93.

Hole, J.A. (1992), Nonlinear high resolution three-dimensional seismic travel time tomography, *J. Geophys. Res.* **97**, 6553-6562.

Komminaho, K. (1998), Software manual for programs MODEL and XRAYS: A graphical interface for SEIS83 program package, Rep. 20, University of Oulu, Dept. of Geophysics, 31 pp.

Korenaga, J., W.S. Holbrook, G.M. Kent, P.B. Kelemen, R.S. Detrick, H.-C. Larsen, J.R. Hopper, and T. Dahl-Jensen (2000), Crustal structure of the southeast Greenland margin from joint refraction and reflection seismic tomography, *J. Geophys. Res.* **105**, 21591-21614.

Koulakov, I. (2009), LOTOS code for local earthquake tomographic inversion: benchmarks for testing tomographic algorithms, *Bull. Seismol. Soc. Am.* **99**, 1, 194-214, DOI: 10.1785/0120080013.

Malinowski, M. (2013), Models of the Earth's crust from controlled-source seismology – where we stand and where we go? *Acta Geophys.* **61**, 6, 1437-1456, DOI: 10.2478/s11600-013-0156-7.

Meléndez, A., J. Korenaga, V. Sallarès, A. Miniussi, and C.R. Ranero (2015), TOMO3D: 3-D joint refraction and reflection traveltime tomography parallel code for active-source seismic data—synthetic test, *Geophys. J. Int.* **203**, 1, 158-174, DOI: 10.1093/gji/ggv292.

Moser, T.J. (1991), Shortest path calculation of seismic rays, *Geophysics* **56**, 1, 59-67, DOI: 10.1190/1.1442958.

Rawlinson, N., and M. Urvoy (2006), Simultaneous inversion of active and passive source datasets for 3-D seismic structure with application to Tasmania, *Geophys. Res. Lett.* **33**, L24313, DOI: 10.1029/2006GL028105.

Starostenko, V., T. Janik, D. Lysynchuk, P. Środa, W. Czuba, K. Kolomiyets, P. Aleksandrowski, O. Gintov, V. Omelchenko, K. Komminaho, A. Guterch, T. Tiira, D. Gryn, O. Legostaeva, H. Thybo, and A. Tolkunov (2013), Mesozoic(?) lithosphere-scale buckling of the East European Craton in southern Ukraine: DOBRE-4 deep seismic profile, *Geophys. J. Int.* **195**, 2, 740-766, DOI: 10.1093/gji/ggt292.

Wessel, P., and W.H.F. Smith (1991), Free software helps map and display data, *EOS Trans.* **72**, 41, 441-446, DOI: 10.1029/90EO00319.

Wessel, P., and W.H.F. Smith (1998), New, improved version of the Generic Mapping Tools released, *EOS Trans.* **79**, 47, 579, DOI: 10.1029/98EO00426.

Zelt, C.A. (1994), Software package ZPLOT, Bullard Laboratories, University of Cambridge, Cambridge.

Zelt, C.A. (1999), Modelling strategies and model assessment for wide-angle seismic traveltime data, *Geophys. J. Int.* **139**, 183-204.

Zelt, C.A., and P.J. Barton (1998), 3D seismic refraction tomography: a comparison of two methods applied to data from the Faeroe Basin, *J. Geophys. Res.* **103**, 7.187-7.210.

Zelt, C.A., and R.B. Smith (1992), Seismic traveltime inversion for 2D crustal velocity structure, *Geophys. J. Int.* **108**, 1, 16-34.

Received 15 February 2016
Received in revised form 1 July 2016
Accepted 23 August 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2020-2050
DOI: 10.1515/acgeo-2015-0098

Crustal Structure Along Sunda-Banda Arc Transition Zone from Teleseismic Receiver Functions

Syuhada SYUHADA[1,2], Nugroho Dwi HANANTO[3], Chalid I. ABDULLAH[4], Nanang T. PUSPITO[5], Titi ANGGONO[1], and Tedi YUDISTIRA[5]

[1]Graduate Research on Earthquake and Active Tectonics (GREAT), Bandung Institute of Technology (ITB), Bandung, Indonesia; e-mail: syuhada@lipi.go.id
[2]Research Centre for Physics – Indonesian Institute of Sciences (LIPI), Tangerang Selatan, Indonesia
[3]Research Centre for Geotechnology – LIPI, Bandung, Indonesia
[4]Faculty of Earth Sciences and Technology, ITB, Bandung, Indonesia
[5]Faculty of Mining and Petroleum Engineering, ITB, Bandung, Indonesia

Abstract

We analyzed receiver function of teleseismic events recorded at twelve Indonesian-GEOFON (IA-GE) broadband stations using non-linear Neighbourhood Algorithm (NA) inversion and *H-k* stacking methods to estimate crustal thickness, *Vp*/*Vs* ratios and S-wave velocity structure along Sunda-Banda arc transition zone. We observed crustal thickness of 34-37 km in Timor Island, which is consistent with the previous works. The thick crust (> 30 km) is also found beneath Sumba and Flores Islands, which might be related to the arc-continent collision causing the thickened crust. In Timor and Sumba Islands, we observed high *Vp*/*Vs* ratio (> 1.84) with low velocity zone that might be associated with the presence of mafic and ultramafic materials and fluid filled fracture zone. The high *Vp*/*Vs* ratio observed at Sumbawa and Flores volcanic Islands might be an indication of partial melt related to the upwelling of hot asthenosphere material through the subducted slab.

Key words: Receiver function, crustal structure, Sunda-Banda arc transition zone.

1. INTRODUCTION

The Sunda-Banda Arc transition zone has been known as one of the most complicated tectonic setting in the world providing one the best modern examples of the early stages of transition from subduction to arc-continent collision (Fig. 1). It developed as a consequence of the interaction between the Australian lithosphere and the Banda Arc in the Pliocene (Hall and Smyth 2008). The structural evolution of this area is characterized by different tectonic episodes including rifting of Timor, Savu Basin and Sumba evolution. This complicated tectonic setting causes this area to be an ideal place to study the lithologic changes in crustal structures associated with the deformation along the plate boundary. A number of studies have been conducted to elucidate the geodynamic and tectonic evolution of the region (Bowin *et al*. 1980, Milsom and Audley-Charles 1986, Shulgin *et al*. 2009). However, information regarding the deep crustal structure of this region is less studied, which is critical to complete our understanding of the geodynamic evolution of the region. Most of the crustal studies were focused on the small part of the region through gravity modelling, seismic reflection and wide-angle

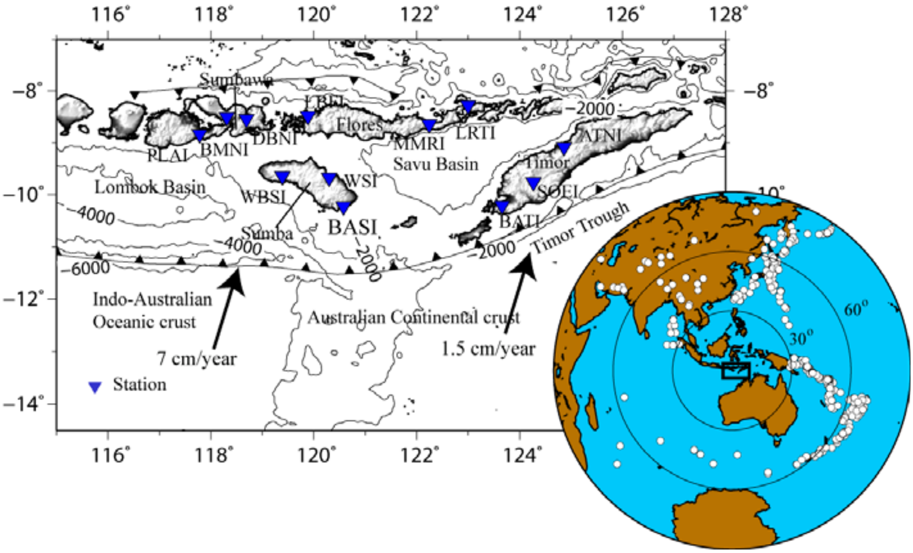


Fig. 1. The tectonic setting of the study area showing the change in tectonic regime from subduction to collision. The inverted blue triangles represent the seismic stations used for the study. The bathymetric contours are marked by the grey lines (Sandwell and Smith 2009). The globe on the right bottom depicts the distribution of teleseismic events used in this study (magnitude greater than 6 and epicentral distance between 30° and 90°). The black solid box represents the study area on the larger map.

refraction studies. For example, gravity modelling has been conducted in Timor Island indicating that the Moho depth in this island ranges from 30 to 40 km (Milsom and Audley-Charles 1986, Bowin *et al.*1980, Kaye 1989). he marine seismic reflection study conducted by Shulgin *et al.* (2009) in the east of Sumba reveals 12-15 thick crust representing the continental character of the Australian lithosphere continuous through the outer Banda arc. The combined analysis of seismic wide-angle refraction, multichannel streamer and gravity data also has been conducted in south of Lombok, suggesting the significant changes in the incoming crustal structure which ranges from 7 km thick of oceanic crust offshore Lombok (Planert *et al.* 2010) to 12 km thick crystalline crust of the Australian continental margin (Shulgin *et al.* 2009).

Receiver function analysis is a method using teleseismic event to extract structure information beneath a station. This technique has been widely used as a powerful tool in characterizing the crustal structure as well as determining the Moho depth in various tectonic environments. For example, Zhu and Kanamori (2000) applied a grid-search method using receiver function data to estimate the Moho depth variation and Poisson's ratio in Southern California, USA. Bannister *et al*. (2004) derived the velocity structure from the inversion of receiver function beneath seismic stations in the volcanically active region, North Island, New Zealand. Ahmed *et al.* (2014) imaged crustal thickness variation from computed receiver function in the eastern Gulf of Aden continental margins.

In recent years, the installation of seismic broadband stations around Banda-Arc transition zone may enable us to study crustal structure through receiver function method for the larger scale coverage of the Sunda-Banda arc transition zone. In this study, we use receiver function analysis to determine the regional variation of the crustal properties (the Moho depths, *Vp*/*Vs* ratios and S-wave velocity models) along the transition zone from subduction to collision using teleseismic receiver function. The study provides new insight into the tectonic development and the processes involved in the transition zone.

2. DATA AND METHODS

We analyzed body wave seismograms obtained from teleseismic events recorded by twelve Indonesian-GEOFON (IA-GE) stations from 2005 to 2014. The twelve seismic stations are distributed in four main islands (Sumbawa, Flores, Sumba and Timor islands) around the Sunda-Banda transition zone reflecting the diverse geologic and tectonic setting (Fig. 1). We selected teleseismic events from ISC catalogue with epicentral distance 30°-90° and magnitude greater than 6 (Fig. 1). This distance range is used to avoid the contamination from regional and core phases and to assure that incoming waves have steep incidence angles. Furthermore, the large moment magnitude selection for the teleseismic events provides waveforms with good signal to noise ratio (Macpherson *et al.* 2012).

The seismograms were pre-processed into the following steps. First, the three component seismograms were inspected manually to select the best quality records as well as to pick the P-arrivals. The seismograms used are baseline-corrected by subtracting the mean and removing the instrument response. The pairs of horizontal components (north-south and east-west components) for each event were then rotated into radial-transverse components. Finally, we selected the time windows started 10 s before and 50 s after the P-wave arrival.

We computed receiver functions in the time domain using the iterative deconvolution technique developed by Ligorria and Ammon (1999) with 500 iterations. This method provides more stable solutions of the receiver functions for noisy data compared to the conventional deconvolution method in frequency domain. We applied the Gaussian filter to control the high frequency noise of the receiver function. After trial and error with different values, we set the Gaussian filter with bandwidths of 1.5. The bandwidth attenuates the frequency contents greater than approximately 0.75 Hz. A least-square misfit criterion was then used as assessment for the quality of receiver function resulted from the iterative deconvolution process. The misfit was computed from the difference between the observed radial seismogram and the calculated receiver function convolved with the observed vertical seismogram. The receiver functions which have more than 90% waveform fit were then used for further analysis.

In total, we obtained 800 individual receiver functions from 354 events which are mostly from N-NE and E-SE directions. The radial and transverse receiver functions were then plotted as a function of the source backazimuth to identify the presence of the complex structure (*e.g.*, dipping interfaces, crustal scattering sources and anisotropic bodies) beneath the stations. The plots of the receiver functions with respect to the event backazimuth for each station are shown in Figs. 2, 3 and a series of figures in the Appendix (Figs. A1-A4). We observed that the radial receiver functions of some stations (*e.g.*, MMRI and LBFI) show strong azimuthal variations in amplitude and time of the Moho PS arrival, indicating the presence of dipping interfaces (*e.g.*, Zheng *et al.* 2005, Linkimer *et al.* 2010). Because station MMRI exhibits the most acceptable coverage in backazimuth, we used the receiver functions from this station to illustrate how the amplitude of the PS conversion varies with backazimuth (Fig. 2). We found that receiver functions at this station show a strong PS arrival at ~6 s for the earthquakes arriving from 0°-75° and 230°-360°, and at ~5 s for the earthquakes arriving from

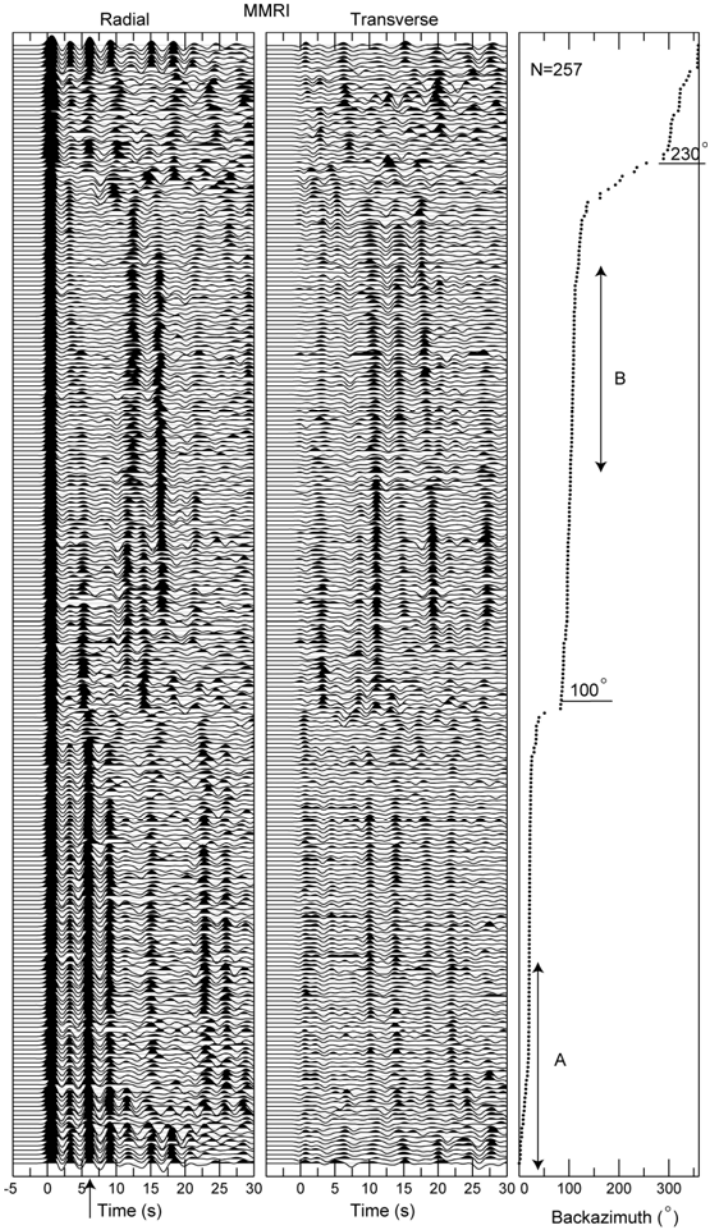


Fig. 2. Receiver functions computed at station MMRI plotted with equal spacing as a function of backazimuth. The black arrow in the left panel marks the PS phase. A and B in the right panel denote the teleseismic events coming from N-NE and E-SE directions, respectively. The N value on top of the right panel represents the total number of receiver functions used for the analysis at the station.

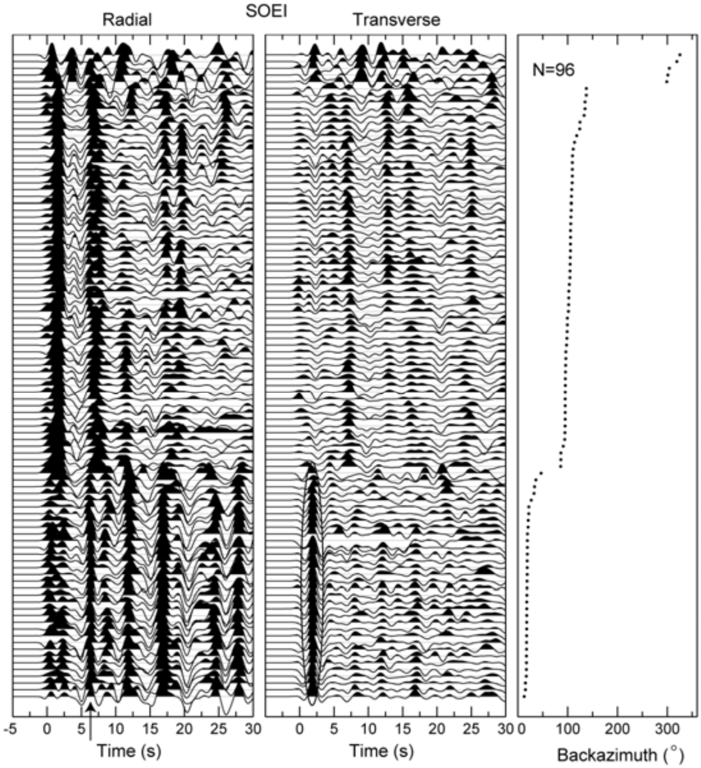


Fig. 3. Receiver functions computed at station SOEI plotted with equal spacing as a function of backazimuth. The black arrow in the left panel marks the P-to-S phase. The black ellipse in the mid panel marks the large amplitude arrivals on the transverse receiver functions. N value on the top of the right panel represents the total number of traces used for the analysis at the station.

~100°-120° backazimuth. The weak or unclear PS arrival is observed for the receiver functions obtained from the earthquakes located at the 120°-230°.

We also found that some stations (*e.g.*, SOEI, ATNI and BATI) exhibit relatively large amplitude arrivals on the transverse receiver functions (Figs. 3 and A1) suggesting the presence of complex structure such as dipping layers, sources of crustal scattering and anisotropy (*e.g.*, Levin and Park 1998, Jones and Phinney 1998). Since we only used the radial receiver functions for our subsequent analysis and the amplitude of the main signals in the radial receiver functions is larger than that in the transverse receiver functions, thus, the velocity model obtained from the 1 D inversion is still considered valid (Darbyshire 2003). In addition, the receiver functions at each site were stacked to suppress random background noise and to enhance the main signal, as well as to reduce the 3-D effects due to lateral structure variation and to provide an average crustal model (Zhu and Kanamori 2000).

Stacked receiver functions in radial components from all stations used in this study are shown in Fig. 4. The number of receiver functions for the stacking varies from about 16 to 257 at each individual station. The quality of the stacks may depend on the number of receiver functions at each station, similarity and coherence signal from subsurface structure on the receiver functions derived from different events at the same station and stability of the computed receiver function (Spasojevic and Clayton 2008). Stacking N number of traces will improve the signal to noise ratio by a factor of √N (Owens 1984). We found that some stations (*e.g.*, BMNI and WSI) have fewer traces compared to other stations, which may affect to the quality of the stacked receiver function. However, we observed that the Moho P to S

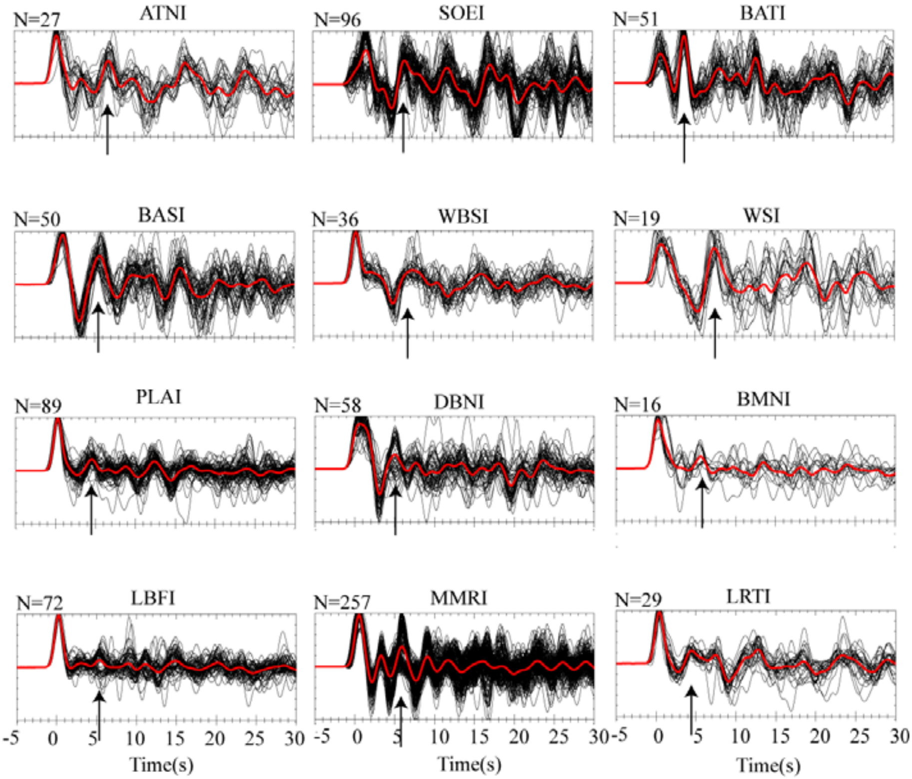


Fig. 4. The stacked radial receiver functions for all stations used in this study are represented by thick red lines. The individual receiver functions are shown by thin black lines and the Moho arrivals are marked by black arrows. The number of traces in each stack is indicated by N value on left top of each panel.

phase is still clearly visible on the stacked receiver functions for those stations.

The stacked receiver function at some seismic stations, such as SOEI and BATI, show time delay on the initial P pulse of the radial receiver functions, indicating the existence of the near surface sedimentary layer beneath the stations. The shift in the direct P pulse (from zero time) is caused by the superposition of the direct P and PS phases generated at the sediment-bedrock interface and cannot be removed by filtering process (*e.g.*, Cassidy 1992, Julia *et al*. 2008). We also observed that the negative pulses after the direct P pulse with large amplitude are quite apparent on the stacked radial receiver functions at some seismic stations (*e.g.*, WBSI, WSI and SOEI), which may indicate the presence of any low velocity layer in the mid crust. The effect of such a low velocity layer on receiver functions has been observed and modelled for other regions, for instance, by Bannister *et al.* (2004). Another prominent feature in the averaged radial receiver functions is the Moho P to S converted phase, which is generally apparent in some stacked radial receiver functions at around 5-7 s. However, few seismic stations also exhibit unclear or small amplitude for the Moho P-S phase on the receiver function (*e.g.*, station LBFI). The amplitude of the Moho P to S converted phase in a receiver function depends on the incident angle of the incoming P wave and the size of the velocity contrast generating the converted wave and its multiples (Ammon 1991).

Due to limitation on geophysical information about crustal structures in this region that can be used as a reference to assess the consistency of the receiver function analysis, we employed two inversion methods in this study. Firstly, the neighbourhood algorithm technique of Sambridge (1999) was utilized to derive the S wave velocity structure beneath the seismic station. To estimate the S wave velocity profiles of the crust and uppermost mantle through the inversion process, first, we stacked all radial receiver functions at each station to enhance the signal to noise ratio. For the inversion process, we selected the receiver function's time window from 5 s before the direct P wave arrival to 25 s after it. This time window was chosen to assure that all crustal and lithosperic phase arrivals were included in the analysis. In this study, we inverted the stacked receiver functions using the non-linear neighbourhood algorithm (NA) (Sambridge 1999). This technique is a fully non-linear approach based on random choice of model samples to search the optimum model using the misfits of the previous models. We divided the crustal structure for the inversions into six horizontal layers: sediment, basement, upper crust, middle crust, lower crust and upper mantle. Each layer contains four parameters describing the thickness of the layer (km), Vp/Vs, the S-wave velocity at the top and bottom of each layer. That provides 24 parameters (4 parameters in each layer). We set the tuneable parameters after a number of trials for each inversion involving 2000 iterations, providing 40,020 velocity models. A large range of initial random seeds and velocity profile parameterizations are set to exploit the dependence of the inversion to the initial choice of model samples.

Secondly, the Zhu and Kanamori (2000) *H-k* stacking method was applied to estimate the average of crustal thickness (*H*) and *Vp*/*Vs* ratio (*k*). We assumed that the crust and upper mantle are isotropic with a flat-lying planar interface due to the limited backazimuthal coverage of the data, which also prevents us from analysing the effect of anisotropy using receiver functions. We applied this *H-k* stacking procedure for the receiver functions to estimate the average thickness (*H*) and *Vp*/*Vs* ratio (*k*) of the crust beneath the stations as well as to test the consistency of the NA inversion results. The method sums the amplitudes of the PS signal and the multiples PPPS and PPSS + PSPS within given range of *H* and *k* values using the stacking function. The maximum value resulted from this coherent stack then provides the best estimate of crustal thickness (*H*) and *Vp*/*Vs* ratio (*k*) at particular station. The advantage using this analysis is that it does not require subjective picking of arrival times of the Moho PS converted phase and its multiple phases. It is necessary to note that this method is only valid in the case of simple crust. In the case of complex structural crust such as the presence of anisotropic or dipping structure and near surface sedimentary layers, this conventional *H-k* staking method may provide ambiguous estimates of crustal thicknesses and *Vp*/*Vs* ratios (Eagar 2011). Some stations in this region are situated in the place where sediment inclusions may dominate the near surface layer. The presence of sediment layer causes the delay arrival time of the direct P signal and the primary Moho conversion phase on the low frequency receiver function, as shown in Fig. 4. Therefore, in order to minimize the sediment effect to the *H-k* grid-searching analysis, we employed the two layer *H-k* stacking method developed by Yeck *et al.* (2013). The method requires the information about the sediment thickness and its *Vp*/*Vs* ratio beneath the seismic stations. The estimated sediment thickness and *Vp*/*Vs* ratio is then applied to correct the arrival of the subsediment interface converted phases using time adjustments due to the presence of sediments (Yeck *et al*. 2013). In this *H-k* analysis, we employed the values of the sediment thickness and its *Vp*/*Vs*, the sediment P wave velocities and the average crustal velocities previously obtained from the NA-inversion.

To assess the accuracy and stability of stacking analysis, we employed bootstrap stacking of the receiver functions for each station using the same procedure as reported by Eagar and Fouch (2012). The bootstrap sampling technique produces a random sample with replacement from the original data (Efron and Tibshirani 1986, 1991). Here, we constructed 1000 resampled data sets selected at random from the original receiver function data set. The value of H and k were determined for every resampled data. We then calculated the mean and standard deviation of the resulting parameters.

3. RESULTS

3.1 S-wave velocity profiles

S-wave velocity models obtained from the inversion for stations located in Timor Island: ATNI, SOEI and BATI are shown in Fig. 5. The inversion result for station ATNI shows S velocities of ~2-3 km/s near the surface, then increasing to 3.2 km/s at 10 km depth. The mid crust in this station is observed at ~10-26 km depth with average Vs of 3.3 km/s. At 26-34 km depth, the velocity profile indicates a low-velocity zone in the lower crust, with an S-wave velocity as low as ~2.6 km/s. The Moho beneath this station is estimated at a depth of ~37 km, where the S-wave velocity reaches 4.0 km/s. For station SOEI, located to the west of ATNI, the inversion results indicate S wave velocity of ~1-2.5 km/s at the surface, increasing to ~3.2-3.4 km/s

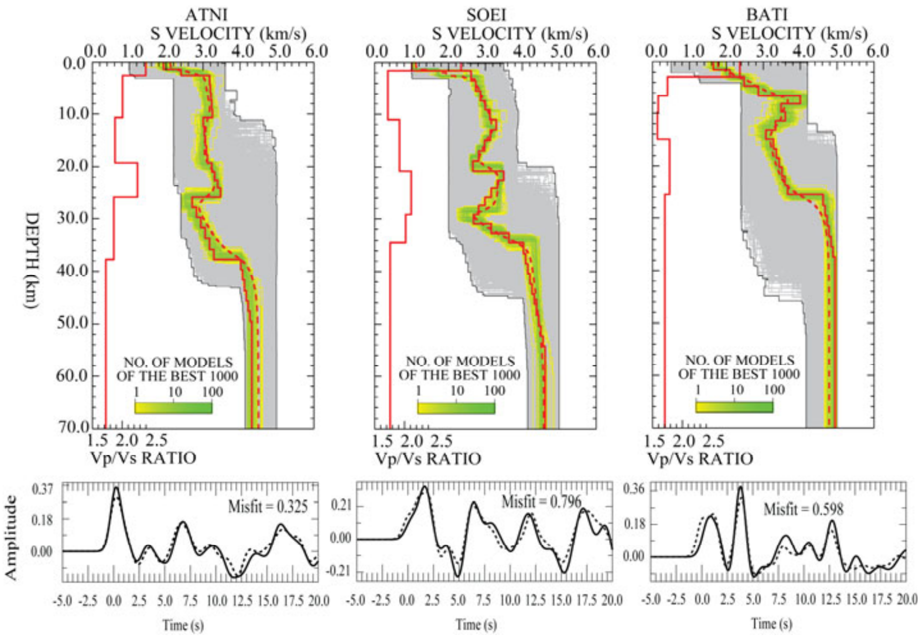


Fig. 5. S wave velocity models for seismic stations located in Timor Island derived from the non-linear inversion. The dashed and solid red lines represent the average and the best fitting model, respectively. The grey shaded area indicates all models (40 020 models) searched in the inversion. The red line on the left represents the best fitting Vp/Vs ratio. Bottom panels show the predicted and observed receiver function indicated by dashed and solid black lines, respectively.

at 10 km depth and fluctuating down to ~29 km depth. The S velocity then increases and becomes larger than 4 km/s below ~34 km depth, reflecting the mantle velocities beneath the station. The low-velocity zone is also observed in the lower crust beneath this station at a depth between ~20-34 km with S-wave velocity as low as ~2.8 km/s. The average *Vp*/*Vs* ratio obtained from the inversion is relatively high for the entire crust for both stations, ATNI and SOEI, which are ~1.9 and ~2.0, respectively. For station BATI, situated in the southwestern part of the island, the S velocity is less than 2 km/s at the first 2 km depth, then increases to 4.0 km/s at 6.0-7.0 km depth. Uppermost mantle velocity is reached at 28 km depth. The crustal *Vp*/*Vs* ratio in this station is lower (~1.7) compared to those at other stations.

The S wave velocity profiles derived from the non-linear inversion for the Sumba stations, BASI, WBSI and WSI, are shown in Fig. 6. At station BASI, the near surface is covered by low velocity material with a velocity less than 2 km/s. The velocities increase rapidly to 3.8 km/s down to 10 km depth, and then decrease slightly to ~3.2-3.4 km/s at a depth between ~14 km and 24 km. The Moho discontinuity is estimated at ~36 km depth

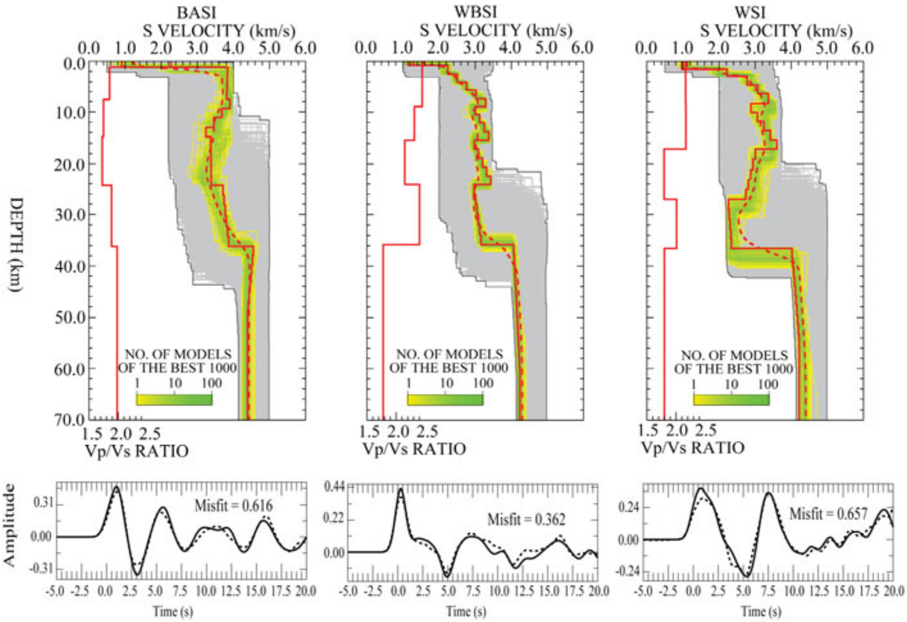


Fig. 6. S wave velocity models for seismic stations located in Sumba Island derived from the non-linear inversion. The dashed and solid red lines represent the average and the best fitting model, respectively. The red line on the left represents the best fitting *Vp*/*Vs* ratio.

with S velocities higher than 4.2 km/s. At station WBSI, the model indicates that the S wave propagates with a velocity less than 2 km/s at the near surface, then increasing to ~3 km/s at 10 km depth. The model also reveals a modest negative velocity gradient that corresponds to a velocity of ~2.9 km/s at depths between ~24-36 km. Below 36 km, the velocity increases to ~4.2 km/s reflecting the S wave velocity of the mantle. The velocity profile for station WSI is started with the low near surface velocity layer with a velocity less than 1 km/s. In this profile, we also observed the presence of low velocity zone at a depth of 20-35 km, characterized by a strong negative velocity gradient with a minimum velocity of 2.2 km/s. The crust-mantle transition is observed at a depth below 36 km marked by a positive velocity gradient. In addition, another interesting feature from the velocity profiles for the Sumba stations (especially at station WBSI and WSI) is that the *Vp*/*Vs* ratio is relatively high (> 2) in the lower crust where the low-velocity zones are also observed.

Inversion results for stations PLAI, DBNI and BMNI, which are located in Sumbawa Island, are shown in Fig. 7. The solutions of the inversion for the three stations have similar characteristics at the near surface with low ve-

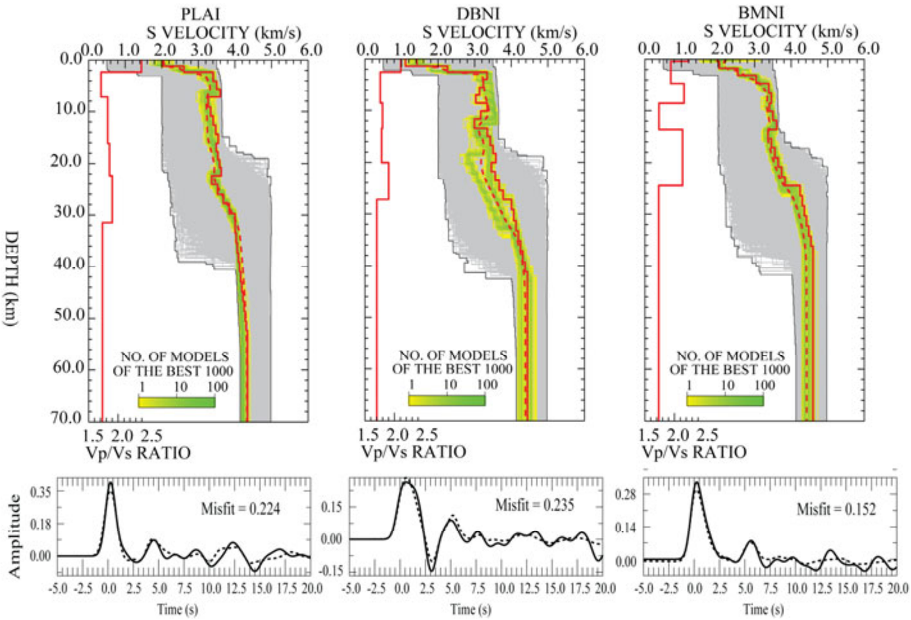


Fig. 7. S wave velocity models for seismic stations located in Sumbawa Island derived from the non-linear inversion. The dashed and solid red lines represent the average and the best fitting model, respectively. The red line on the left represents the best fitting *Vp*/*Vs* ratio.

locity below 2 km/s. Slightly negative velocity gradients are observed in the mid crust beneath the three stations. The velocities appear to reach mantle-crust discontinuity layer at depths of ~30, 28, and 27 km for station PLAI, DBNI and BMNI, respectively. The *Vp*/*Vs* ratios are high (> 1.84) for all seismic stations. Furthermore, we observed that although the velocity model obtained for station BMNI is produced by inverting the stacked receiver functions consisting of only few traces (16 traces), the waveform misfit between the synthetic waveform and the observed radial receiver function is quite reasonable. Figure 8 displays S wave velocity profiles obtained from the non-linear inversion for stations located in the Flores Island, namely LBFI, MMRI and LRTI. At station LBFI, the inversion solution shows a low S wave velocity of ~2 km/s near the surface indicating the presence of sediment layer. The velocity increases gradually and seems to reach mantle velocities of ~3.7 km/s at 30 km depth, even though the crust-mantle transition is not well resolved. For station MMRI, the low velocity of less than 2 km/s is observed at the near surface. The velocity then increases to about 3.4 km/s at 10 km depth. The modest low velocity zone is observed between 12 and 22 km depth and Moho is detected at a depth of 32-34 km. For station LRTI,

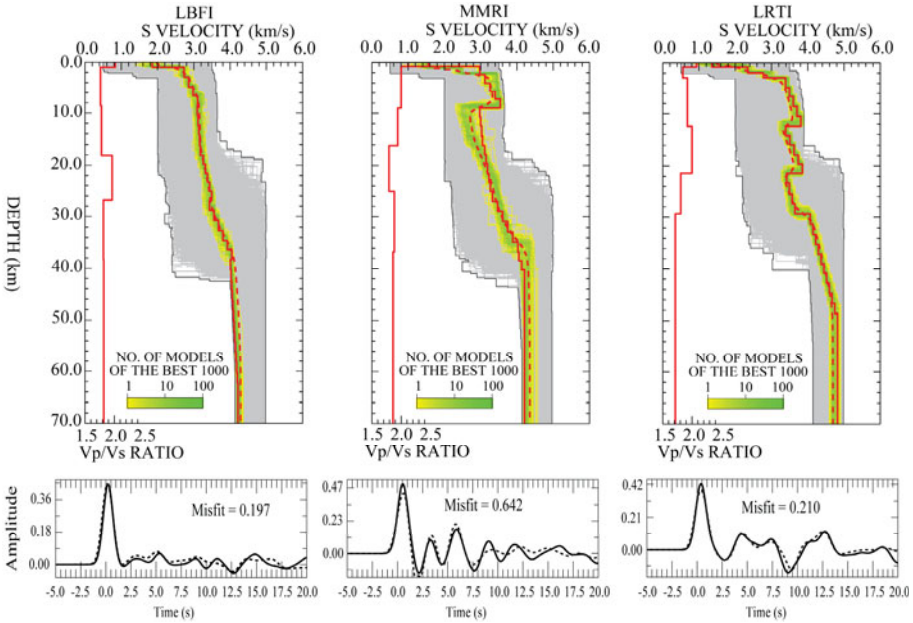


Fig. 8. S wave velocity models for seismic stations located in Flores Island derived from the non-linear inversion. The dashed and solid red lines represent the average and the best fitting model, respectively. The red line on the left represents the best fitting *Vp*/*Vs*.

we observed that the S wave velocity fluctuates down to Moho depth at ~35 km depth. The high *Vp*/*Vs* ratios are found in the crust beneath these three stations.

It is important to note that we found the difficulties in modelling the stacked receiver function for some stations where the complex structure is present. As discussed earlier, some stations (*e.g.*, ATNI, SOEI and BATI) exhibit the high level of energy on the transverse receiver functions, which indicates the presence of complex structure such as sources of crustal scattering and anisotropy (*e.g.*, Levin and Park 1998, Jones and Phinney 1998). Strong backazimuth variations in amplitude and time of PS arrival are also observed on the radial receiver functions for some stations, which suggest the presence of dipping structures (*e.g.*, Zheng *et al.* 2005, Linkimer *et al.* 2010). We note that although the stacked receiver function for some stations (*e.g.*, MMRI and SOEI) consists of reasonable number of traces, the inversion result shows poor inversion solution. A number of trials involving a range of incidence angles, and a range of model parameters have been used to improve the waveform fit on these stations, but we still obtained unsatisfactory solutions. However, the waveform fits are generally quite reasonable, matching the initial PS phases in the first 5-6.5 s, but poorly matching the later phases. This suggests that the interpretation for the crustal model may still be valid.

3.2 Crustal thickness and *Vp*/*Vs* ratio

The results of the crustal thickness and *Vp*/*Vs* ratio estimation computed from the *H-k* stacking method for all stations used in this study are summarized and depicted in Table 1 and Fig. 9, respectively. In general, the Moho depth and *Vp*/*Vs* ratio computed from *H-k* stacking method are comparable with those obtained from the non-linear inversion. At stations ATNI and SOEI located in Timor Island, the *H-k* analysis resulted in similar Moho depth of ~33-35 km with *Vp*/*Vs* ratio ranging from 2 to 2.2. We observed that at station BATI, located to the west of ATNI and SOEI stations, the Moho depth is shallower, that is, about 27 km. We found that receiver functions at station BATI contain unusual strong reflection at 3-4 s and the amplitude is higher than the direct P amplitude. This strong reflection may cause ambiguity in the measurement and we will discuss this in the following section. Hence, we suggest that this could cause bias in the *H-k* analysis. In Sumba Island, we inferred the comparable Moho depth of ~32-36 km at all seismic stations. The *Vp*/*Vs* ratios in this island are relatively high, varying between 1.88 and 2.4. We also observed consistent Moho depth with high *Vp*/*Vs* (1.9-2.2) beneath seismic stations in Sumbawa Island, which have ~28 km deep Moho. The Moho depth obtained for Flores Island ranges from ~28 to 34 km with consistent *Vp*/*Vs* around ~ 1.95.

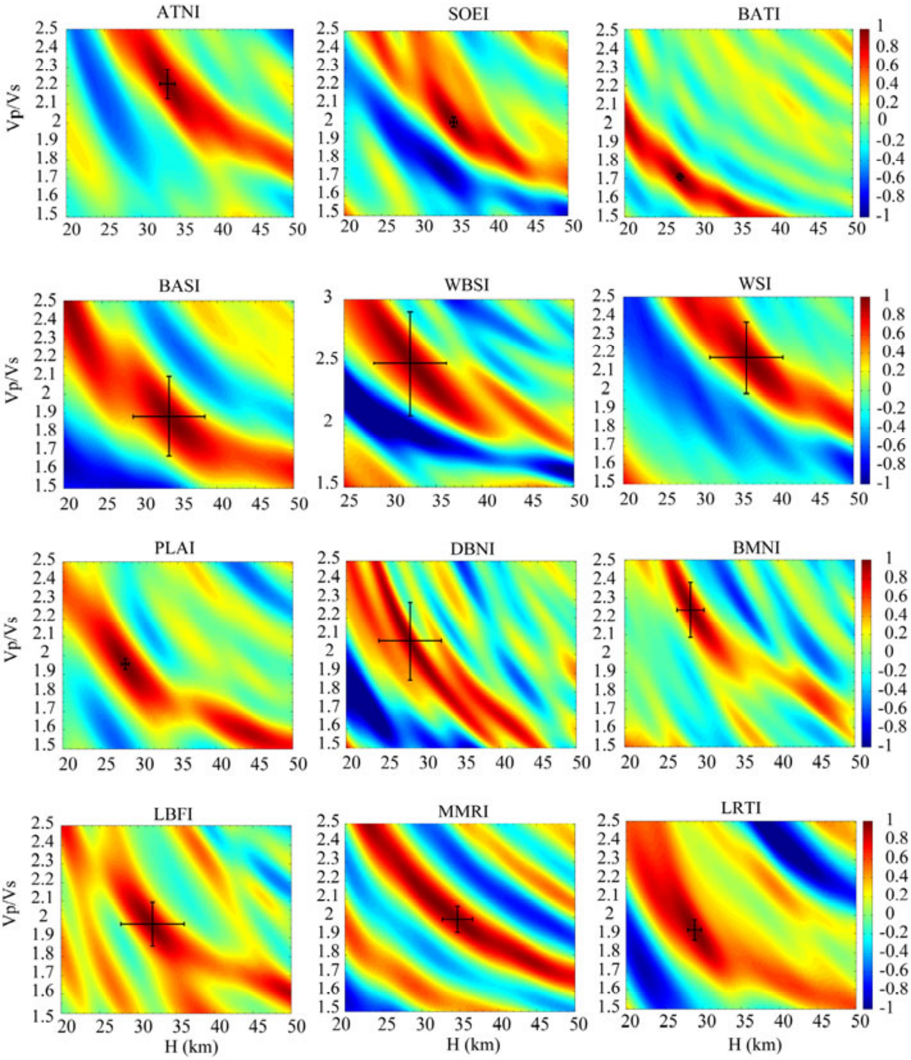


Fig. 9. Results of *H*-*k* stacking analysis to search the best estimate of crustal thickness and *Vp*/*Vs* ratio for all stations. The black crossbars indicate the maximum stacking amplitude. Small and large crossbars represent the uncertainty of measurements derived from the bootstrap sampling method.

Table 1

Summary of the crustal thickness and *Vp*/*Vs* with bootstrap error obtained from *H*-*k* stacking analysis

Island	Station	Crustal thickness *H* [km]	*Vp*/*Vs* ratio	Number of records
Timor	ATNI	33.36 ± 1.00	2.21 ± 0.08	27
	SOEI	34.67 ± 0.45	2.01 ± 0.03	96
	BATI	27.28 ± 0.39	1.71 ± 0.02	51
Sumba	BASI	33.73 ± 4.68	1.88 ± 0.21	50
	WBSI	32.19 ± 3.94	2.49 ± 0.42	36
	WSI	36.04 ± 4.77	2.18 ± 0.19	19
Sumbawa	PLAI	28.16 ± 1.95	1.95 ± 0.03	89
	DBNI	28.35 ± 4.07	2.07 ± 0.21	58
	BMNI	28.58 ± 1.76	2.24 ± 0.15	16
Flores	LBFI	31.91 ± 4.14	1.97 ± 0.12	72
	MMRI	34.67 ± 1.96	1.99 ± 0.07	257
	LRTI	28.95 ± 0.88	1.92 ± 0.05	29

4. DISCUSSION

We analysed receiver functions using two different methods, the NA-non-linear inversion and the *H*-*k* stacking analysis, to estimate the velocity structures, *Vp*/*Vs* ratios and the average of crustal thickness beneath seismic stations along the Sunda-Banda Arc transition zone. In general, the results obtained from the two methods indicate the consistency and robustness of the measurements.

4.1 Timor Island

Timor Island is fold-thrust mountain built during initial stage of the arc-continent collision (Harris and Audley-Charles 1987). The geological structure of the island is complex, involving widely different rock types such as sediment derived from an accretionary wedge, metamorphic rocks including high pressure types, ophiolites and continental crystalline fragments (Barber *et al*. 1977, Barber 1981, Charlton 1989). This complicated structure results from the collision of the Australian continent margin with the Banda arc. In the near surface, the island is generally covered by thick sediment deposited up to 2.5 km depth since 5.5 Ma (Haig 2012). In the deeper part of the crust, beneath station ATNI and SOEI, the results obtained from the NA-inversion and *H-k* stacking indicate the presence of material with relatively high *Vp*/*Vs* ratios. The low and high *Vp*/*Vs* ratio values represent the average crustal composition and they depend mainly on the lithology, the presence of partial melting, temperature, cracks and pore fluid (Fountain and Christensen 1989, Zandt and Ammon 1995, Koch 1992). Christensen (1996) has classified the typical values of *Vp*/*Vs* for various rock types. For example, the low values (~1.70) represent felsic rocks (*i.e.*, granite), the intermediate values (~1.78) and the high values (~1.84) are values for intermediate rocks (*e.g.*, diorite) and mafic rocks (*e.g.*, basalt), respectively. Thus, we suggest that the crust in Timor Island might be dominated by mafic and ultramafic bodies which might be derived from the Banda forearc basement or lower plate of the Australian continental margin, as suggested by geological and gravity studies (Kaye 1989, Ishikawa *et al.* 2007, Harris 2011). Furthermore, the low velocity gradient found in the lower crust with high *Vp*/*Vs* ratio might be related to the serpentinisation of some basic rocks during collision between Australian passive margin and volcanic arc, which may produce a mafic and ultramafic precursor for the movement of thrust sheets (Kaye 1989). Christensen and Salisbury (1975) reported that the typical values of the *Vp*/*Vs* ratio are higher than 2 for serpentinized and very serpentinized ultra mafic rocks and less than 1.95 for mafic rocks. In West Timor, a geological investigation reported the existence of spinel peridotites, recrystallized partly to tremolite-talc and serpentine schist, and mafic rocks (Helmers *et al.* 1989). The gravity study suggests that these rocks might have been carried ashore by obducting mantle during the later stage of collision (Kaye 1989).

We found that the Moho depth beneath two stations (ATNI and SOEI) is of about 34-37 km obtained from the NA inversion and *H-k* stacking. These results are comparable to those obtained by other studies and reflect the influence of continental character (Milsom and Audley-Charles 1986, Bowin *et al.* 1980). However, for station BATI located in the southwestern part of the island, the results show a shallower Moho depth, of ~28 km, with low *Vp*/*Vs* ratio. This can be explained that strong reflection appearing at ~3-4 s on the receiver functions might correspond to the Australian lower crust inserting beneath southern part of Timor, as suggested by Kaye (1989) rather than Moho discontinuity. Thus, this strong reflection masks the Moho conversion phase causing ambiguous results for the NA and *H-k* inversions.

4.2 Sumba Island

Sumba Island is located to the south of volcanic arc within the forearc basin and lies at the border of the transition zone between the subduction of the Indo-Australian oceanic lithosphere along the Sunda margin in the west and the collision of the Australian continent margin with the Banda arc in the east. The location of Sumba Island is separated by two deep forearc basins, the Lombok Basin to the west and the Savu Basin to the east. The major geologic formation of Sumba Island contains volcanic, plutonic, and volcaniclastis rocks (Rutherford *et al*. 2001). These rock formations record the volcanic activities that began at around 80 Ma or earlier and ceased at around 31 Ma (Rutherford *et al*. 2001, Abdullah *et al*. 2000). Tectonic studies suggest that Sumba Island is considered as a micro-continent or continental fragment detached from its origin and drift away to its current position (Hamilton 1979, Abdullah *et al*. 2000). Sumba Island ceased moving and started to rise when the Australian continental crust collided with the Banda Arc at 8 Ma (Rutherford *et al*. 2001, Keep *et al*. 2003). Keep *et al.* (2003) suggested that the uplift of Sumba might be governed by the continent-arc collision tectonic process, as the buoyant of promontory Australian continental crust resisted to subduct and underplated beneath Sumba. As a consequence, this continued tectonic process has led to the Australian continental crust colliding with the thickened continental crust (as result of underplated product) beneath the island.

The Moho depth beneath the Island obtained from the NA inversion and *H-k* stacking analysis varies from approximately 32 to 36 km. This estimation is consistent with the typical values of the continental character (Christensen and Mooney 1995). The island is also characterized by the presence of low velocity zone in the lower crust with high *Vp*/*Vs*. Low velocity with high negative gradient on velocity profile indicates the presence of fluids (Mosalve 2013). In this case, the fluids can be water or partial melts. However, it is still difficult to say that the lower crustal low-velocity zone beneath this Island is attributed to high temperature fluids related to the geothermal activities, because the volcanic activities in this area ceased at 31 Ma (Abdullah *et al.*, 2000). As the uplifted Sumba Island might be influenced by the Australian Plate, the interaction between the deeper part of crustal root underneath Sumba and the finger of Australian continental crust might cause frictional strength at the plate boundary. Thus, we consider that the frictional process created the weak or fracture zone in the lower crust. The presence of fracture zone containing fluid might increase the *Vp*/*Vs* ratio. Conversely, the porous fracture zone might reduce the *Vp*/*Vs* ratio (Wang 2012, Kaypak 2008). High *Vp*/*Vs* ratios due to the crack fabric have been reported in other regions, for example, by Matsubara *et al*. (2008) and Kodaira *et al.* (2004). Matsubara *et al.* (2008) observed high *Vp*/*Vs* in the lower crust of southwestern (SW) Japan through seismic tomography, which they interpreted as due to the high pore-fluid pressure resulting from the dehydration of fluid from the oceanic crust of the Philippine Sea plate beneath SW Japan. Other seismic tomography studies performed in the subduction zone in Japan and Cascadia also found low velocity zone with high *Vp*/*Vs* interpreted as regions with high pore fluid pressure (Kodaira *et al*. 2004, Audet *et al*. 2009). Furthermore, laboratory measurements proposed that high *Vp*/*Vs* with low velocity may be due to high pore fluid pressure and crack anisotropy or mineral alignment (Bezacier *et al*. 2010, Wang *et al*. 2012).

The presence of low velocity zone in the lower crust beneath Sumba Island may affect the genesis of earthquakes. The plot of the shallow events extracted from the EHB catalogue (Engdahl *et al*. 1998, http://www.isc.ac.uk) shows that the earthquakes are mainly concentrated above the lower crust in the eastern part of the island (Fig. 10). This suggests that the frictional process, as a result of the interaction between the crustal root under-

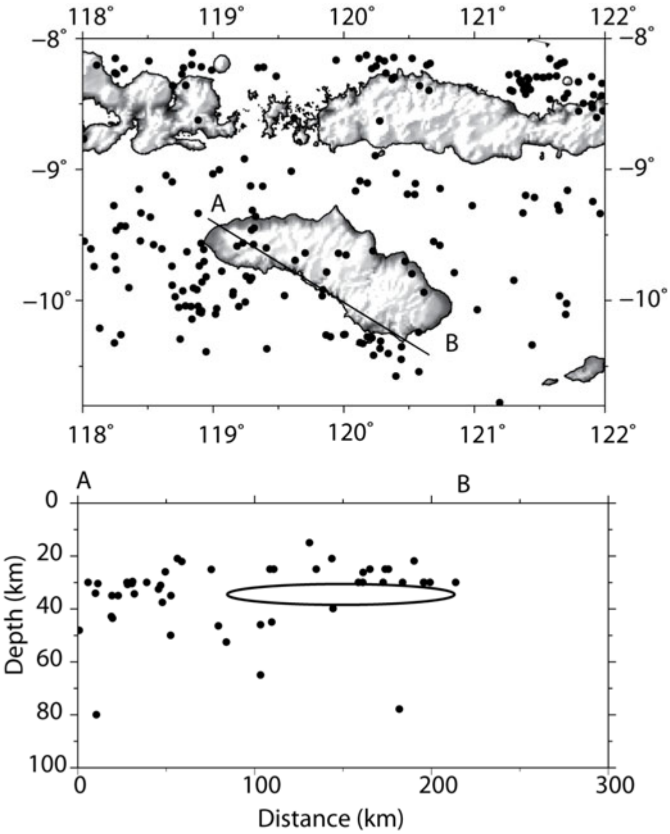


Fig. 10. Top panel shows the epicenters of the shallow events extracted from the EHB catalogue around Sumba Island. The A-B black line represents the position of cross section shown in bottom panel. The black ellipse in bottom panel indicates the low crustal low velocity layer.

neath Sumba and the finger of Australian continental crust, may increase the temperature along the crustal boundary, causing the lower crust to deform aseismically. As also observed by other studies (*e.g.*, Bai *et al.* 2015, Chiarabba and Amato 1996), the seismogenic layer is almost confined above the low crustal low velocity layer. Bai *et al.* (2015) observed that there are no earthquakes in the lower crust beneath Central Tibet, which they suggest that this could be due to the presence of partial melting in the lower crust as a result of the plate convergence. Chiarabba and Amato (1996) also found that the relocated events in the Northern and Southern Apennines are mainly concentrated above the low velocity of the lower crust. These researches lead to the conclusion that the lower crust under such condition may not be cold enough to produce earthquakes. However, further investigation involving larger local seismic data around Sumba Island is necessary in order to substantiate the role of this velocity zone to the earthquake generation process.

4.3 Sumbawa and Flores Islands

Sumbawa and Flores Islands are part of active magmatic inner Sunda-Banda arc. The volcanic arc has been accommodated by northward subduction of Indo-Australian crust. Recently, in eastern part of the arc, including Sumbawa and Flores, collision between Australian continent crust with Banda arc modified the shape of the arc, leading to the idea of subduction polarity reversal north of Flores and Sumbawa (Hamilton 1977, McCaffrey 1988). Bowin *et al.* (1980) suggested that the eastern part of Sunda-Banda arc lies on the oceanic crust. It was suggested that the crustal thickness in this region is more than 22 km, which is greater than that of normal oceanic crust (Prasetyo 1992). However, other works suggested that the insertion of Australian continental character underneath southern part of the Banda arc might elevate the crustal thickness (Richardson and Blundell 1996, Keep 2003). From our analysis, we observed that the Moho depth derived from the NA inversion and *H-k* stacking is about 28 km and 30-34 km beneath Sumbawa and Flores Islands, respectively. The deeper Moho depth beneath Flores Island might be due to a consequence of arc-continent collision. This inner arc region is also characterized high Vp/Vs ratio (> 1.8) and low velocity zone in the upper and mid crust as observed in stations beneath Flores Island. We suggest that those might be due to the presence of magmatic materials ascending from the subducted slab to the volcanic islands.

5. CONCLUSION

Here we presented the crustal properties and velocity structure derived from teleseismic receiver function using the non-linear NA inversion and *H-k* stacking from twelve permanent seismic stations around Sunda-Banda arc transition zone. We obtained that the Moho depth at this transition zone is ranging from 28 km beneath Sumbawa Island to 37 km beneath Timor Island. We suggest that the thicker crustal thickness observed beneath Sumba and Flores Islands might be related to the arc-continent collision, in which the underplating of buoyant Australian crust during the arc-continent collision in southern part of the Banda arc increases the crustal thickness, as suggested by Richardson and Blundell (1996) and Keep (2003). In Timor and Sumba Islands, we observed relatively high *Vp*/*Vs* ratios with the presence of low velocity zone in the lower crust. In Timor Island, we suggest that they might be related to the serpentinisation of some basic rocks during the arc-continent collision producing mafic and ultramafic bodies, as suggested by previous studies (Helmers *et al*. 1989, Kaye 1989). Whereas in Sumba they can be associated with the presence of fluid filled fracture zone due to interaction in the plate boundary between the deeper crust of Sumba and the underplated Australian continental crust (Keep 2003, Fleury *et al*. 2009). High *Vp*/*Vs* ratios are also observed in Sumbawa and Flores Island which may correlate to the magmatic activies in the subducted slab.

Acknowledgments. This research was supported by the AIFDR-GREAT-ITB and LIPI. We acknowledge Indonesian Agency for Meteorology, Climatology, and Geophysics (BMKG) and GFZ Potsdam for providing the data used in this study. We also thank R.B. Herrmann, C.J. Ammon, Malcolm Sambridge and K.C. Eagar for providing softwares used in this study. We also benefited from discussion with J. Julio.

Appendix

Receiver function plots as a function of event backazimuth

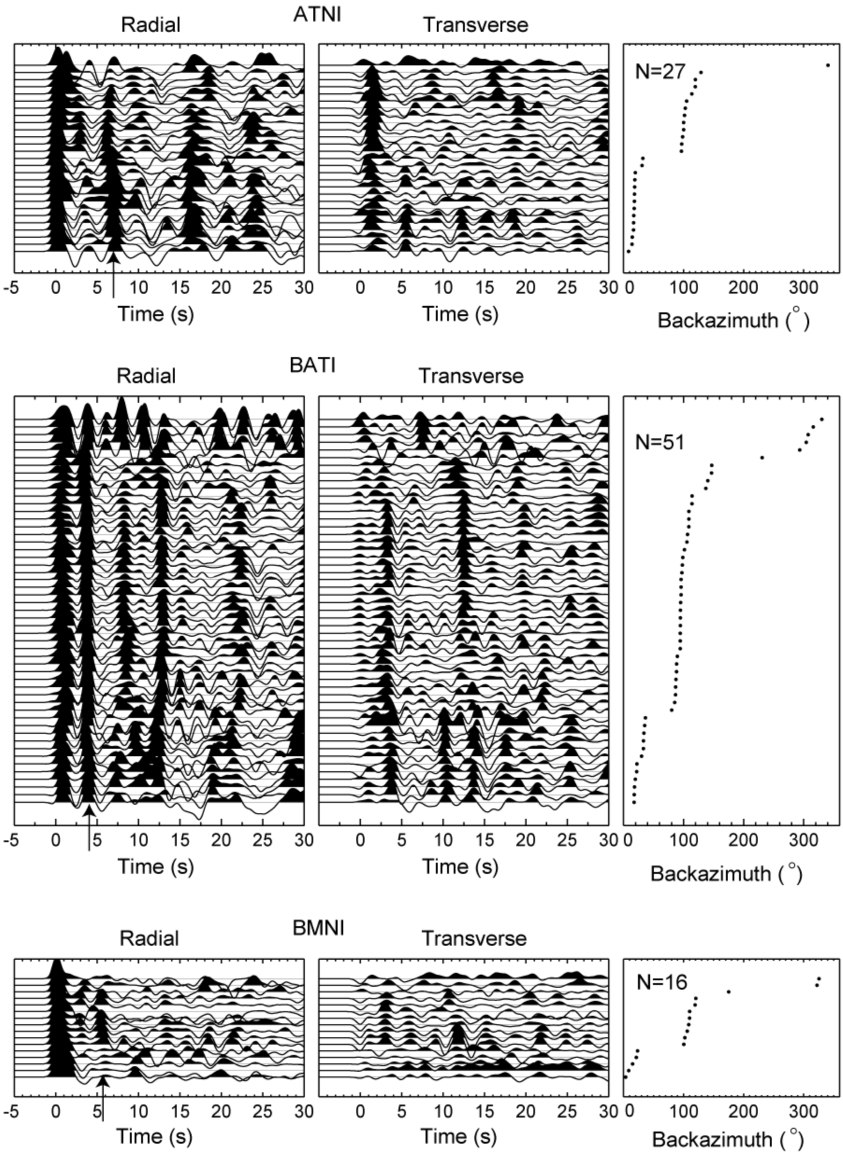


Fig. A1. Receiver functions at stations ATNI, BATI and BMNI plotted with equal spacing as a function of backazimuth. The black arrow in the left panel marks the PS phase.

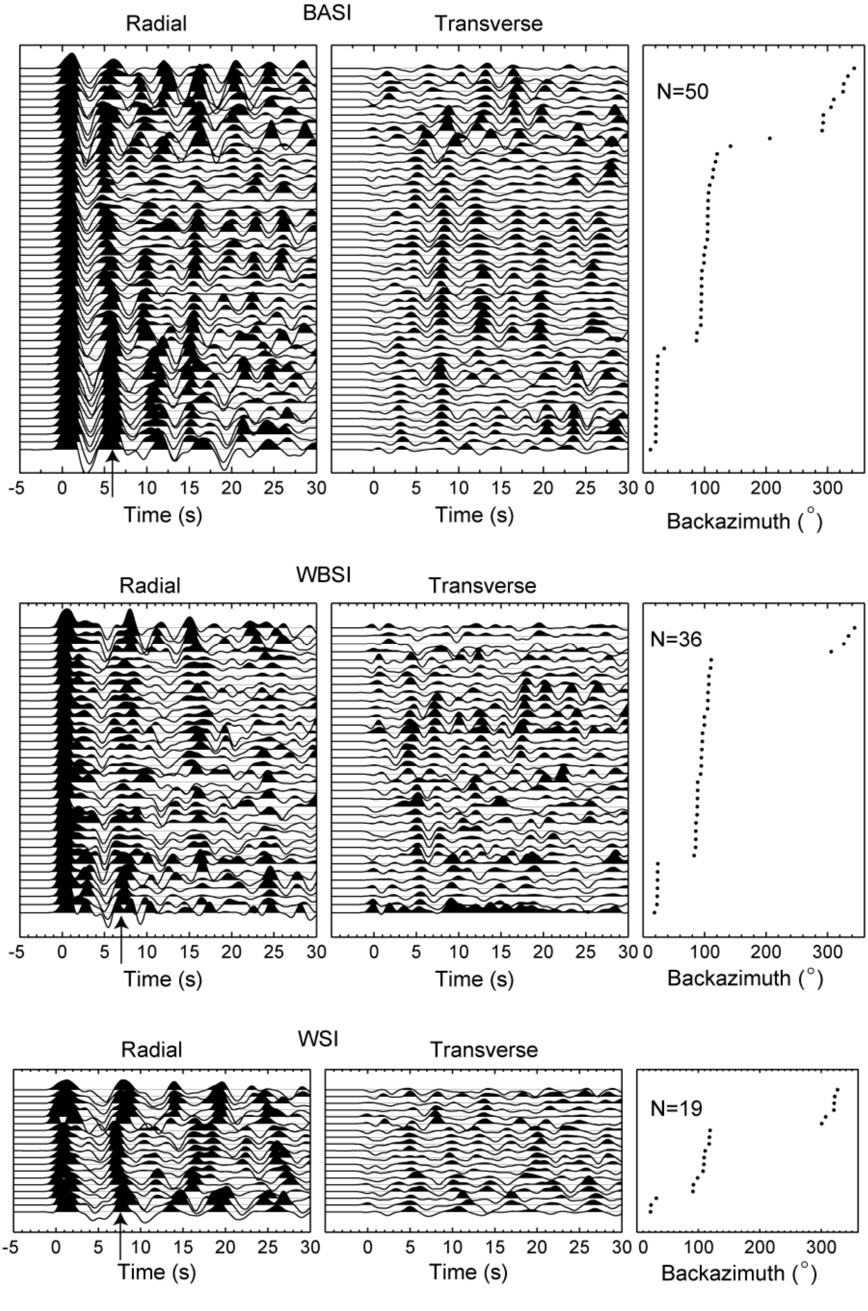


Fig. A2. Receiver functions at stations BASI, WBSI and WSI plotted with equal spacing as a function of backazimuth. The black arrow in the left panel marks the PS phase.

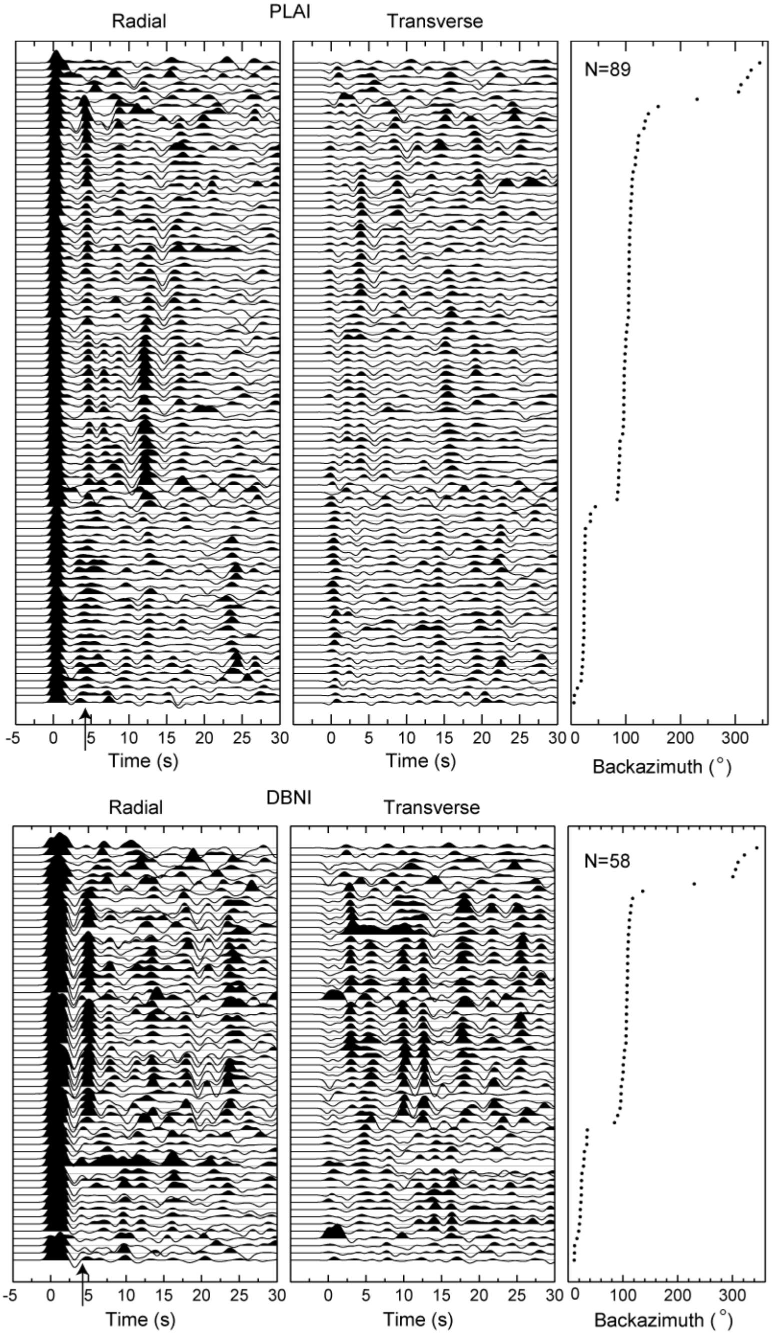


Fig. A3. Receiver functions at stations PLAI and DBNI plotted with equal spacing as a function of backazimuth. The black arrow in the left panel marks the PS phase.

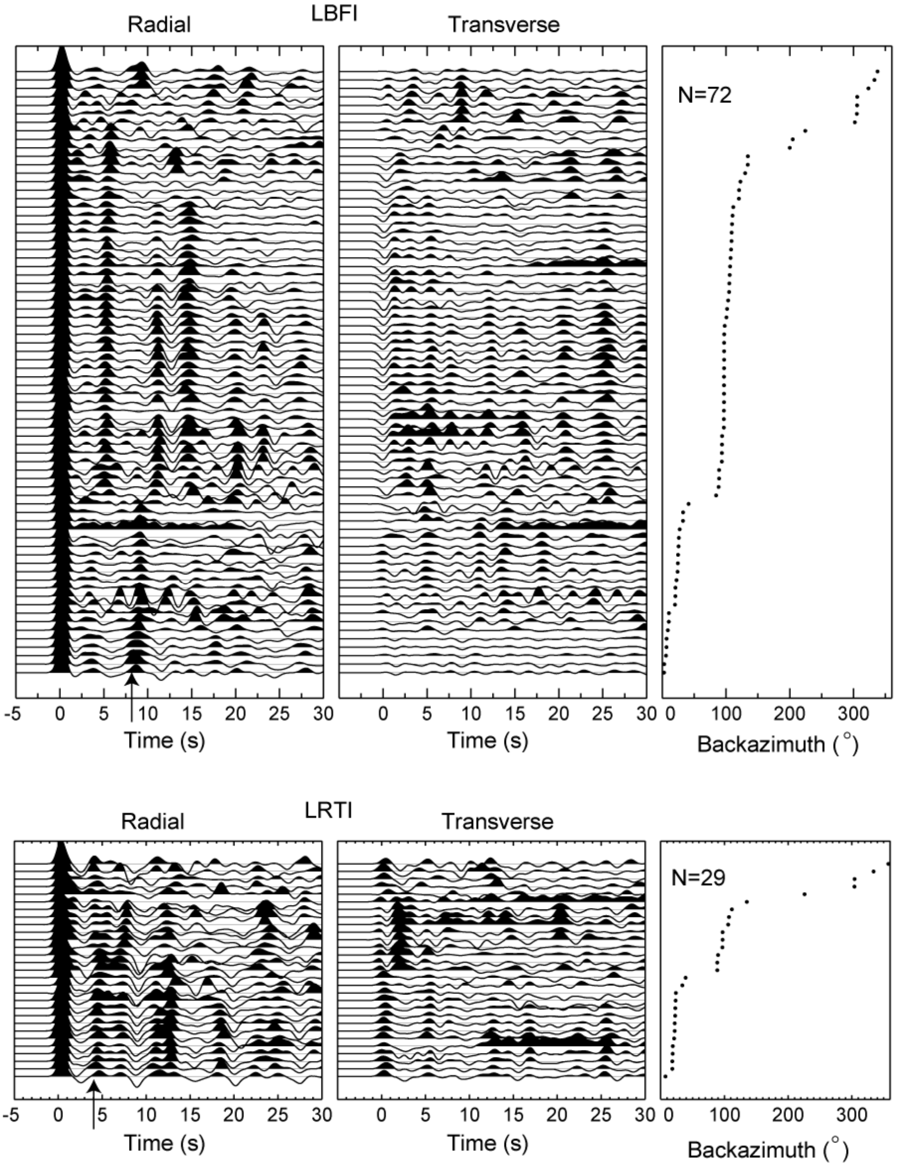


Fig. A4. Receiver functions at stations LBFI and LRTI plotted with equal spacing as a function of backazimuth. The black arrow in the left panel marks the PS phase.

References

Abdullah, C.I., J.-P. Rampnoux, H. Bellon, R.C. Maury, and R. Soeria-Atmadja (2000), The evolution of Sumba Island (Indonesia) revisited in the light of new data on the geochronology and geochemistry of the magmatic rocks, *J. Asian Earth Sci.* **18**, 5, 533-546, DOI: 10.1016/S1367-9120(99)00082-6.

Ahmed, A., S. Leroy, D. Keir, F. Korostelev, K. Khanbari, F. Rolandone, G. Stuart, and M. Obrebski (2014), Crustal structure of the Gulf of Aden southern margin: Evidence from receiver functions on Socotra Island (Yemen), *Tectonophysics* **637**, 251-267, DOI: 10.1016/j.tecto.2014.10.014.

Ammon, C.J. (1991), The isolation of receiver effects from teleseismic P waveforms, *Bull. Seismol. Soc. Am.* **81**, 6, 2504-2510.

Audet, P., M.G. Bostock, N.I. Christensen, and S.M. Peacock (2009), Seismic evidence for overpressured subducted oceanic crust and megathrust fault sealing, *Nature* **457**, 7225, 76-78, DOI: 10.1038/nature07650.

Bai, L., G. Li, N.G. Khan, J. Zhao, and L. Ding (2015), Focal depths and mechanisms of shallow earthquakes in the Himalayan–Tibetan region, *Gondwana Res*., DOI: 10.1016/j.gr.2015.07.009 (in press).

Bannister, S., C.J. Bryan, and H.M. Bibby (2004), Shear wave velocity variation across the Taupo Volcanic Zone, New Zealand, from receiver function inversion, *Geophys. J. Int.* **159**, 1, 291-310, DOI: 10.1111/j.1365-246X.2004.02384.x.

Barber, A.J. (1981), Structural interpretations of the island of Timor, eastern Indonesia. **In:** A.J. Barber and S. Wiryosujono (eds.), *The Geology and Tectonics of Eastern Indonesia*, Special Publication of the Geological Research and Development Centre, Bandung, 183-197.

Barber, A.J., M.G. Audley-Charles, and D.J. Carter (1977), Thrust tectonics in Timor, *J. Geol. Soc. Australia* **24**, 1, 51-62, DOI: 10.1080/00167617708728966.

Bezacier, L., B. Reynard, J.D. Bass, C. Sanchez-Valle, and B.V. Moortele (2010), Elasticity of antigorite, seismic detection of serpentinites, and anisotropy in subduction zones, *Earth Planet. Sci. Lett.* **289**, 198-208, DOI: 10.1016/j.epsl.2009.11.009.

Bowin, C., G.M. Purdy, C. Johnston, G. Shor, L. Lawver, H.M.S. Hartono, and P. Jezek (1980), Arc-continent collision in Banda Sea region, *Am. Assoc. Petrol. Geol. Bull.* **64**, 6, 868-915.

Cassidy, J.F. (1992), Numerical experiments in broadband receiver function analysis, *Bull. Seismol. Soc. Am*. **82**, 3, 1453-1474.

Charlton, T.R. (1989), Stratigraphic correlation across an arc-continent collision zone: Timor and the Australian Northwest Shelf, *Aust. J. Earth Sci.* **36**, 2, 263-274, DOI: 10.1080/08120098908729485.

Chiarabba, C., and A. Amato (1996), Crustal velocity structure of the Apennines (Italy) from P-wave travel time tomography, *Ann. Geofis*. **39**, 6, 1133-1148, DOI: 10.4401/ag-4042.

Christensen, N.I. (1996), Poisson's ratio and crustal seismology, *J. Geophys. Res*. **101**, B2, 3139-3156, DOI: 10.1029/95JB03446.

Christensen, N.I., and W.D. Mooney (1995), Seismic velocity structure and composition of the continental crust: A global view, *J. Geophys. Res*. **100**, B6, 9761-9788, DOI: 10.1029/95JB00259.

Christensen, N.I., and M.H. Salisbury (1975), Structure and constitution of the lower oceanic crust, *Rev. Geophys*. **13**, 1, 57-86, DOI: 10.1029/ RG013i001p00057.

Darbyshire, F.A. (2003), Crustal structure across the Canadian High Arctic region from teleseismic receiver function analysis, *Geophys. J. Int*., **152**, 2, 372-391, DOI: 10.1046/j.1365-246X.2003.01840.x.

Eagar, K.C., and M.J. Fouch (2012), FuncLab: A matlab interactive toolbox for handling receiver function datasets, *Seismol. Res. Lett*. **83**, 3, 596-603, DOI: 10.1785/gssrl.83.3.596.

Eagar, K.C., M.J. Fouch, D.E. James, and R.W. Carlson (2011), Crustal structure beneath the High Lava Plains of eastern Oregon and surrounding regions from receiver function analysis, *J. Geophys. Res*. **116**, B2, B02313, DOI: 10.1029/2010JB007795.

Efron, B., and R. Tibshirani (1986), Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy, *Stat. Sci.* **1**, 1, 54-77.

Efron, B., and R. Tibshirani (1991), Statistical data analysis in the computer age, *Science* **253**, 5018, 390-395, DOI: 10.1126/science.253.5018.390.

Engdahl, E.R., R. Hilst, and R. Buland (1998), Global teleseismic earthquake relocation with improved travel times and procedures for depth determination, *Bull. Seismol. Soc. Am*. **88**, 3, 722-743.

Fleury, J.M., M. Pubellier, and M. Urreiztieta (2009), Structural expression of forearc crust uplift due to subducting asperity, *Lithos* **113**, 1-2, 318-330, DOI: 10.1016/j.lithos.2009.07.007.

Fountain, D.M., and N.I. Christensen (1989), Composition of the continental crust and upper mantle: A review. **In:** L.C. Pakiser Jr. and W.D. Mooney (eds.), *Geophysical Framework of the Continental United States*, Geological Society of America, Memoir 172, Boulder, Colorado, 711-742.

Haig, D.W. (2012), Palaeobathymetric gradients across Timor during 5.7-3.3Ma (latest Miocene-Pliocene) and implications for collision uplift, *Palaeogeogr. Palaeoclim. Palaeoecol.* **331-332**, 50-59, DOI: 10.1016/ j.palaeo.2012.02.032.

Hall, R., and H.R. Smyth (2008), Cenozoic arc processes in Indonesia: Identification of the key influences on the stratigraphic record in active volcanic arcs, *Geol. Soc. Am. Spec. Pap*. **436**, 27-54, DOI: 10.1130/2008.2436(03).

Hamilton, W. (1977), Subduction in the Indonesian region. **In:** M. Talwani and W.C. Pitman (eds.), *Island Arcs, Deep Sea Trenches, and Back-ArcBasins*, Maurice Ewing Series 1, American Geophysical Union 1, Washington, 15-31.

Hamilton, W. (1979), Tectonics of the Indonesian Region, Geological Survey Professional Paper 1078. U.S. Government Printing Office, Washington D.C.

Harris, R. (2011), The nature of the Banda arc-continent collision in the Timor region. **In:** D. Brown and P.D. Ryan (eds.), *Arc-Continent Collision, Frontiers in Earth Sciences*, Springer, Berlin Heidelberg, 163-211, DOI: 10.1007/978-3-540-88558-0_7.

Harris, R.A., and M.G. Audley-Charles (1987), Taiwan and Timor neotectonics: a comparative review, *Mem. Geol. Soc. China* **9**, 45-61.

Helmers, J., J. Sopaheluwaken, S. Tjokrosapoetro, and E.S. Nila (1989), High grade metamorphism related to peridotite emplacement near Atapupu, Timor, with reference to the Kaibobo peridotite on Seram, Indonesia, *Neth. J. Sea Res.* **24**, 2-3, 357-371, DOI: 10.1016/0077-7579(89)90161-0.

Ishikawa, A., Y. Kaneko, A. Kadarusman, and T. Ota (2007), Multiple generations of forearc mafic-ultramafic rocks in the Timor-Tanimbar ophiolite, eastern Indonesia, *Gondwana Res.* **11**, 1-2, 200-217, DOI: 10.1016/j.gr.2006.04.007.

Jones, C.H., and R.A. Phinney (1998), Seismic structure of the lithosphere from teleseismic converted arrivals observed at small arrays in the southern Sierra Nevada and vicinity, California, *J. Geophys. Res*. **103**, 10065-10090, DOI: 10.1029/97JB03540.

Julia, J., M. Assumpcao, and M.P. Rocha (2008), Deep crustal structure of the Parana´ Basin from receiver functions and Rayleigh-wave dispersion: Evidence for a fragmented cratonic root, *J. Geophys. Res*. **113**, B8, B08318, DOI: 10.1029/2007JB005374.

Kaye, S.J. (1989), The structure of Eastern Indonesia: an approach via gravity and other geophysical methods, Ph.D. Thesis, University of London.

Kaypak, B. (2008), Three-dimensional VP and VP/VS structure of the upper crust in the Erzincan basin (eastern Turkey), *J. Geophys. Res*. **113**, B7, B07307, DOI: 10.1029/2006JB004905.

Keep, M., I. Longley, and R. Jones (2003), Sumba and its effect on Australia's north-western margin, *Geol. Soc. Austral. Spec. Publ.* **372**, 309-318, DOI: 10.1130/0-8137-2372-8.309.

Koch, M. (1992), Bootstrap inversion for vertical and lateral variations of the S wave structure and the Vp/Vs-ratio from shallow earthquakes in the

Rhinegraben seismic zone, Germany, *Tectonophysics* **210**, 1-2, 91-115, DOI: 10.1016/0040-1951(92)90130-X.

Kodaira, S., T. Iidaka, A. Kato, J.-O. Park, T. Iwasaki, and Y. Kaneda (2004), High Pore fluid pressure may cause silent slip in the Nankai Trough, *Science* **304**, 5675, 1295-1298, DOI: 10.1126/science.1096535.

Levin, V., and J. Park (1998), P-SH conversions in layered media with hexagonally symmetric anisotropy: a cookbook, *Pure Appl. Geophys*. **151**, 669-697, DOI: 10.1007/s000240050136.

Ligorría, J.P., and C.J. Ammon (1999), Iterative deconvolution of teleseismic seismograms and receiver function estimation, *Bull. Seismol. Soc. Am*. **89**, 5, 1395-1400.

Linkimer, L., S.L. Beck, S.Y. Schwartz, G. Zandt, and V. Levin (2010), Nature of crustal terranes and the Moho in northern Costa Rica from receiver function analysis, *Geochem. Geophys*. **11**, Q01S19, DOI: 10.1029/2009GC002795.

Macpherson, K.A., D. Hidayat, and S.H. Goh (2012), Receiver function structure beneath four seismic stations in the Sumatra region, *J. Asian Earth Sci.* **46**, 161-176, DOI: 10.1016/j.jseaes.2011.12.005.

Matsubara, M., K. Obara, and K. Kasahara (2008), Three-dimensional P- and S-wave velocity structure beneath the Japan Islands derived from the high-density seismic stations by seismic tomography, *Tectonophysics* **454**, 1-4, 86-103, DOI: 10.1016/j.tecto.2008.04.016.

McCaffrey, R. (1988), Active tectonics of the eastern Sunda and Banda arcs, *J. Geophys. Res*. **93**, B12, 15163-15182, DOI: 10.1029/JB093iB12p15163.

Milsom, J., and M.G. Audley-Charles (1986), Post-collision isostatic readjustment in the southern Banda Arc. **In:** M.P. Coward, and A.C. Ries (eds.), *Collision and Tectonics*, Geological Society, Spec. Pub., Vol. 19, 353-364, DOI: 10.1144/GSL.SP.1986.019.01.20.

Mosalve, H., J.F. Pacheco, C.A. Vargas, and Y.A. Morales (2013), Crustal velocity structure beneath the western Andes of Colombian using receiver-function inversion, *J. South Am. Earth Sci.* **48**, 106-122, DOI: 10.1016/j.jsames.2013.09.001.

Owens, T.J. (1984), Determination of crustal and upper mantle structure from analysis of broadband teleseismic P-waveforms. Ph.D. Thesis, Department of Geology and Geophysics, The University of Utah, USA.

Planert, L., H. Kopp, E. Lueschen, C. Mueller, E.R. Flueh, A. Shulgin, Y. Djajadihardja, and A. Krabbenhoeft (2010), Lower plate structure and upper plate deformational segmentation at the Sunda-Banda arc transition, Indonesia, *J. Geophys. Res*. **115**, B8, B08107, DOI: 10.1029/2009JB006713.

Prasetyo, H. (1992), The Bali-Flores Basin, geological transition from extensional to subsequent compressional deformation. **In:** *Proc. 21st Annual Convention Indonesian Petroleum Association*, 455-478.

Richardson, A.N., and D.J. Blundell (1996), Continental collision in the Banda arc, **In:** R. Hall and D.J. Blundell (eds.), *Tectonic Evolution in Southeast Asia*, Geological Society of London, Spec. Publ. **106**, 47-60, DOI: 10.1144/GSL.SP.1996.106.01.05.

Rutherford, E., K. Burke, and J. Lytwyn (2001), Tectonic history of Sumba Island, Indonesia, since the Late Cretaceous and its rapid escape into forearc in the Miocene, *J. Asian Earth Sci.* **19**, 4, 453-479, DOI: 10.1016/S1367-9120(00)00032-8.

Sambridge, M. (1999), Geophysical inversion with a neighbourhood algorithm– I. Searching a parameter space, *Geophys. J. Int.* **138**, 2, 479-494, DOI: 10.1046/j.1365-246X.1999.00876.x.

Sandwell, D.T., and W.H.F. Smith (2009), Global marine gravity from retracked Geosat and ERS-1 altimetry: Ridge Segmentation v. spreading rate, *J. Geophys. Res* **114**, B01411, DOI: 10.1029/2008JB006008.

Shulgin, A., H. Kopp, C. Mueller, E. Lueschen, L. Planert, M. Engels, E.R. Flueh, A. Krabbenhoeft, and Y. Djajadihardja (2009), Sunda-Banda arc transition: Incipient continent-island arc collision (northwest Australia), *Geophys. Res. Lett.* **36**, 10, L10304, DOI: 10.1029/ 2009GL037533.

Spasojevic, S., and R.W. Clayton (2008), Crustal structure and apparent tectonic underplating from receiver function analysis in South Island, New Zealand, *J. Geophys. Res*. **113**, B4, B04307, DOI: 10.1029/2007JB005166.

Wang, H.Q., A. Schubnel, J. Fortin, E.C. David, Y. Gueguen, and H.K. Ge (2012), High Vp/Vs ratio: Saturated cracks or anisotropy effects?, *J. Geophys. Res*. **39**, 11, L11307, DOI: 10.1029/2012GL051742.

Yeck, W.L., A.F. Sheehan, and V. Schulte-Pelkum (2013), Sequential H-k Stacking to obtain accurate crustal thicknesses beneath sedimentary basins, *Bull. Seismol. Soc. Am*. **103**, 3, 2142-2150, DOI: 10.1785/0120120290.

Zandt, G., and C.J. Ammon (1995), Continental crust composition constrained by measurements of crustal Poisson's ratio, *Nature* **374**, 6518, 152-154, DOI: 10.1038/374152a0.

Zheng, T., L. Zhao, and L. Chen (2005), A detailed receiver function image of the sedimentary structure in the Bohay Bay Basin, *Phys. Earth Planet. Int*. **152**, 3, 129-143, DOI: 10.1016/j.pepi.2005.06.011.

Zhu, L., and H. Kanamori (2000), Moho depth variation in southern California from teleseismic receiver functions, *J. Geophys. Res*. **105**, B2, 2969-2980, DOI: 10.1029/1999JB900322.

Received 20 August 2015
Received in revised form 17 December 2015
Accepted 21 December 2015

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2051-2076
DOI: 10.1515/acgeo-2016-0086

Shear Wave Velocity Estimates through Combined Use of Passive Techniques in a Tectonically Active Area

Rajib BISWAS[1] and Saurabh BARUAH[2]

[1]Department of Physics, Tezpur University, Tezpur, Assam, India;
e-mail: rajib@tezu.ernet.in

[2]Geoscience Division, CSIR-NEIST, Jorhat, Assam, India

Abstract

We made an attempt to assess the shear wave velocity values V_S and, to a lesser extent, the V_P values from ambient noise recordings in an array configuration. Five array sites were situated in the close proximity to borehole sites. Shear wave velocity profiles were modeled at these five array sites with the aid of two computational techniques, viz. spatial autocorrelation (SPAC) and H/V ellipticity. Out of these five array sites, velocity estimates could be reliably inferred at three locations. The shear wave velocities estimated by these methods are found to be quite consistent with each other. The computed V_S values up to 30 m depth are in the range from 275 to 375 m/s in most of the sites, which implies prevalence of a low velocity zone at some pocket areas. The results were corroborated by evidence of site geology as well as geotechnical information.

Key words: array recordings, SPAC, ellipticity.

1. INTRODUCTION

Seismic hazard estimation in recent years has received vast attention from all levels, starting from geo-scientists, civil engineers and policy makers (Hartzell *et al.* 1996, Yamanaka *et al.* 1993, 1994, 1998; Köhler *et al.* 2007,

Ownership: Institute of Geophysics, Polish Academy of Sciences;

Papadopoulou-Vrynioti *et al.* 2013, Pavlou *et al*. 2013, Kassaras *et al.* 2015). The interest in this area is motivated by the notion of minimizing damage by accurate hazard estimation, rather than averting it. One of the important steps in hazard estimates is to reliably determine shear wave velocity profiles. This parameter is basically frequency dependent which is dispersive in nature (Seligson 1970). The dispersive character of shear wave velocity can be efficiently exploited to reveal an underlying one dimensional velocity model, pertaining to a specific study area (Borcherdt 1970, Campbell 1976, Ohnberger *et al.* 2004a, b; Herak 2008). Estimation of a shear wave velocity profile at a site of interest is essential towards the assessment of seismic hazard. The estimations have been performed by making use of data accrued through array implementations.

In our recent work, we estimated the site effect of Shillong area through modified method of Nakamura (Biswas and Baruah 2011). More specifically, Biswas *et al.* (2013) investigated attenuation and site effect in Shillong area using microtremors. More recently, mapping of sediment thickness has also been accomplished (Biswas *et al.* 2015). Microtremor data obtained by an array of sensors have been proven an effective tool for estimation of shear wave velocity (Williams *et al.* 2003). The most widely used techniques for the evaluation of shear wave velocity from dispersive velocity curves of microtremor propagation are the spatial auto-correlation technique (Okada 2006), the frequency wavenumber method (Seligson 1970, Bozdag and Kocaoglu 2005) and the *H*/*V* spectral ratio can be modeled by the theoretical ellipticity of layered velocity models (Claprood and Asten 2007a, b). All these three passive techniques possess different methodology for attaining depth profiles. Despite this, certain researchers combined two of the three to determine the V_S profiles (Fäh *et al.* 2003, Picozzi *et al.* 2005, Di Giulio *et al.* 2006, Claprood and Asten 2007a, b; 2010) and fewer articles regarding adoption of all three passive techniques to attain the shear wave velocity structure (Kuo *et al.* 2009, Boore and Asten 2008).

In this article, we also endeavor to exploit these computing schemes to produce a reliable estimate of velocity to depth profile. Here, we first derive velocity curves through spatial autocorrelation technique. Further, to get more refined and validated results, we compare our finding with all sorts of available geophysical data. The one-dimensional velocity models estimated through these techniques are analyzed in terms of frequency band. Finally, the attained velocity structures are compared for consistency with the models obtained from inversion of ellipticities of Rayleigh wave modes computed from single station three-component horizontal-to-vertical ratio method with simultaneous comparison with the available geotechnical information.

2. GEOLOGICAL BACKGROUND

The study area is located within the Shillong Plateau (SP). The Shillong City, with an area coverage of 6430 km^2 and an average elevation of 1000 m has an approximate population of 180 000. The SP with an Archaean gneissic basement and late Cretaceous–Tertiary sediments along its southern margin is bounded by the Brahmaputra river-fault to the north and the Dauki fault to the south (Kayal *et al.* 2006, Kayal 2008, Rao and Rao 2008). The

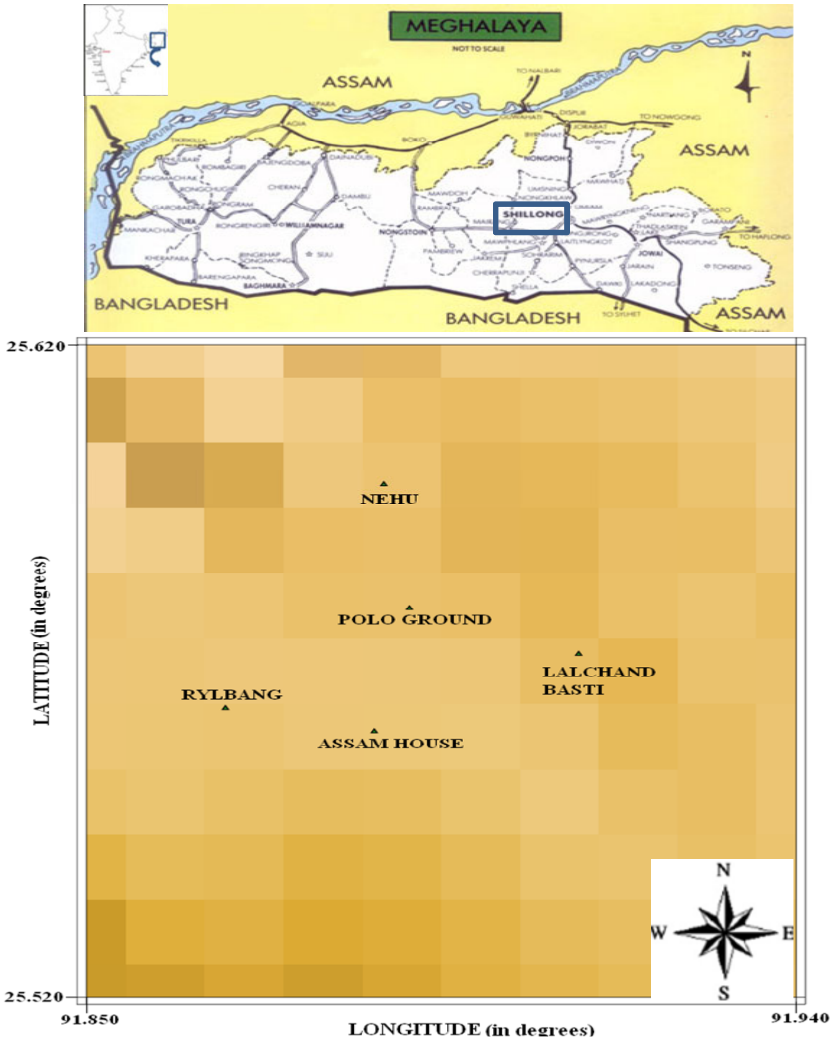


Fig. 1. Locations of the five array sites in Shillong city which is represented by the filled triangles. In inset, map of India is given along with the study area in Meghalaya state.

study area is marked by the Shillong series of parametamorphites, which include mostly quartzites and sandstones, followed by schist, phyllites, slates, *etc.* (GSI 1985). A conglomerate bed containing cobbles and boulders of Archaean crystalline mainly constitutes Shillong series of rocks. The Shillong series grew as depositions in shallow marine conditions over these Archean crystalline rocks (Sar 1973, Mitra and Mitra 2001). The Shillong groups of rocks are intruded by epidiorite rocks, known as Khasi Greenstone as outlined in Fig. 1. The Khasi Greenstone is a group of basic intrusives in the linear to curvilinear form occurring as concordant and discordant bodies within the Shillong group of rocks and suffered metamorphism (Srinivasan *et al.* 1996). These rocks are widely weathered and the degree of weathering is mainly found in the topographic depressions. The metabasic rocks are more prone to weathering than the quartzite rocks. Additionally, the low lying areas are filled with valley fill sediments. Numerous lineaments trend in NE-SW, N-S, and E-W directions in the area (Chattopadhaya and Hashimi 1984).

3. ARRAY CONFIGURATION AND DATA

Array records of ambient noise are used to obtain the shear wave velocity structure of both shallow and deep sedimentary layers. In order to record ambient noise, an array consisting of four sensors was laid out at five selected locations which were located in close proximity to the borehole logs in Shillong City. These five locations were chosen as they cover most characteristic subsurface profiles pertinent to Shillong City (see Table 1 for station locations). For each of them, a homogeneous instrumentation was implemented with good soil sensor coupling (see Bard 2004). It comprised of three Kinematrics Trillium 120P sensors and one short period S-13 Teledyne Geotech, all operating at a sampling rate of 100 samples/s. The time synchronization was provided for each station by GPS receivers. We implemented an equilateral triangular array with three sensors laid out at the

Table 1

Station locations

Station name	Latitude [°]	Longitude [°]	Elevation [m]
Assam House	25.567	91.892	1544
Lalchand Basti	25.590	91.915	1456
Rylbang	25.574	91.880	1562
Nehu	25.611	91.901	1427
Pologround	25.580	91.888	1444

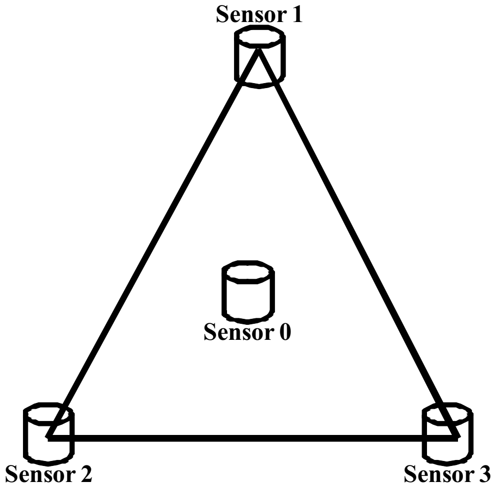


Fig. 2a. Deployed array layouts at the five noise recording site. Here, four sensors are provided. The formation is an equilateral triangle. At the incentre, there remains the central sensor, surrounded by the remaining three sensors. The aperture is 8.67 m whereas the array radius is kept at 5 m.

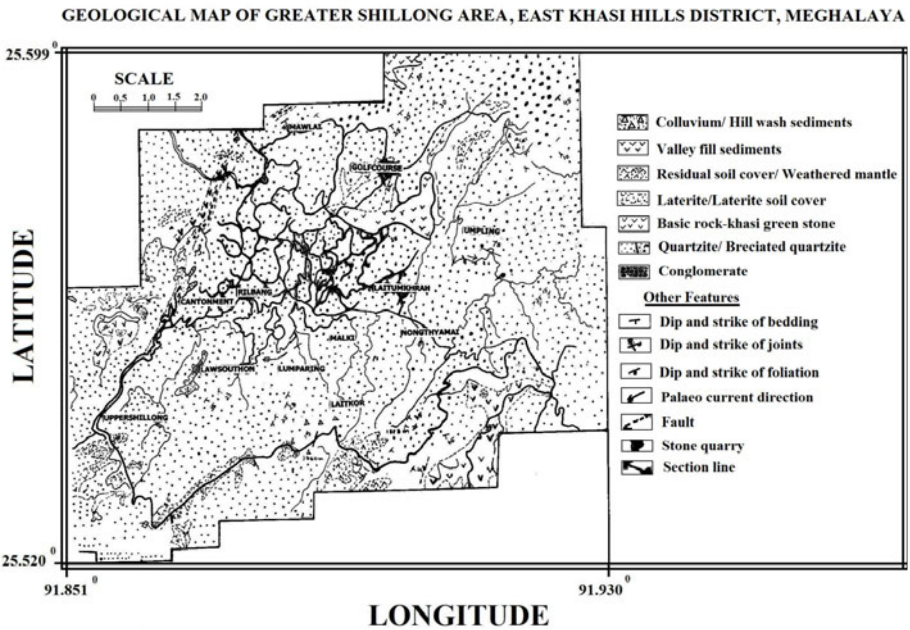


Fig. 2b. Geological map of Shillong City (after GSI 1985). The scale shown in the figure is in kilometres.

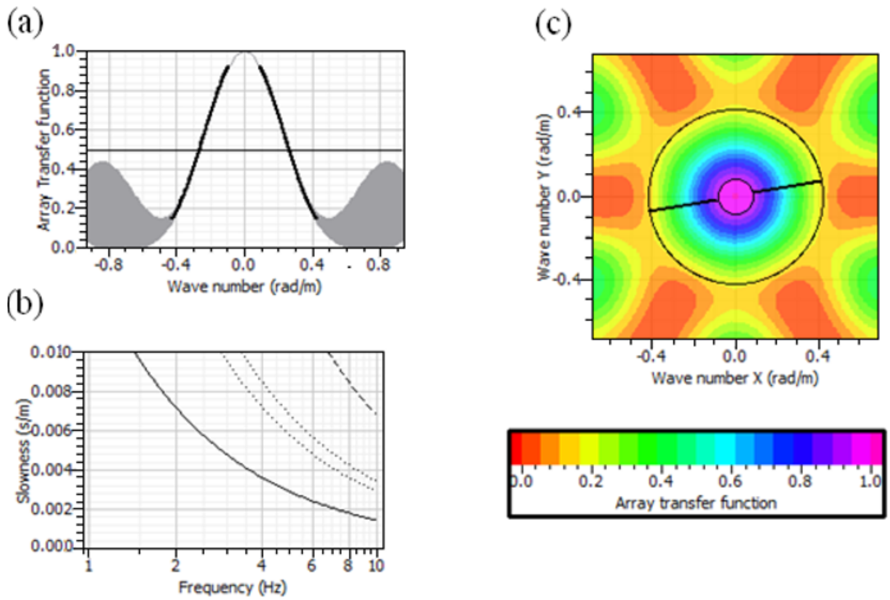


Fig. 3. Theoretical array response for the adopted array layout: (a) Plot of array transfer function *versus* wavenumber. The single peak appearing here corresponds to the main peak. The other side lobes represent the aliasing peaks; (b) Array responses computed for the whole frequency band. k_x and k_y values are plotted along the horizontal and vertical axis, respectively. The color scale indicates the values of array transfer function; (c) Slowness *versus* frequency curve within the defined limits. The four exponential lines represent the constant wavenumber values. $k_{min}/2$ (continuous line), k_{min} (dot-dash line), $k_{max}/2$ (dots), and k_{max} (dashed line).

vertices, while the fourth was placed at the centre of the triangle. Initially, two different arrays were designed with radius of 1 and 5 m. However, surface wave dispersion characteristics obtained by the array of radius of 1 m could not be utilized due to poor resolution in all five sites, while better resolution of dispersion characteristics was achieved for the array of radius of 5 m. Thus, the ambient noise wavefields from 5 m radius array was the prime input towards the estimation of the velocity profile. Consequently, the aperture is kept equal to 8.67 m. Figure 2a shows the adopted layout in selected locations in Shillong City.

The arrays with similar configuration were deployed at five locations, *i.e.*, ASSAM HOUSE, POLOGROUND, NEHU, LALCHAND BASTI, and RYLBANG, as demonstrated in Fig. 2b. The array transfer function, as proposed by Woods and Lintz (1973) is defined within the wave number limits k_{min} and k_{max} and is described in the k_x and k_y plane. Additionally, the corresponding theoretical array response is represented by Fig. 3.

4. SPATIAL AUTOCORRELATION METHOD

The spatial auto-correlation techniques take advantage of the random distribution of sources in time and space to link auto-correlation ratios to phase velocities. In the case of a single-valued phase velocity per frequency band, Aki (1957) demonstrated that these ratios have the shape of Bessel functions of 0 order, the argument of which depends upon the dispersion curve values and the array aperture to reveal the nature of the background seismic noise and also the characteristics of the propagation medium. Bettig *et al.* (2001) brought slightly modified the original formula to extend the method for irregular arrays and urban investigations.

The spatial autocorrelation function of a single plane wave polarized in x direction, $u(x, t)$ for region $x \in [0, X]$ in time domain $t \in [0, T]$ is defined, after Wathelet *et al.* (2004), as follows:

$$\left\langle \varphi(\xi,t) \right\rangle_x = \frac{1}{X} \int_0^X u(x,t)u(x+\xi,t)dx \; . \tag{1}$$

Considering SPAC to be stationary both in space and time, Eq. 1 can be written, after Aki (1957), as

$$\varphi(\xi) = \frac{1}{\pi} \int_0^\infty \varphi(\omega) \cos\left(\frac{\omega}{c(\omega)} \xi \right) d\omega \; , \tag{2}$$

where $\varphi(\omega)$ is the autocorrelation frequency spectrum, ω is the angular frequency, and $c(\omega)$ is the frequency dependent velocity.

From this basic equation, pertaining to the frequency dependent velocity and after adopting the theoretical procedure after Bettig *et al.* (2001), we computed each spectrum pertaining to the array sites deployed at those selective locations. We utilized a window span of ten minutes to evaluate spatial autocorrelation ratio for the respective five sites. After obtaining the dispersion curves, we intend to derive shear wave velocity models, accompanied by V_P values at these five array sites. In addition, the ellipticity peak of fundamental mode of Rayleigh wave corresponding to the estimate of H/V ratio were inverted in order to check the consistency of the inverted results by SPAC. While doing so, as input parameter for the required modeling, we enlist the values of three sites in Tables 2, 3 and 4, respectively, in synchrony with the available borehole information. For the other two sites where no prior information was available regarding depth, the inversion was confined to depth of 30 m only, as given in Table 5. Thus, we attained velocity models for all the five array sites by inverting the dispersion curves of SPAC with the aid of modified neighbourhood algorithom (Sambridge 1999a, b) by Wathelet *et al.* (2004); the results of which are detailed below.

Table 2

Parameterized model for inversion up to a depth of 65 m

Layer	Thickness [m]	V_P [m/s]	V_S/V_P	Poisson's ratio	Density [t/m^3]
Sediments	65 No. of layers 8	200-1275	0.1 to 0.707	0.2 to 0.5	2
Half space	–	2000-3000	0.1 to 0.707	0.2 to 0.5	2

Table 3

Parameterized model for inversion up to a depth of 51 m

Layer	Thickness [m]	V_P [m/s]	V_S/V_P	Poisson's ratio	Density [t/m^3]
Sediments	1 to 51 No. of sub-layers 5	300-1000	0.1 to 0.707	0.2 to 0.5	2
Half space	–	2000-3000	0.1 to 0.707	0.2 to 0.5	2

Table 4

Parameterized model for inversion up to a depth of 100 m

Layer	Thickness [m]	V_P [m/s]	V_S/V_P	Poisson's ratio	Density [t/m^3]
Sediments	1 to 100 No. of sub-layers 5	450-1475	0.1 to 0.707	0.2 to 0.5	2
Half space	–	2250-4200	0.1 to 0.707	0.2 to 0.5	2

Table 5

Parameterized model for inversion for an arbitrary depth of 30 m

Layer	Thickness [m]	V_P [m/s]	V_S/V_P	Poisson's ratio	Density [t/m^3]
Sediments	30 No. of layers 6	200-1000	0.1 to 0.707	0.2 to 0.5	2
Half space	–	2000-3500	0.1 to 0.707	0.2 to 0.5	2

5. VELOCITY PROFILES FROM SPAC FOR EACH ARRAY ASSAM HOUSE

The compressional wave (V_P) and shear wave (V_S) velocity profiles computed for this site are demonstrated in Fig. 4a. The depth of the profile is restricted to 30 m due to lack of *a priori* information of local geology. The shear wave velocity varies between 200 and 400 m/s up to a depth of 30 m, corresponding to the lowest misfit. Gradual increase is observed in V_S, starting from 1 m depth. Corresponding to 1 m stratum of top layer, the shear wave velocity has been estimated at 220 m/s, whereas the V_P value is 500 m/s. In the intermediate layers, with thicknesses of 2 and 4 m, the V_S values are 260 and 310 m/s, respectively. The bottom layer whose thickness is 14 m yields the highest value of shear wave velocity equal to 430 m/s. The dispersion curve yielded by this inversion is also provided in the same figure. The slowness varies between 0.0020 and 0.0032 s/m in the frequency band 0.5 to 10 Hz.

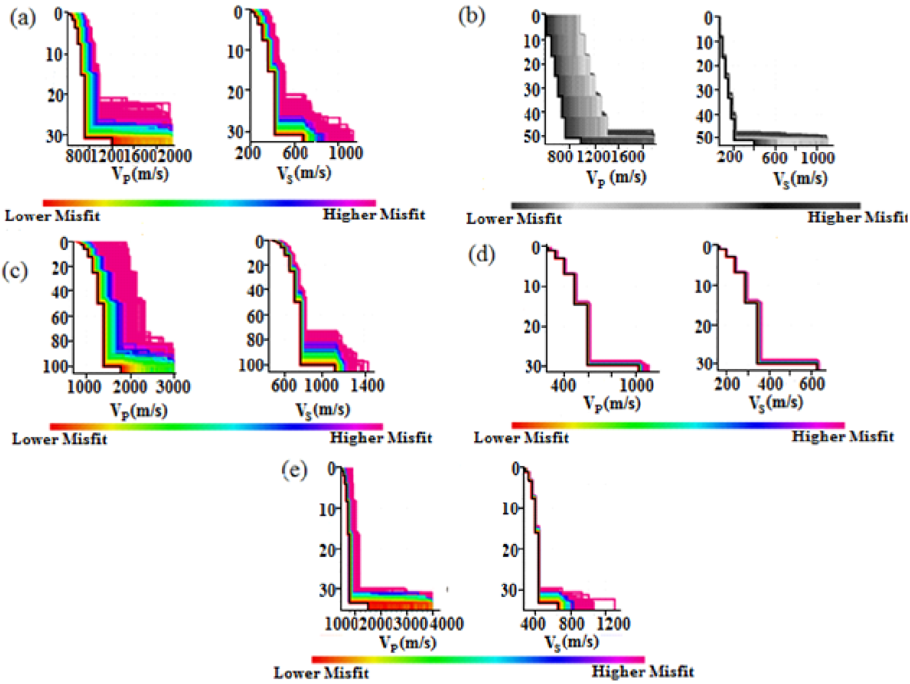


Fig. 4. Shear wave velocity profile estimated from spatial autocorrelation ratios: (a) Assam House, (b) Nehu, (c) Rylbang, (d) Polo ground, and (e) Lalchand Basti.

6. NEHU CAMPUS

This site, as the Assam House, is a plain land formation but with borehole information in its vicinity. Thus, lithological information towards direct inversion of SPAC curves can be incorporated. The velocity profile constrained up to depth of 55 m is illustrated in Fig. 4b. The uppermost layer, having thickness of 8.5 m, produces a very low value of Vs which has been estimated at 120 m/s. As for the V_P, it is found equal to 390 m/s. Towards deeper layers, the shear wave velocity is observed to be slowly rising, reaching a value of 210 m/s for the bottom layer. The same trend is observed to the obtained V_P values. These results indicate a low velocity zone at this site. Concerning the computation of the dispersion curve, the slowness is characterized by higher values for the fundamental mode of Rayleigh waves, starting from 0.004 s/m.

7. RYLBANG

This site is located on the outskirts of Shillong City. It is worth noting that there have been borehole drillings in the immediate neighborhood of this site. Owing to this, *a priori* information can be incorporated in parameterization to acquire a reliable velocity structure underneath this site after computation of SPAC curves. On inverting the SPAC curves, the depth profiles obtained are displayed in Fig. 4c. For the uppermost layer, having a thickness of 4 m, V_S is estimated to be in the range of 385 to 525 m/s. Similarly, V_P is observed to be in the range of 600 to 700 m/s for the same layer. Below, along the estimated profile, Vs increases in regular intervals, a trend also apparent in the estimates of V_P. For the bottom layer, V_S attains a value of 720 m/s.

The dispersion curve covers the frequency band from 2 to 10 Hz. The slowness estimates for the fundamental mode of Rayleigh wave is found to be lower, compared to the previously mentioned sites.

8. POLO-GROUND

Regarding this site, the shear wave velocity profile has been estimated up to depth of 30 m, as illustrated in Fig. 4d. It is evident that the topmost layer reveals a shear wave velocity estimate of 120 m/s. Towards deeper formations, the values of V_S increase up to 325 m/s. A similar trend in estimates of V_P values has also been observed. The V_P is found to be 300 m/s for the uppermost layer, having a thickness of 1 m.

9. LALCHAND BASTI

Figure 4e demonstrates the results of the observed SPAC curves. The uppermost layer which is of 3 m thickness shows a V_S value of 275 m/s, whereas the estimate of V_P value for same layer is 500 m/s. The strata characterized by higher thickness yield the largest V_S value equal to 410 m/s. Similarly, the corresponding value of V_P is estimated to be 700 m/s. The shear wave velocity increases with depth along the profile. The corresponding dispersion curve for the fundamental mode of Rayleigh wave is provided in the same figure. The dispersion curve encompasses slowness estimates of 0.0026 to 0.0032 s/m.

Apart from the inversion of spatial autocorrelation ratios, we have extended our evaluation of shear wave velocity profiles through another robust technique.

10. INVERSION OF *H*/*V* ELLIPTICITY

The H/V ratio generally exhibits a peak that corresponds more or less to the fundamental frequency of the site ($f_0 = V_S/4h$; Bonnefoy-Claudet *et al.* 2004). Initially, the ratio is influenced by the S_H resonance in the superficial layers. If large contribution comes from Rayleigh surface waves, the theoretical ellipticity dictates the observed one as observed by Nogoshi and Igarashi (1970), Fäh *et al.* (2001, 2003), Scherbaum *et al.* (2003). Malischewsky and Scherbaum (2004) developed an analytical formulation for two-layer models. They plotted the differences of the peak frequency between the aforementioned surmises *versus* the magnitude of the velocity contrast. At intermediate and low contrasts, a drastic gap may exist between the two interpreta-tions. Additionally, the observed H/V peak better fits with the extremes of the *SH* transfer function, as inferred by Bonnefoy-Claudet *et al.* (2004). The usefulness of the H/V ratio method has been emphasized in several works. As pointed out by Nakamura (2008), this ratio is capable of yielding reliable estimates of the predominant frequency irrespective of the location of the site involving either ambient noise or earthquake motion as input. The H/V spectrum contains valuable information concerning the underlying structure, such as the relation between the V_S of the sediments and their thickness (Boore and Toksoz 1969, Scherbaum *et al.* 2003). On the other hand, the elliptiicty or S_H transfer function does not provide reliable information related to the site-specific amplification (Wathelet *et al.* 2005). The peak frequency is utilized to serve the objective of attaining the shear wave velocity structure. The V_P profile could be constrained to a good limit by exploiting the ellipticity amplitude.

11. INVERSION RESULTS FROM ELLIPTICITY OF *H*/*V* PEAK

Out of the five array sites, peak resonant frequency has been estimated through the horizontal to vertical ratio methodology only for three sites, namely NEHU campus, Rylbang and Polo-ground. The resonant frequencies for these were previously estimated through single station method (Biswas and Baruah 2011) while for the remaining sites, we lacked resonant frequency estimates. The estimated peak resonant frequencies corresponding to these three sites are inverted in order to obtain the shear wave velocity model. The results for each array site are described below.

12. POLO-GROUND

When the *H*/*V* peak frequency that corresponds to this site is inverted, it produces a shear wave velocity profile characterized by increased values, as illustrated in Fig. 5a. The V_P and V_S values for the top layer, with a thickness of 1 m, are estimated to be 460 and 135 m/s, respectively. The next layer reveals a shear wave velocity of 210 m/s, corresponding to the V_P value of 515 m/s. The shear wave velocity to depth profile has been modeled up to the depth of 31 m for this site. The highest shear wave velocity has been found to be 375 m/s at the bottom stratum whose thickness is ~15 m. The dispersion curve resulting from the inversion shows an increase of slowness values with frequency. The slowness estimates range from 0.001 to 0.005 s/m.

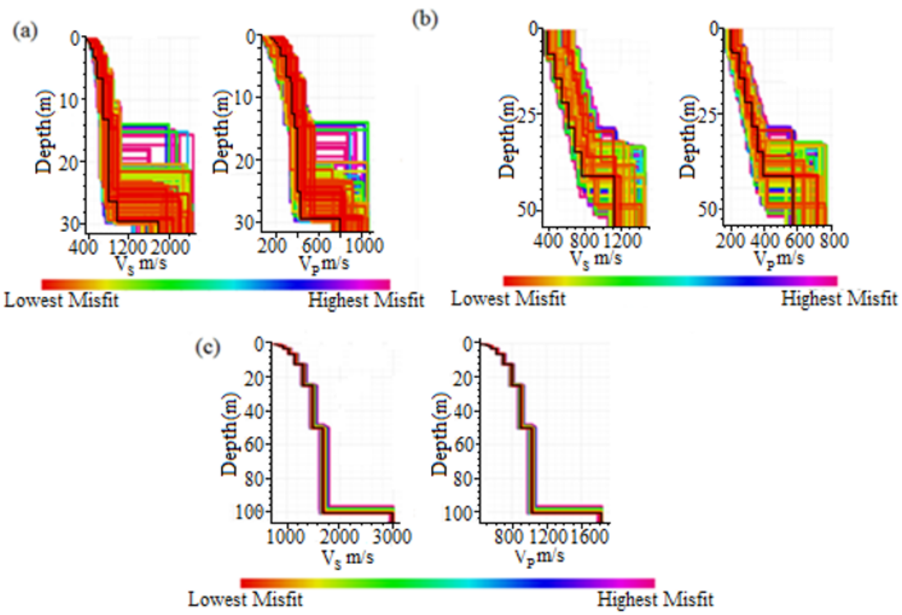


Fig. 5. Shear wave velocity profile estimated from *H*/*V* ellipticity: (a) Polo Ground, (b) Nehu, and (c) Rylbang.

13. NEHU

With the objective of attaining V_P and V_S profiles for NEHU, the corresponding peak frequency has been inverted. The obtained shear wave velocity model is displayed in Fig. 5b. The V_P and V_S values pertaining to the lowest misfit are considered. In this site, the shear wave velocity has been found to be 160 m/s for the top layer with a thickness of 8.5 m. Afterwards, a slow increase in the estimates of shear wave velocity pertaining to the site is observed. Similarly, the V_P values are also found to be very low for this site, having magnitude of 385 m/s for the surface layer.

14. RYLBANG

So far the H/V ratio estimates have been concerned; the fundamental frequency was observed to be at 7 Hz. The inversion results are demonstrated in Fig. 5c. The V_S shows a constant increase revealing a value of 525 m/s corresponding to the top layer, having a thickness of 4 m, whereas the bottom layer of thickness equal to 51 m is characterized by a value of 925 m/s. Concerning V_P, the top layer exhibits a value of 975 m/s, whereas the layer having higher thickness yields a value of 1560 m/s. Here, in accordance with the higher estimates of phase velocities, the curve provides small slowness values, starting from 0.0008 s/m.

All these observations suggest that the results from the inversion of peak resonant frequencies for these array sites are in a general agreement with the shear wave velocity profiles and the dispersion curves. The application of all procedures produces comparable results, irrespective of their variant inherent computation procedures.

15. CORRELATION WITH GEOPHYSICAL AND GEOTECHNICAL PARAMETERS

The estimated shear wave velocity models obtained through these different techniques are compatible. However, this estimation requires validation through correlation with available geophysical and geotechnical information. In this regard, such information like borehole, resistivity and gravity data is compared with the shear wave velocity models estimated for Shillong City.

Concerning geotechnical information, the report on the resistivity profile, available for the Raj-bhaban Area by the Central Ground Water Board (CGWB 2008), Shillong, indicates higher resistivity values at depths of 20-30 m (Fig. 6). The values obtained in these depths are observed to be in the range of 2650 to 6500 ohm-m, as indicated in the table of Fig. 6. These higher values of resistivity imply the existence of stiff soil strata overlain by basement rock (Lay and Wallace 2001), where the shear wave velocity increases with the compactness of the strata. In accordance with the obtained

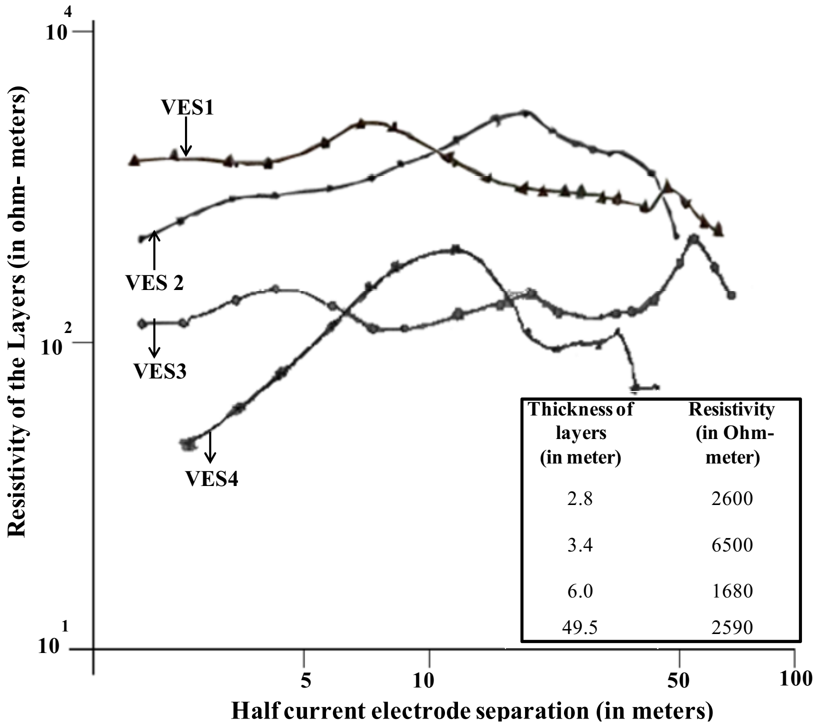


Fig. 6. Resistivity profile of the Rajbhaban Area (after Report by CGWB 2008). In the bottom right corner, the corresponding resistivity values are provided for this site. Resistivity is provided against each layer in ohm-meter units.

results, higher shear wave velocity values are obtained at Rylbang, which is in the vicinity of Rajbhaban Area.

Other geophysical investigations have also been carried out both in and around Shillong City. For example, geophysical investigations, including magmatic, resistivity, and gravity ones, were executed in an area covering 6 km^2 by the Geological Survey of India (GSI 1985, Dasgupta and Biswas 2000, Kalita 1998). Their study reveals a zone of low resistivity and moderate gravity anomaly of 0.3 mGal in the NEHU campus. It is attributed to the existence of a localized region of carbonaceous phylites up to depths of 102 to 139 m. This observation is found to be consistent with the derived shear wave velocity models in this work. According to their conjecture, this region extends up to 3.2 km in the Shillong area. This observation is consistent with the shear wave velocity models derived in this work. Generally, areas yielding low resistivity are considered as low velocity zones, *i.e.*, shear wave velocity decline because of lower porosity and lower density pertinent to stratum. The V_S values estimated by the inversion of the SPAC, corresponding to the NEHU site, are observed to be relatively low, in the range of 150 to 300 m/s. However, towards the other side of the Shillong area, a study of GSI (1985) inferred the abundance of metabasic rocks such as quartzite and phyllites, which present higher resistivity values. In accordance with this observation, the derived model exhibits higher values of shear wave velocities in accordance with the results of geophysical investigation.

Additionally, the shear wave estimates computed through empirical relationship incorporating geotechnical parameters in the form of N values from Standard Penetration Test are found to be in good agreement with the estimates obtained through these inversion techniques (see Biswas *et al.* 2015). As an example, the Lalchand Basti site can be considered. As shown in Fig. 7, the empirically determined (Ohta and Goto 1978) shear wave velocity profile matches reasonably well the models obtained through the inversion. It is evident in this figure that V_S has been empirically estimated to be in the range of 250 to 475 m/s, while in the shear wave velocity models evaluated

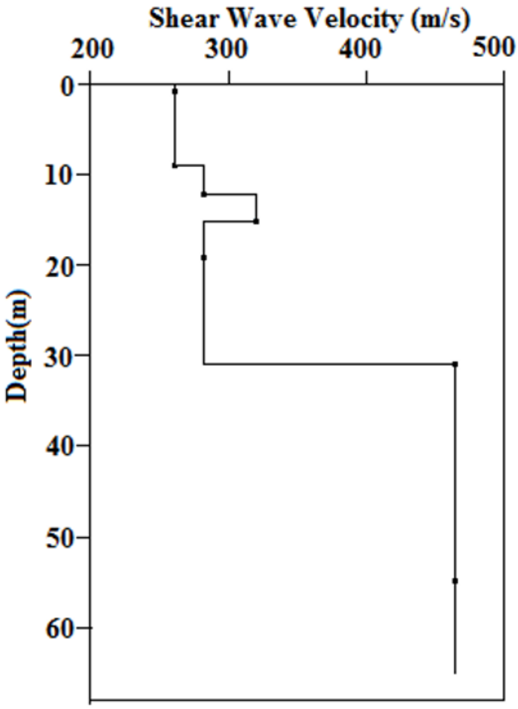


Fig. 7. Shear wave velocity profile for Lalchand Basti site estimated through empirical relationship. Depth is plotted along the vertical axis whereas the empirically computed shear wave velocities are provided along the horizontal axis.

through inversion a similar range of V_S values is obtained, *i.e.*, 245 to 435 m/s. Thus, the derived shear wave velocity models present good correlation with the geotechnical parameters.

Simultaneously, the geological map, as well as the available borehole information, is found to be in good agreement with the estimates of V_P and V_S profiles as obtained by the inversion in the other array sites, such as POLO-GROUND or ASSAM HOUSE.

16. DISCUSSIONS

In this study, we endeavor to assess the shallow shear wave velocity structure of Shillong City with simultaneous determination of subsurface V_P values, to a lesser extent, using ambient noise wavefield. Two different methodologies are adopted, namely the Spatial Autocorrelation Method and the *H*/*V* peak inversion where the phase velocity dispersion curves are inverted to estimate the desired profile. The inverted profiles (V_S and V_P values) are separately determined from the SPAC curves and frequency-wave-number curves for the five sites. These velocity structures, computed using these methods, are compared for consistency with the velocity profile estimated from *H*/*V* peak frequency inversion. The estimated velocity profiles are then correlated with available geophysical information.

The computed SPAC estimates exhibit a varying pattern. In all the five sites, the average autocorrelation ratios are found to be in the range of ± 0.3. As per Okada (2006), three stations or more is the best array deployment towards estimating SPAC coefficients corresponding to fundamental mode. Here, we also constrain the estimates to the first minimum by adoption of four station strategy. The obtained results can be related to the discussion of errors in SPAC coefficient by Henstridge (1979) and Matsuoka *et al.* (1996). However, there is plausibility in augmenting more coherent arrivals for determining SPAC coefficients by adding more stations.

The SPAC ratios in two of the sites, Rylbang and Assam House, reveal high scatter. In a similar work by Asten *et al.* (2004), noticeable scatter in the SPAC curves was observed, which had been attributed to a lack of source azimuthal coverage. In these two sites, the observed high standard deviations could be correlated to shorter time windows which actually generates side-effects. Another plausible explanation of this observation may be related to the insufficient energy at low frequencies (Wathelet *et al.* 2005). However, relatively less scatter is observed in Lalchand Basti, Polo-ground, and Nehu. These sites are characterized by less transients in the evaluated frequency band arising from ambient noise.

Subsequent to the inversion of the autocorrelation curves and ellipticity curves computed through SPAC and *H*/*V*, respectively, the shear wave velocity to depth profile has also been estimated at each of the array sites. Simultaneously, V_P profiles are also estimated to a reliable extent. While deriving the shear wave velocity models, it is emphasized that the incorporation of the lithological information greatly improves the results (Wathelet *et al.* 2005). Out of the five array sites, sufficient geotechnical information in the form of borehole data was available for three sites (Polo-ground, Nehu, and Lalchand Basti). Shear wave velocity models are derived for these three sites based on the available information. Because of the inadequacy of lithological information at the other two sites (Rylbang and Assam House), the shear wave velocity models are interpolated from the available lithological information pertaining to three sites. In these two sites, the estimate of the sediment thickness is not well constrained. Further, for these two sites the results have lower accuracy. Similar observation was also reported by Scherbaum *et al.* (2003).

The computation of the estimated models through SPAC and *H*/*V* technique are found to be quite similar for four sites, except Rylbang. Additionally, the inversion of ellipticity of *H*/*V* peak resonant frequencies at the sites POLO-GROUND, RYLBANG, and NEHU yields shear wave models which are quite consistent with the results obtained through inversion of SPAC. The observations found from *H*/*V* peak inversion are quite comparable with other studies such as Fäh *et al.* (2003) and Bonnefoy-Claudet *et al.* (2008). Concerning the shear wave models, the estimates of V_P and V_S values are found to be of higher magnitude compared to the inversion results from SPAC. As pointed out by Ohnberger *et al.* (2004b), this may be related to the bias in the estimate of *H*/*V* peak frequency, which is the prime input for the inversion process, caused by external perturbation, such as strong wind or rain at the time of ambient noise survey (Hartzell 1992, Milana *et al*. 1996). It should be mentioned that there was continuous rain and wind during the ambient noise survey at Rylbang station.

Bonnefoy-Claudet *et al.* (2006) pointed out that there is a good agreement between the *H*/*V* ratio peak frequency, the fundamental frequency and ellipticity peak frequency of the fundamental Rayleigh Mode. In this study, almost all the inverted models, derived from SPAC curves and ellipticity inversion of *H*/*V* peak frequencies, show estimates of V_P and V_S values which are found to increase with depth. Several researchers reported this trend of subsurface shear wave velocities (Wathelet *et al.* 2004, Scherbaum *et al.* 2003). As in the work of Tokimatsu *et al.* (1992) and Tokimatsu (1997), it was postulated that there existed a mixture of fundamental and higher-mode dispersion curves, which would result in overestimation or underestimation of phase velocity estimates in the dispersion data and could remain undetected by the analyst's quality control. Even though it cannot be completely ruled out that there may be an overestimation of shear wave velocities in the proposed ambient vibration model as obtained in Rylbang, most likely due to bias in the estimation of sediment/bedrock interface depth.

The thickness variations underneath the arrays significantly perturb the dispersion curves, so that the laterally unvarying 1D assumption is too strong and affects the results (Cornou *et al.* 2003, Cho *et al.* 2004). Since the fundamental mode of Rayleigh wave is taken into account due to incorporation only of the vertical component in all the inversion process through these two different methodologies, it is possible that higher modes of surface waves can be present in ambient vibrations, as observed by Tokimatsu (1997) and Zywicki (1999). On the other hand, some small contribution of higher modes at higher frequencies may occur, as indicated by Asten *et al.* (2004). Therefore, the derived shear wave velocity models and V_P values are constrained using only the vertical component for the five array sites. However, it can be stated with emphasis that the shear wave velocity models derived for each respective sites by these two techniques are in conformity.

In addition to the derivation of shear wave velocity models through these techniques, the obtained models are correlated with the available geotechnical and geophysical information. As for the NEHU site, the inverted model is in general agreement with resistivity values, as in the case of the Rylbang site. Apart from this, the shear wave velocity profiles empirically determined

Table 6

Comparison of shear wave velocities from different literature sources

Reference	Site geology	Computed velocity profiles V_S	In this study
Claprood and Asten 2010	Quarternary and Tertiary sediments	400-500 m/s within 0-50 m	Average 250-650 m/s up to 50 m depth
Chávez-Garcia *et al.* 2006	Sediments with clayey soil	530 m/s at 35 m; 700 m/s at 78 m	
Garcia-Jerez *et al.* 2007	Similar	250-750 m/s up to 200 m depth	
Roberts and Asten 2008	Quaternary sediments	200-500 m/s	
Rayhani *et al.* 2008	Same	100-600 m/s up to 35 m depth	
Louie 2001	Quaternary sediments	200-600 m/s	
Asten 2006	Sediments	300-800 m/s up to 280 m depth	

by us, following Ohta and Goto (1978), tally with the computed models. Another validation of the shear wave velocity estimates comes from the comparison of similar work by different researchers, as illustrated in Table 6. In Table 6, the estimates of shear wave velocities computed by them are listed. While presenting their results, it is kept in mind that the shear wave velocity estimates must correspond to the same lithology, irrespective of the geographic location of the sites. All these studies yield values which are found to be in good agreement with the computed shear wave velocities for the study area. Apart from this, it is observed that all SPAC curves yield SPAC ratios that exhibit a decaying pattern when approaching the higher frequency band. Several researchers, *e.g.*, Wathelet *et al.* (2004), Bard (2004), Bettig *et al.* (2001), and Raptakis and Makra (2010), have reported such observations. However, the stratification observed in the velocity profiles estimated by three techniques matches quite well the stratification of the sites Pologround, Rylbang, and Nehu, as observed in the lithology (as, for example, Table 7 highlighting NEHU).

Table 7a

Estimates of parameters for each layer as per spatial autocorrelation technique for array site NEHU

Layer	Thickness [m]	V_P [m/s]	V_S [m/s]
1	8.5	390	120
2	7	485	130
3	9	525	155
4	8	555	170
5	8	595	185
6	8	685	210

Table 7b

Interpretation of each layer as per ellipticity inversion of *H*/*V* peak for array site NEHU

Layer	Thickness [m]	V_P [m/s]	V_S [m/s]
1	8.5	400	110
2	7	480	120
3	9	525	145
4	8	585	168
5	8	645	195
6	8	695	215

Although we did implement these two techniques, which are based upon different working principles, still, each of them is endowed with its merits and demerits. For example, SPAC procedure is basically suitable for spatial and time coherent signals. It can reliably infer one dimensional velocity estimates with a moderate impedance contrast. On the other hand, *H*/*V* method is found to suitably yield dispersion curves for sites which are characterized by a sharp contrast. Since our study area is characterized by sharp to moderate contrast, varying from site to site, we have been encouraged to adopt the joint inversion of dispersion curves yielded by both these methods to attain converging results.

17. CONCLUSION

In this work, we analyze ambient vibrations recorded by three component arrays located in Shillong City. Our aim is to exploit the wave field characteristics of the ambient vibrations and thereby obtain the inverted shear wave velocity models. Two processing methods have been adopted to retrieve the dispersion characteristics from the recorded ambient vibrations. In all cases, the vertical components of the ambient vibrations were processed to obtain the V_S profile. The combined application of two techniques allowed us to attain a suitable shear wave velocity model compatible to the study region (Shillong City). The investigation carried out in this work has shown that it is possible to derive a velocity model of *S*-wave velocity structure for the study area by means of low cost measurements of ambient noise. Our results are consistent with other available data. The computed V_S values up to a depth of 30 m are found to be ~ 375 m/s, in most of the sites. This low value implies prevalence of a low velocity zone at some pocket areas, as supported by evidence of site-geology and geotechnical information. Moreover, in a recent publication (Biswas *et al.* 2015), we have established prevalence of low velocity zones from empirical estimation of observed *H*/*V* resonance frequencies. Thus, the smaller V_S estimates, as found in this study, eventually complements those as reported in Biswas *et al.* (2015).

More input from seismic reflection and refraction studies would enhance the reliability of the results. Despite the lack of *a priori* information of V_P and V_S values at two of the sites, we could retrieve the velocity to depth profiles.

The results provide fair insight towards understanding the subsurface and seismic hazard of the region. A further extension to a 2D and 3D model of the *S*-wave velocity profile obtained in this study would allow for improved knowledge of the interior of the depth profiles, which would be the focus of future research.

References

Aki, K. (1957), Space and time spectra of stationary waves with special reference to micro tremors, *Bull. Earth. Res. Inst. Univ. Tokyo* **35**, 415-456 (in Japanese).

Asten, M.W., T. Dhu, and N. Lam (2004), Optimized array design for microtremor array studies applied to site classification; comparison of results with SCPT logs. **In:** *Proc. 13th World Conf. on Earthquake Engineering, Vancouver, Canada*, Paper No. 2903.

Asten, M.W. (2006), On bias and noise in passive seismic data from finite circular array data processed using SPAC methods, *Geophysics* **71**, 6, V153-V162.

Bard, P.Y. (2004), The SESAME project: an overview and main results. **In:** *13th World Conf. on Earthquake Engineering, August 2004, Vancouver, Canada*, paper No. 2207.

Bettig, B., P.Y. Bard, F. Scherbaum, J. Riepl, F. Cotton, C. Cornou, and D. Hatzfeld (2001), Analysis of dense array noise measurements using the modified spatial auto-correlation method (SPAC): Application to the Grenoble area, *Boll. Geofis. Teor. Appl.* **42**, 3-4, 281-304.

Biswas, R., and S. Baruah (2011), Site response estimation by Nakamura method: Shillong City, Northeast India, *Mem. Geophys. Soc. India* **77**, 173-183.

Biswas, R., S. Baruah, and D.K. Bora (2013), Influence of attenuation and site on microearthquakes' spectra in Shillong region of Northeast India: A case study, *Acta Geophys*. **61**, 4, 886-904, DOI: 10.2478/s11600-013-0129-x.

Biswas, R., S. Baruah, and D.K. Bora (2015), Mapping sediment thickness in Shillong City of northeast India, through empirical relationship, *J. Earthq.* **2015**, 572619, DOI: 10.1155/2015/572619.

Bonnefoy-Claudet, S., C. Cornou, J. Kristek, M. Ohrnberger, M. Wathelet, P.-Y. Bard, P. Moczo, D. Fah, and F. Cotton (2004), Simulation of seismic ambient noise: I. Results of H/V and array techniques on canonical models. **In:** *Proc. 13th World Conf. on Earthquake Engineering.*

Bonnefoy-Claudet, S., F. Cotton, and P.Y. Bard (2006), The nature of the seismic noise wave field and its implication for site effects studies, a literature review, *Earth Sci. Rev*. **79**, 205-227.

Bonnefoy-Claudet, S., F. Leyton, S. Baize, C. Berge-Thierry, L.F. Bonilla, and J. Campos (2008), Potentiality of microtremor to evaluate site effects at shallow depths in the deep basin of Santiago de Chile. **In:** *Proc. 14th World Conf. on Earthquake Engineering, 12-17 October 2008, Beijing, China.*

Boore, D.M., and M.W. Asten (2008), Comparisons of shear-wave slowness in the Santa Clara Valley, California, using blind interpretations of data from invasive and noninvasive methods, *Bull. Seismol. Soc. Am.* **98**, 4, 1983-2003, DOI: 10.1785/0120070277.

Boore, D.M., and M.N. Toksoz (1969), Rayleigh wave particle motion and crustal structure, *Bull. Seismol. Soc. Am.* **59**, 1, 331-346.

Borcherdt, R.D. (1970), Effects of local geology on ground motion near San Francisco Bay, *Bull. Seismol. Soc. Am.* **60**, 1, 29-61.

Bozdag, E., and A.H. Kocaoglu (2005), Estimation of site amplification from shear wave velocity profiles in Yesilyurt and Avcilar, Istanbul by frequency-wavenumber analysis of microtremors, *J. Seismol.* **9**, 1, 87-98, DOI: 10.1007/s10950-005-5271-8.

Campbell, K.W. (1976), A note on the distribution of earthquake damage in Long Beach, 1933, *Bull. Seismol. Soc. Am.* **66**, 3, 1001-1006.

CGWB (2008), Geophysical profiling of Greater Shillong Area, Central Ground Water Board, Shillong (unpublished report).

Chattopadhaya, N., and S. Hashimi (1984), The Sung valley alkaline ultramaffic carbonatite complex, East Khasi Hills district, Meghalaya, *Rec. Geo. Surv. In.* **113**, 4, 24-33.

Chávez-García, F.J., M. Rodríguez, and W.R. Stephenson (2006), Subsoil structure using SPAC measurements along a line, *Bull. Seismol. Soc. Am.* **96**, 2, 729-736, DOI: 10.1785/0120050141.

Cho, I., T. Tada, and Y. Shinozaki (2004), Suggestion and theoretical evaluations on the performance of a new method to extract phase velocities of Rayleigh waves from microtremor seismograms obtained with a circular array. **In:** *13th World Conference on Earthquake Engineering*, Paper No. 647.

Claprood, M., and M.W. Asten (2007a), Use of SPAC, HVSR and strong motion analysis for site hazard study over the Tamar Valley in Launceston, Tasmania. **In:** *Earthquake Engineering in Australia Conference*.

Claprood, M., and M.W. Asten (2007b), Combined use of SPAC, FK and HVSR microtremor survey methods for site hazard study over the Tamar Valley in Launceston, Tasmania. **In:** *ASEG 19th Geophysical Conference and Exhibition, Perth, Australia*, Extended abstracts.

Claprood, M., and M.W. Asten (2010), Statistical validity control on SPAC microtremor observations recorded with a restricted number of sensors, *Bull. Seismol. Soc. Am.* **100**, 2, 776−791, DOI: 10.1785/0120090133.

Cornou, C., P.-Y. Bard, and M. Dietrich (2003), Contribution of dense array analysis to the identification and quantification of basin-edge-induced waves. Part I: Methodology, *Bull. Seismol. Soc. Am.* **93**, 6, 2604-2623, DOI: 10.1785/0120020139.

Dasgupta, A., and A. Biswas (2000), *Geology of Assam*, Geological Society of India, Bangalore.

Di Giulio, G., C. Cornou, M. Ohrnberger, M. Wathelet, and A. Rovelli (2006), Deriving wavefield characteristics and shear velocity profiles from two dimensional small aperture arrays analysis of ambient vibrations in a small size alluvial basin, Colfiorito, Italy, *Bull. Seismol. Soc. Am.* **96**, 5, 1915-1933, DOI: 10.1785/0120060119.

Fäh, D., F. Kind, and D. Giardini (2001), A theoretical investigation of average H/V ratios, *Geophys. J. Int.* **145**, 2, 535-549.

Fäh, D., F. Kind, and D. Giardini (2003), Inversion of local S-wave velocity structures from average H/V ratios, and their use for the estimation of site effects, *J. Seismol.* **7**, 4, 449-467, DOI: 10.1023/B:JOSE.0000005712.86058.42.

Garcia-Jerez, A., M. Navarro, F.J. Alcala, F. Luzon, J.A. Perez Ruiz, T. Enomoto, F. Vidal, and E. Ocana (2007), Shallow velocity structure using joint inversion of array and h/v Spectral ratio of ambient noise, The case of Mula Town (SE of Spain), *Soil Dyn. Earthq. Eng*. **27**, 907-919.

GSI (1985), Geology mapping in greater Shillong area, Meghalaya, Memoir Geolog. Surv. India.

Hartzell, S.H. (1992), Site response estimation from earthquake data, *Bull. Seismol. Soc. Am.* **82**, 6, 2308-2327.

Hartzell, S.H., P. Liu, and C. Mendoza (1996), The 1994 Northridge, California earthquake: Investigation of rupture velocity, rise time, and high-frequency radiation, *J. Geophys. Res.* **101**, B9, 20091-20108, DOI: 10.1029/96JB01883.

Henstridge, D.J. (1979), A signal processing method for circular arrays, *Geophysics* **44**, 2, 179-184, DOI: 10.1190/1.1440959.

Herak, M. (2008), ModelHVSR: a Matlab_ tool to model horizontal-to-vertical spectral ratio of ambient noise, *Comput. Geosci.* **34**, 11, 1514-1526, DOI: 10.1016/j.cageo.2007.07.009.

Kalita, B.C. (1998), Ground water prospects of Shillong Urban Aglomerate, Central Ground Water Board, Meghalaya (unpublished report).

Kassaras, I., D. Kalantoni, Ch. Benetatos, G. Kaviris, K. Michalaki, N. Sakellariou, and K. Makropoulos (2015), Seismic damage scenarios in Lefkas old town (W. Greece), *Bull. Earthq. Eng.* **13**, 12, 3669-3711, DOI: 10.1007/s10518-015-9789-z.

Kayal, J.R. (2008). *Microearthquake Seismology and Seismotectonics of South Asia*. Springer, Dordrecht.

Kayal, J.R., S.S. Arefiev, S. Baruah, D. Hazarika, N. Gogoi, A. Kumar, S.N. Chowdhury, and S. Kalita (2006), Shillong Plateau Earthquakes in northeast India region: Complex tectonic model, *Curr. Sci.* **91**, 1, 109-114.

Köhler, A., M. Ohrnberger, F. Scherbaum, M. Wathelet, and C. Cornou (2007), Assessing the reliability of the modified three-component spatial autocorrelation technique, *Geophys. J. Int.* **168**, 2, 779-796, DOI: 10.1111/j.1365-246X.2006.03253.x.

Kuo, C.H., D.S. Cheng, H.H. Hsieh, T.M. Chang, H.J. Chiang, C.M. Lin, and K.L. Wen (2009), Comparison of three different methods in investigating shallow shear-wave velocity structures in Ilan, Taiwan, *Soil Dyn. Earthq. Eng.* **29**, 1, 133-143.

Lay, T., and T.C. Wallace (2001), *Modern Global Seismology*, Academic Press.

Louie, L.N. (2001), Faster, better: Shear wave velocity to 100 meters depth from refraction microtremor arrays, *Bull. Seismol. Soc. Am.* **91**, 347-364.

Malischewsky, P.G., and F. Scherbaum (2004), Love's formula and H/V-ratio (ellipticity) of Rayleigh waves, *Wave Motion* **40**, 1, 57-67, DOI: 10.1016/j.wavemoti.2003.12.015.

Matsuoka, T., N. Umezawa, and H. Makishima (1996), Experimental studies on the applicability of the spatial autocorrelation method for estimation of geological structures using microtremors, *Butsuri Tansa* **49**, 26-41.

Milana, G., S. Barba, D.E. Pezzo, and E. Zambonelli (1996), Site response from ambient noise measurements: new perspectives from an array study in Central Italy, *Bull. Seismol. Soc. Am.* **86**, 2, 320-328.

Mitra, S., and C. Mitra (2001), Tectonic setting of the precambrians of the north-eastern India, Meghalaya Plateau, and age of Shillong group of rocks, *Geol. Surv. India Spec. Publ.* **64**, 653-658.

Nakamura, Y. (2008), On the H/V Spectrum. **In:** *14th World Conf. on Earthquake Engineering.*

Nogoshi, M., and T. Igarashi (1970), On the propagation characteristics of microtremors, *J. Seismol. Soc. Jap.* **23**, 264-280.

Ohnberger, M., E. Schissele, C. Cornou, M. Wathelet, A. Savvaidis, F. Scherbaum, D. Jongmans, and F. Kind (2004a), Microtremor array measurements for site effect investigations: comparison of analysis methods for field data crosschecked by simulated wavefields. **In:** *13th World Conf. on Earthquake Engineering, 1-6 August 2004, Vancouver, Canada*, Paper No. 0940.

Ohnberger, M., F. Scherbaum, F. Krüger, R. Pelzing, and Sh.-K. Reamer (2004b), How good are shear wave velocity models obtained from inversion of ambient vibrations in the Lower Rhine Embayment (NW-Germany), *Boll. Geof. Teor. Appl.* **45**, 3, 215-232.

Ohta, Y., and N. Goto (1978), Empirical shear wave velocity equations in terms of soil characteristics soil indexes, *Earthq. Eng. Struct. Dyn.* **6**, 167-187.

Okada, H. (2006), Theory of efficient array observations of microtremors with special reference to the SPAC method, *Explor. Geophys.* **37**, 1, 73-84, DOI: 10.1071/EG06073.

Papadopoulou-Vrynioti, K., G. Bathrellos, H. Skilodimou, G. Kaviris, and K. Makropoulos (2013), Karst collapse susceptibility mapping considering peak ground acceleration in a rapidly growing urban area, *Eng. Geol.* **158**, 77-88, DOI: 10.1016/j.enggeo.2013.02.009.

Pavlou, K., G. Kaviris, K. Chousianitis, G. Drakatos, V. Kouskouna, and K. Makropoulos (2013), Seismic hazard assessment in Polyphyto Dam area (NW Greece) and its relation with the "unexpected" earthquake of 13 May 1995 (Ms = 6.5, NW Greece), *Nat. Hazards Earth Syst. Sci.* **13**, 141-149, DOI: 10.5194/nhess-13-141-2013.

Picozzi, M., S. Parolai, and D. Albarello (2005), Statistical analysis of noise horizontal-to-vertical spectral ratios (HVSR), *Bull. Seismol. Soc. Am*. **95**, 5, 1779-1786, DOI: 10.1785/0120040152.

Rao, J.M., and G.V.S.P. Rao (2008), Geology, geochemistry and palaeomagnetic study of cretaceous mafic dykes of Shillong Plateau and their evolutionary history. **In:** R.K. Srivastava, Ch. Sivaji, and N.V. Chalapathi Rao (eds.), *Indian Dykes: Geochemistry, Geophysics and Geomorphology*, Narosa Publishing House, 589-607.

Raptakis, D., and K. Makra (2010), Shear wave velocity structure in western Thessaloniki (Greece) using mainly alternative SPAC method, *Soil Dyn. Earthq. Eng*. **30**, 4, 202-214, DOI: 10.1016/j.soildyn.2009.10.006.

Rayhani, M.H.T., M.H. El Naggar, and S.H. Tabatabi (2008), Nonlinear analysis of local site effects on seismic ground response in the Bam earthquake, *Geotech. Geolog. Eng.* **21**, 1, 91-100.

Roberts, J., and M.W. Asten (2008), A study of near source effects in array-based (SPAC) microtremor surveys, *Geophys. J. Int*. **174**, 1, 159-177, DOI: 10.1111/j.1365-246X.2008.03729.x.

Sambridge, M. (1999a), Geophysical inversion with a neighborhood algorithm. I. Searching a parameter space, *Geophys. J. Int.* **138**, 2, 479-494, DOI: 10.1046/j.1365-246X.1999.00876.x.

Sambridge, M. (1999b), Geophysical inversion with a neighborhood algorithm. II. Appraising the ensemble, *Geophys. J. Int.* **138**, 3, 727-746, DOI: 10.1046/j.1365-246x.1999.00900.x.

Sar, S.N. (1973), An interim report on ground water exploration in the Greater Shillong area, Khasi Hills District, Meghalaya, Memo report, Central Ground Water Board.

Scherbaum, F., K.G. Hinzen, and M. Ohrnberger (2003), Determination of shallow shear wave velocity profiles in the Cologne, Germany area using ambient vibrations, *Geophys. J. Int.* **152**, 3, 597-612, DOI: 10.1046/j.1365-246X.2003.01856.x.

Seligson, C.D. (1970), Comments on high-resolution frequency wavenumber spectrum analysis, *Proc. IEEE* **58**, 6, 947-949, DOI: 10.1109/PROC.1970.7825.

Srinivasan, P., S. Sen, and P.C. Bandopadhaya (1996), Study of variation of Paleocene-Eocene sediments in the shield areas of Shillong Plateau, *Rec. Geol. Surv. India* **129**, 77-78.

Tokimatsu, K. (1997), Geotechnical site characterization using surface waves. **In:** *Proc. 1st Int. Conf. on Earthquake Geotechnical Engineering*, Vol. 3, 1333-1368.

Tokimatsu, K., K. Shinzawa, and S. Kuwayama (1992), Use of short-period micro tremors for V_s profiling, *J. Geotech. Eng*. **118**, 10, 1554-1558, DOI: 10.1061/(ASCE)0733-9410(1992)118:10(1544).

Wathelet, M., D. Jongmans, and M.Ohrnberger (2004), Surface-wave inversion using a direct search algorithm and its application to ambient vibration measurements, *Near Surf. Geophys.* **2**, 4, 211-221, DOI: 10.3997/1873-0604.2004018.

Wathelet, M., D. Jongmans, and M. Ohrnberger (2005), Direct inversion of spatial autocorrelation curves with the neighborhood algorithm, *Bull. Seismol. Soc. Am*. **95**, 5, 1787-1800, DOI: 10.1785/0120040220.

Williams, R.A., W.J. Stephenson, and J.K. Odum (2003), Comparison of P- and S-wave velocity profiles obtained from surface seismic refraction/reflection and downhole data, *Tectonophysics* **368**, 1-4, 71-88, DOI: 10.1016/S0040-1951(03)00151-3.

Woods, J.W., and P.L. Lintz (1973), Plane waves at small arrays, *Geophysics* **38**, 6, 1023-1041, DOI: 10.1190/1.1440393.

Yamanaka, H., M. Dravinski, and H. Kagami (1993), Continuous measurements of microtremors on sediments and basement in Los Angeles, California, *Bull. Seismol. Soc. Am.* **83**, 5, 1595-1609.

Yamanaka, H., M. Takemura, H. Ishida, and M. Niwa (1994), Characteristics of long-period microtremors and their applicability in exploration of deep sedimentary layers, *Bull. Seismol. Soc. Am.* **84**, 6, 1831-1841.

Yamanaka, H., K. Irikura, K. Kudo, H. Okada, and T. Sasatani (1998), Geophysical explorations of sedimentary structures and their characterization. **In:** *Proc. 2nd Int. Symp. on the Effects of Surface Geology on Seismic Motion* **1**, 15-33.

Zywicki, D.J. (1999), Advanced signal processing methods applied to engineering analysis of seismic surface waves, Ph.D. Thesis, Georgia Institute of Technology, Atlanta, USA.

Received 28 April 2015
Received in revised form 3 February 2016
Accepted 8 February 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2077-2091
DOI: 10.1515/acgeo-2016-0090

Variations of Strength, Resistivity and Thermal Parameters of Clay after High Temperature Treatment

Qiang SUN, Weiqiang ZHANG, Yuliang ZHANG, and Lining YANG

School of Resources and Geosciences, China University of Mining and Technology, Xuzhou, Jiangsu Province, P.R. China
e-mails: zhangweiqiang1204@163.com (corresponding author), sunqiang04@126.com

Abstract

This paper reports the variations of strength, resistivity and thermal parameters of clay after high-temperature heating. Experiments were carried out to test the physical properties of clay heated at temperatures ranging from room temperature to 800°C in a furnace. The experiment results show that below 400°C the uniaxial compressive strength and resistivity change very little. However, above 400°C, both increase rapidly. At a temperature under 400°C, the thermal conductivity and specific heat capacity decrease significantly. The thermogravimetric analysis (TG) and differential scanning calorimeter (DSC) test indicate that a series of changes occur in kaolinite at temperatures from 400 to 600°C, which is considered the primary cause of the variation of physical and mechanical properties of clay under high temperatures.

Key words: high temperature, clay, strength, resistivity, thermal parameters.

1. INTRODUCTION

Clay is a common material, with broad applications in geomechanical engineering. Under the influence of high temperature, the mineral composition

Ownership: Institute of Geophysics, Polish Academy of Sciences;

and microstructure of clay change significantly (Sato *et al.* 1992, Aparicio and Galan 1999, Wu and Zhou 2008, Zhu *et al.* 2008). Meanwhile, various physical and mechanical properties are changed in the clay matrix. After being cooled down to room temperature, these thermally-induced variations are to some extent irreversible. Hence, the macroscopic physical and mechanical properties of clay depend on the temperature history, especially the maximum temperature that they were exposed to. Knowledge on the variation of the mechanical and physical properties of clay which are affected by temperature is important to understand and simulate many processes in various engineering projects, such as nuclear water storage (O'Flaherty and Gray 1974, Dixon *et al.* 1985, Radhokrishra and Chan 1989, Witherspoon 2001, Abuel-Naga *et al.* 2009, Sundberg *et al.* 2009, Gens *et al.* 2010, Dupray *et al.* 2013), material modification (Sato *et al.* 1992, Zhang and Yuan 1993, Li *et al.* 2012), pollution controlling (Zheng *et al.* 2010, Sánchez *et al.* 2011), oil or gas industry (Cai 2003).

Numerous studies have shown that some properties of clay or claystone are correlated with thermal damage, such as mass, mechanical strength (*e.g.*, compressive strength), Poisson's ratio, elastic modulus (Laloui and Cekerevac 2003, Zhang 2012, Tian *et al.* 2014, Mao *et al.* 2015), porosity, permeability (Ślizowski *et al.* 2003, Nelskamp *et al.* 2008, Melenevsky *et al.* 2009, Peltonen *et al.* 2009), thermal parameters (Monfared *et al.* 2011), chemical and mineral composition (Sun *et al.* 2015a), and wave velocity (Zhang 2012).

Therefore, the research on the thermo-physical-mechanical of clay is extremely meaningful on a wide range. In this paper, the variations of uniaxial compressive strength, resistivity and thermal parameters are analyzed after high-temperature heating.

2. EXPERIMENTAL TESTS

Clay samples used in these experiments are tan-eluvial soils collected from farmland in Xuzhou, Jiangsu Province. As shown in Fig. 1, the main composition was quartz and kaolinite.

The clay samples were first dried in a high temperature drying oven (keeps 2 hours under the temperature of 105°C), then ground into powder and screened by a 1-mm sieve to remove the impurities. The screened clay was shaped into cylindrical specimens of $\Phi\ 60 \times 70$ mm by adding a small amount of water with a compaction device. The average of moisture content and density of the cylindrical specimens were 5% and 2.11 g/cm^3, respectively. Finally, the samples experienced the several steps of heat treatment by an MTS652.02 high-temperature furnace. The temperature was set into fourteen levels (25, 60, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, and 800°C), with four samples as a set. The heating process was com-

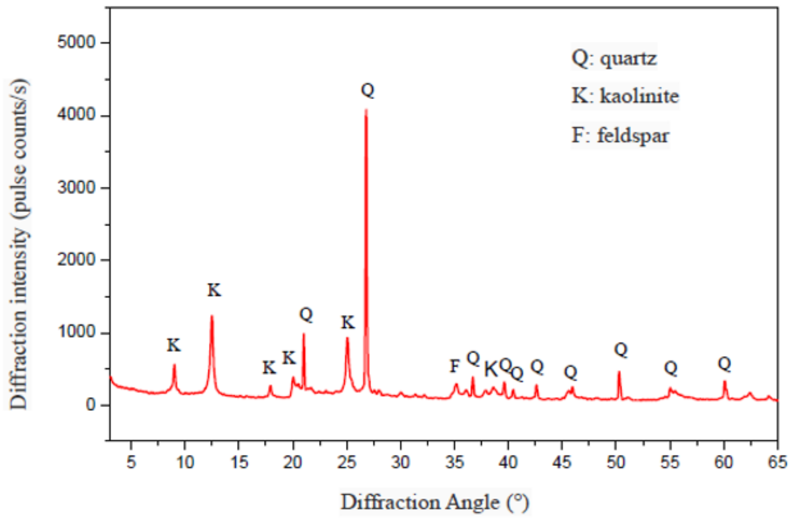


Fig. 1. XRD spectrum of a clay sample.

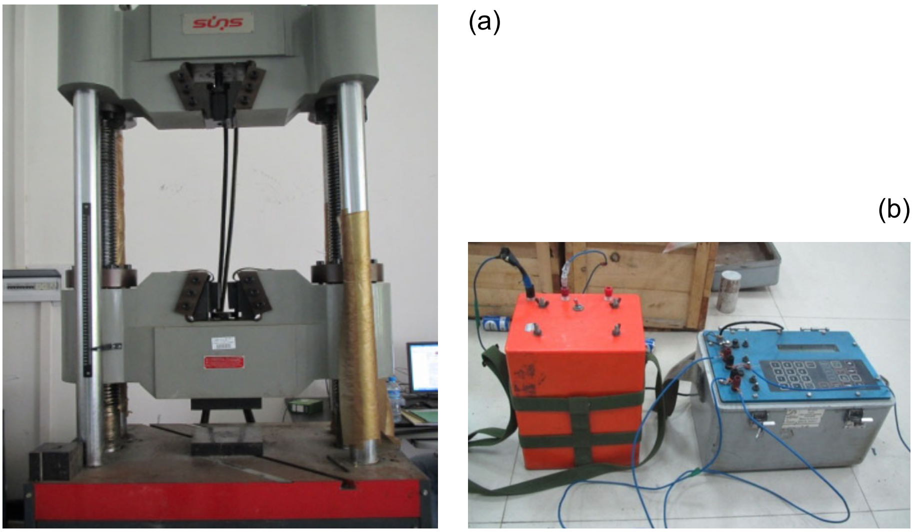


Fig. 2. The testing machine and resistivity detector: (a) loading apparatus, and (b) digital electrical instrument (SYSCAL-R2).

posed of three stages: ① clay specimens were heated in the high-temperature furnace at the heating rate 30°C/min until reaching the targeted temperature; ② the targeted temperature was kept for 3 min to ensure that the inside of the samples was heated evenly; ③ the furnace was cooled down to room temperature by cutting off the power of furnace. The strength, resistivity, and thermal parameters of the specimens were tested before and after heating. The strength was tested with an electro-hydraulic servo-controlled testing machine (EHC-3000, produced by Sansi Instruments Co. Ltd. of Zhuhai, China, shown in Fig. 2a) at the loading speed of 0.5 kN/s. The load was measured by an oil pressure sensor connected to the hydraulic pressure cylinder, and a cylindrical capacitor displacement sensor fixed to the upper plate was used to measure the displacement of clay samples in the test. The electrical resistivity was measured with a digital electrical instrument (SYSCAL-R2, produced by IRIS Instrument Co. Ltd, France, shown in Fig. 2b), and the information on thermal parameters was simultaneously collected by a TPS test machine produced by Hot Disk Equipment Co. Ltd. In addition, the TG curve (thermogravimetric analysis) of clay specimen

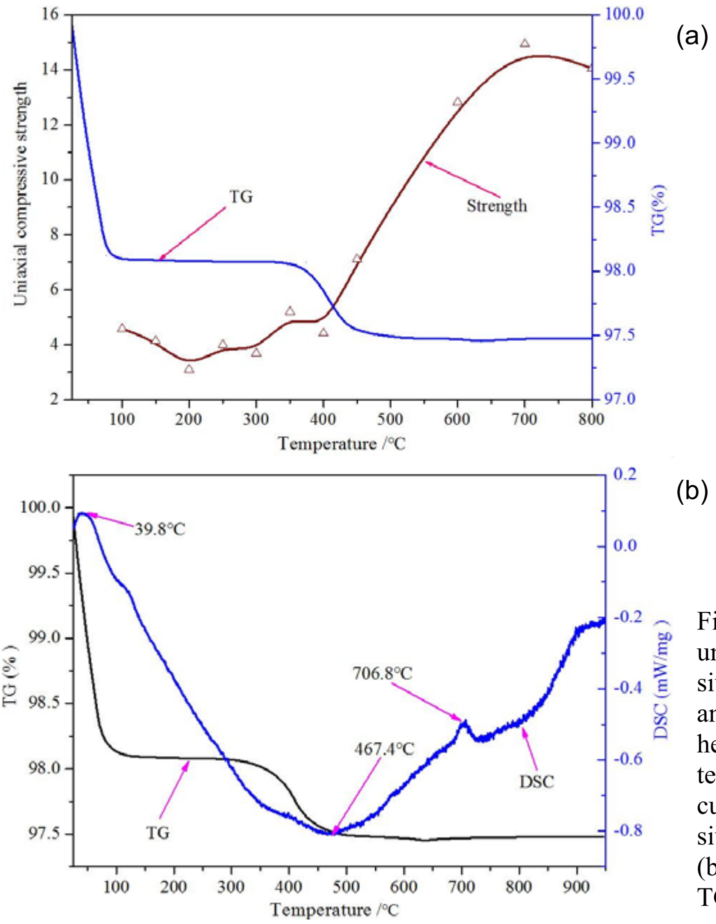


Fig. 3. Variation of uniaxial compressive strength, TG and DSC, after heating at different tempera-tures: (a) curve of compressive strength, and (b) curve of DSC-TG.

(shown in Fig. 3) was tested with a synchronous comprehensive thermal analyzer (type STA409C) produced by NETZSCH Co. Ltd., and in this test, the heating rate was set as 5 k/min and the gas flow rate was 100 mL/min.

3. THE TEST RESULT ANALYSIS

3.1 Variation of uniaxial compressive strength

The variation of uniaxial compressive strength with temperature is shown in Fig. 3. Two stages of variation can be identified: (i) Below 400°C, the strength changes subtly with the increase of temperature (the strength is around 4 MPa); (ii) When temperatures are higher than 400°C, the strength drastically increases with temperature. Combining this with the TG and DSC (differential scanning calorimeter) measurement curves shown in Fig. 4 (Zhu *et al.* 2008) and Fig. 5 (Wu and Zhou 2008), it can be deduced that when the temperature is higher than 400°C, the temperature effect is mainly caused by the alteration of mineral content. From differential thermal analysis (DTA) of clay (as shown by Figs. 4 and 5), the dehydroxylation reaction of clay is obviously verified at the range from 400 to 650°C. Therefore, in the heating process, kaolinite dehydroxylation yielding metakalinite occurs in the 400-650°C range (De Aza *et al.* 2014), as shown by the equations:

$$Al_2O_3 \bullet 2SiO_2 \bullet 2H_2O \rightarrow Al_2O_3 \bullet 2SiO_2 \bullet \frac{1}{2}H_2O + \frac{3}{2}H_2O\,, \tag{1}$$

$$Al_2O_3 \bullet 2SiO_2 \bullet \frac{1}{2}H_2O \rightarrow Al_2O_3 \bullet 2SiO_2 + \frac{1}{2}H_2O\,. \tag{2}$$

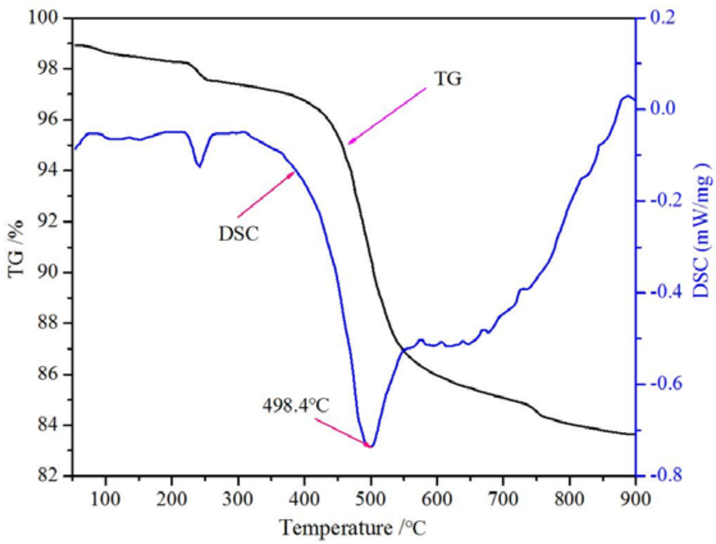


Fig. 4. DSC-TG diagram of Suzhou kaolinite (Zhu *et al.* 2008).

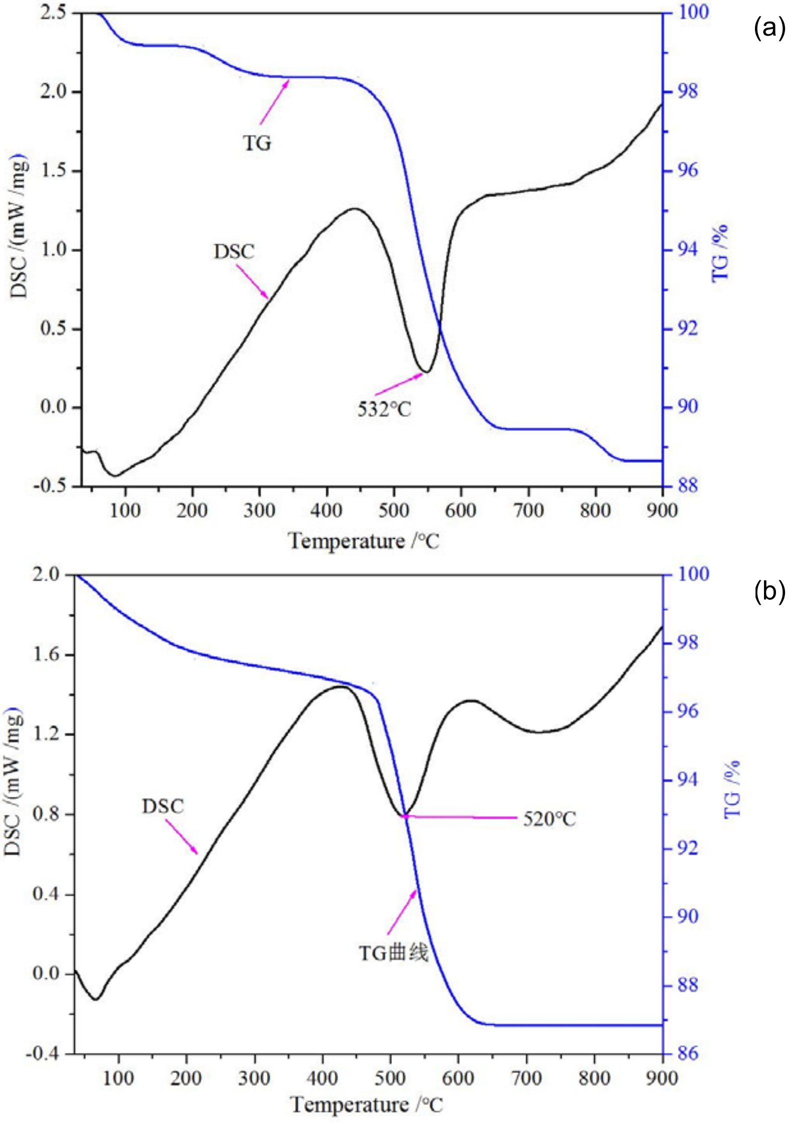


Fig. 5. DSC-TG diagram of kaolinite (Wu and Zhou 2008): (a) Tanwu kaolinite, and (b) Huling kaolinite.

In the heating process, kaolinite loses some hydroxyl (-OH) and the bond between particles is improved (Belloto *et al.* 1995, Lee *et al.* 1999). As a result, the improved cohesion between particles causes the increase of uniaxial compressive strength with temperature.

3.2 Variation of resistivity

The principle of resistivity measurement is shown in Fig. 6: powering both ends of the rock sample and then, using the measuring electrodes at the upper and lower ends, we observe the potential difference between MN when the electrical current passes through the rock sample, so the resistivity can be calculated with the following Eq. 3. As the clay resistivity is relatively large, the intensity of current is too small to be measured directly. It is obtained indirectly by measuring the voltages at both ends of a 1 kΩ standard resistor in the power supply circuit with the switch K (Fig. 6).

$$\rho = K\frac{\Delta U_{\mathrm{MN}}}{I}\,; \quad K = A / L\,, \tag{3}$$

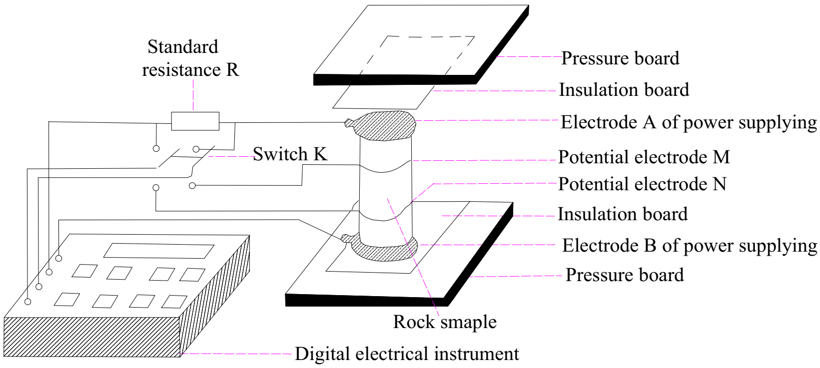


Fig. 6. Schematic diagram of the experimental system for stress and electrical resistance of clay under uniaxial compression (Sun *et al*. 2015b).

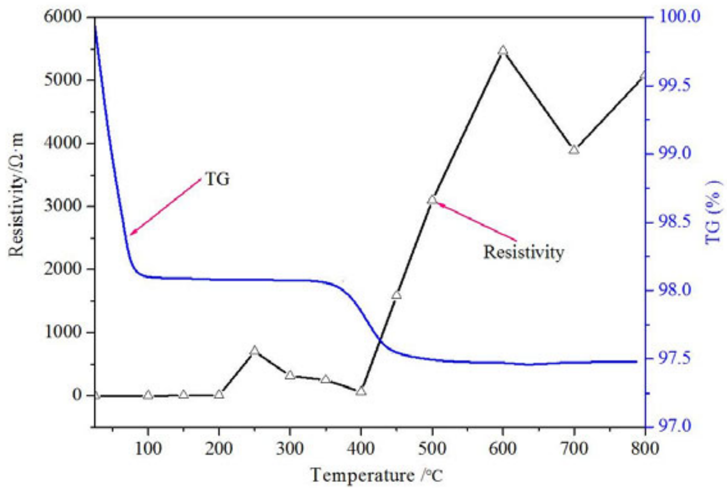


Fig. 7. Variation of specific resistivity at different temperatures.

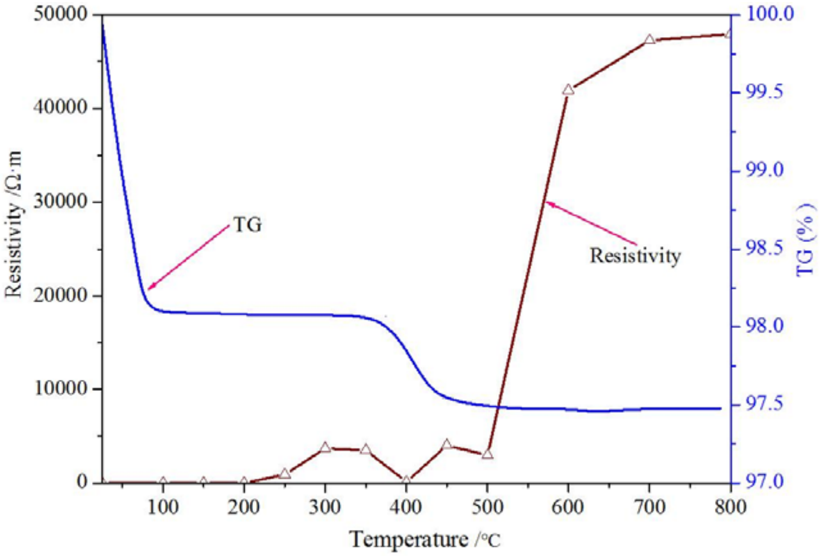


Fig. 8. Variation of resistivity (at the peak strength) at different temperatures.

where K is a coefficient determined by device, A is the cross-sectional area of rock sample; L is the distance between MN; ΔU_{MN} is the potential difference between MN; and I is current.

The variation of resistivity with temperature is shown in Figs. 7 and 8. Below 400°C, the resistivity changes very little with the increase of temperature. When the temperature is higher than 400°C, the resistivity quickly increases with temperature. This pattern of variation suggests that when the temperature is higher than 400°C, the varied thermal effect results from the changing mineral content of the sample (shown by Eqs. 1 and 2).

3.3 Variation of thermal parameters

After high-temperature heating, the thermal conductivity, specific heat capacity and thermal diffusivity of clay specimens change significantly (as shown in Figs. 9-11) in the temperature range of 20-800°C. From 25 to 400°C, conductivity decreases rapidly with temperature, which is mainly due to growing numbers of pores. Above 400°C, the thermal conductivity and specific heat capacity decrease a little bit. However, the thermal diffusivity presents a certain fluctuation, but the variation level is not apparent compared to thermal conductivity and specific heat capacity. It should be pointed out that the properties of clay in this phase (above 400°C) are close to rock. The variation of thermal conductivity with temperature is shown by Fig. 9 and the following Eq. 4

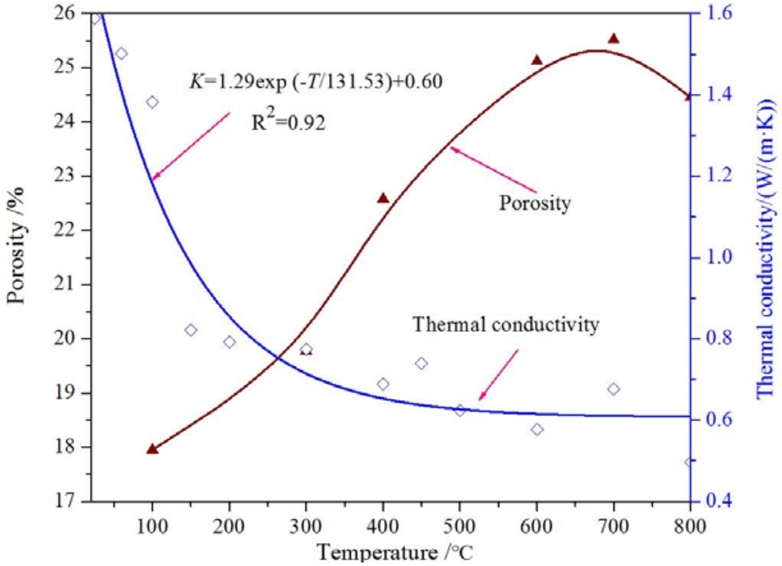


Fig. 9. Variations of thermal conductivity and porosity at different temperatures.

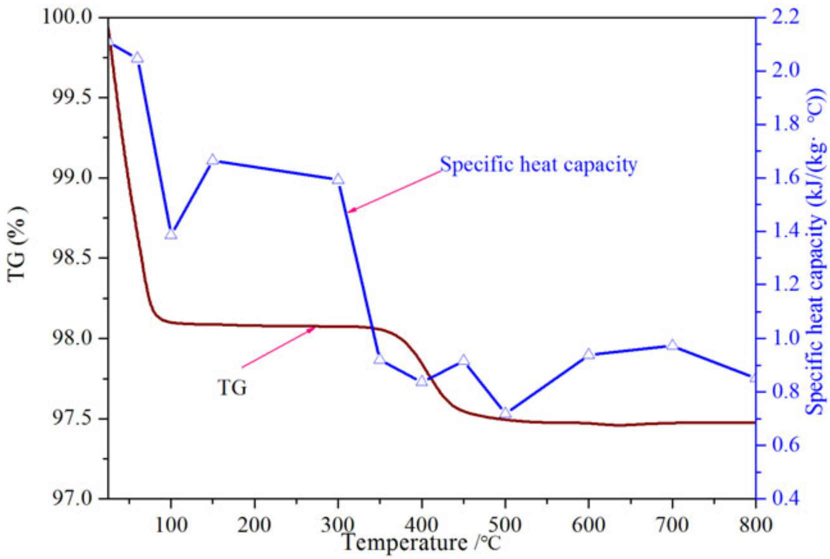


Fig. 10. Variations of specific heat capacity and TG at different temperatures.

$$K = 1.29\exp\left(-\frac{T}{131.53}\right) + 0.60\,; \quad R^2 = 0.92\,, \tag{4}$$

where T *is* temperature, and K is thermal conductivity.

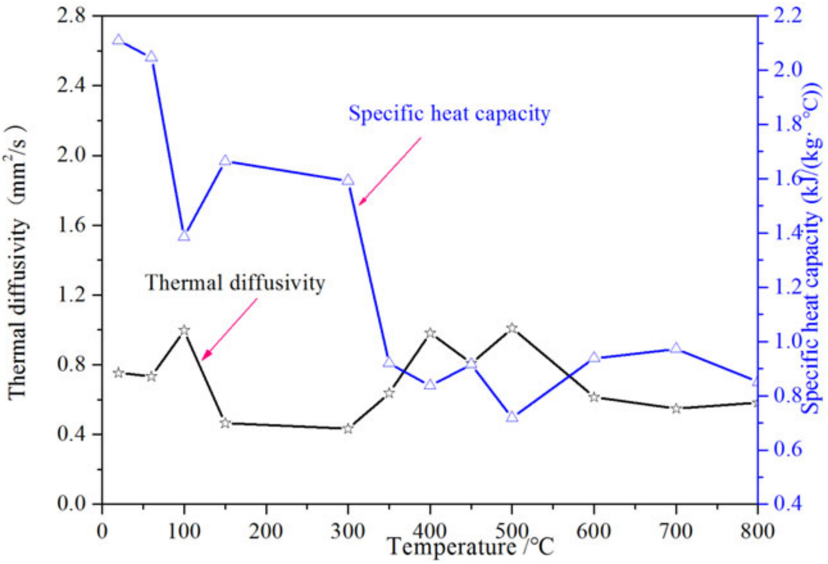


Fig. 11. Variations of specific heat capacity and thermal diffusivity at different temperatures.

4. DISCUSSION

The main mechanism of changes in physical properties of clay after thermal treatment is the damage of clay structure caused by thermal reactions. Generally speaking, clay is composed of different minerals. In the heating progress, there is a series of chemical reactions. Adsorbed and interlayer water can evaporate (*i.e.*, desorption) at the temperature range of 100-250°C, and the structural water is lost when the treatment temperature is higher than 300°C. When the heating temperature is lower than the threshold value, the variation of clay is mainly due to the change of all kinds of water. The existence of water seriously affects the porosity. Another reason: in the progress of losing structural water, some minerals are decomposed and evaporated (for example, the dehydroxylation of clay minerals), which may cause the increase of pore and fracture.

Organic matter is comparatively enriched in muddy sediments (Hunt 1996). Above 300°C, this organic matter begins to turn into char residue. The oxidation/decomposition reaction of organic matters is obviously evidenced in the range from 300°C (especially 400°C) to 500°C. Results of Fig. 12 show that the exothermal peak at 350°C is more strongly marked (Cai 2003, Cai *et al.* 2007), while the exothermal peak above 350°C disappeared. It indicates that organic matter is incorporated with clay minerals closely and stably. Therefore, oxidation/decomposition reaction of organic

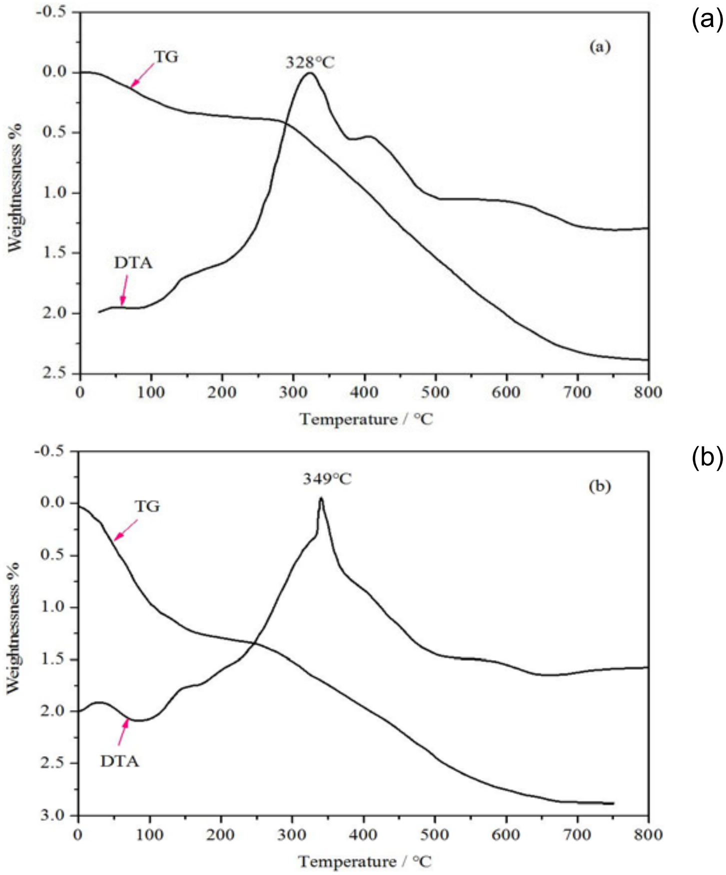


Fig. 12. The TG and DTA curves of clay samples No. 26 (a) and No. 3 (b) (Cai *et al.* 2007).

matters is obviously evidenced in the period ranging from 300 to 500°C, which leads to decrease of mass and increase of porosity. When the heating temperature reached the critical value, some of the carbonate minerals (such as calcium carbonate, magnesium carbonate, magnesite, dolomite) starts to decompose and results in the decrease of elastic modulus, compressive strength, tensile strength, and fracture toughness. Moreover, when the treatment temperature is higher than 400°C, the oxidation of organic matter is quickly increasing.

5. CONCLUSION

The study of characteristics of strength, resistivity, and thermal parameters of clay heated at different temperatures is reported in this paper. Based on the results, the following conclusions can be drawn:

- Temperature has a significant impact on the physical and mechanical properties of clay;
- Between 100 and 400°C, the uniaxial compressive strength and resistivity of clay change very little with temperature. Above 400°C, the strength and resistivity rise rapidly. The variations of uniaxial compressive strength and resistivity are significantly affected by altered mineral clay composition between 400 and 600°C;
- The thermal conductivity and specific heat capacity of clay specimens change significantly after treatment at high temperatures. From 20 to 400°C, they decrease rapidly with temperature. Above 400°C, both thermal conductivity and specific heat capacity decrease slightly.

Finally, it is worth mentioning that these characteristics of variation of clay properties were observed in laboratory tests. Permeability and ultrasonic measurement will also be included in future work.

Acknowledgments. This research was supported by "the Fundamental Research Funds for the Central Universities" (No. 2015XKMS033) and the Priority Academic Program Development of Jiangsu Higher Education Institutions.

References

Abuel-Naga, H.M., D.T. Bergado, A. Bouazza, and M.J. Pender (2009), Thermal conductivity of soft Bangkok clay form laboratory and field measurements, *Eng. Geol.* **105**, 3-4, 211-219, DOI: 10.1016/j.enggeo.2009.02.008.

Aparicio, P., and E. Galan (1999), Mineralogical interference on kaolinite crystallinity index measurements, *Clay. Clay Miner.* **47**, 1, 12-27, DOI: 10.1346/CCMN.1999.0470102.

Belloto, M., A. Gualtieri, G. Artioli, and S.-M. Clark (1995), Kinetic study of the kaolinite-mullite reaction sequence. Part II: Mullite formation, *Phys. Chem. Minerals* **22**, 4, 215-222, DOI: 10.1007/BF00202254.

Cai, J.G. (2003), Oregano-clay complexes in muddy sediments and mudstones, Ph.D. Thesis, Tongji University, Shanghai (in Chinese).

Cai, J.G., Y.J. Bao, S.Y. Yang, X.X. Wang, D.D. Fan, J.L. Xu, and A.P. Wang (2007), Research on preservation and enrichment mechanisms of organic matter in muddy sediment and mudstone, *Sci. China D* **50**, 5, 765-775, DOI: 10.1007/s11430-007-0005-0.

De Aza, A.H., X. Turrillas, M.A. Rodriguez, T. Duran, and P. Pena (2014), Time-resolved powder neutron diffraction study of the phase transformation sequence of kaolinite to mullite, *J. Eur. Ceram. Soc*. **34**, 5, 1409-1421, DOI: 10.1016/j.jeurceramsoc.2013.10.034.

Dixon, D.A., M.N. Gray, and A.W. Thomas (1985), A study of the compaction properties of potential clay-sand buffer mixtures for use in nuclear fuel waste disposal, *Eng. Geol.* **21**, 3-4, 247-255, DOI: 10.1016/0013-7952(85)90015-8.

Dupray, F., C. Li, and L. Laloui (2013), Thermal conductivity of soft Bangkok clay form laboratory and field measurements, *Eng. Geol*. **163**, 113-121, DOI: 10.1016/j.enggeo.2013.05.019.

Gens, A., L. do Guimarães, S. Olivella, and M. Sánchez (2010), Modelling thermo-hydro-mechano-chemical interactions for nuclear waste disposal, *J. Rock Mech. Geotech. Eng*. **2**, 2, 97-102, DOI: 10.3724/SP.J.1235.2010.00097.

Hunt, J.M. (1996), *Petroleum Geochemistry and Geology*, 2nd ed., W.H. Freeman and Co., New York, 100 pp.

Laloui, L., and C. Cekerevac (2003), Thermo-plasticity of clays an isotropic yield mechanism, *Comp. Geotech.* **30**, 8, 649-660, DOI: 10.1016/j.compgeo.2003.09.001.

Lee, S., Y.J. Kim, and H.S. Moon (1999), Phase transformation sequence from kaolinite to mullite investigated by an energy-filtering transmission electron microscope, *J. Am. Ceram. Soc.* **82**, 10, 2841-2848, DOI: 10.1111/j.1151-2916.1999.tb02165.x.

Li, Y., Q.C. Yu, B. Yang, and Y. Dai (2012), Characterization of vacuum thermal decomposed kaolin vacuum, *Chin. J. Vacuum Sci. Tech*. **32**, 599-604 (in Chinese).

Mao, R.R., X.B. Mao, L.Y. Zhang, and R.X. Liu (2015), Effect of loading rates on the characteristics of thermal damage for mudstone under different temperatures, *Int. J. Min. Sci. Technol*. **25**, 5, 797-801, DOI: 10.1016/j.ijmst.2015.07.015

Melenevsky, V.N., A.E. Kontorovich, and W.L. Huang, A.I. Larichev, and T.A. Bul'bak (2009), Hydrothermal pyrolysis of organic matter in Riphean mudstone, *Geochem. Int.* **47**, 5, 476-484, DOI: 10.1134/S0016702909050048.

Monfared, M., J. Sulem, P. Delage, and M. Mohajerani (2011), A laboratory investigation on thermal properties of the opalinus claystone, *Rock Mech. Rock Eng*. **97**, 735-747, DOI: 10.1007/s00603-0110-0171-4.

Nelskamp, S., P. David, and R. Littke (2008), A comparison of burial, maturity and temperature histories of selected wells from sedimentary basins in the Netherlands, *Int. J. Earth Sci*. **97**, 5, 931-953, DOI: 10:1007/s00531-007-0229-4.

O'Flaherty, C.A., and M.N. Gray (1974), The influence of alkali compounds on the compaction and early strength properties of lime-soil mixtures, *Austral. Road Res.* **5**, 5, 4-15.

Peltonen, C., Ø. Marcussen, Bjørlykke, and J. Jahren (2009), Clay mineral diagenesis and quartz cementation in mudstones: The effects of smectite to illite reaction on rock properties, *Mar. Petrol. Geol.* **26**, 6, 887-898, DOI: 10.1016/j.marpelgeo.2008.01.021.

Radhokrishra, H.S., and H.T. Chan (1989), Thermal and physical properties of candidate buffer-backfill material for a nuclear fuel waste disposal vault, *Can. Geotech*. **26**, 6, 629-639, DOI: 10.1016/0148-9062(90)92830-8.

Sánchez, M., A. Shastri, and A. Gens (2011), Transient behavior of a clay barrier subjected to high temperature changes, *Geo-Frontiers* **2011**, 4156-4165, DOI: 10.1061/41165(397)425.

Sato, T., T. Watanable, and Otsuka (1992), Effects of layer charge location and energy change on expansion properties of dioctahedral smectite, *Clay. Clay Miner*. **40**, 1, 103-113, DOI: 10.1346/CCMN.1992.0400111.

Ślizowski, K., J. Janeczek, and K. Przewłocki (2003), Suitability of salt-mudstones as a host rock in salt domes for radioactive-waste storage, *Appl. Energ*. **75**, 1-2, 119-128, DOI: 10.1016/S0140-6701(04)91754-7.

Sun, L.N., Z.N. Zhang, Y.D. Wu, L. Su, Y.Q. Xia, Z.D. Gao, Y.W. Zheng, and Z.X. Wang (2015a), Effect of temperature and pressure on hydrocarbon yield of source rock HTHP simulation experiment in semi-open system, *Nat. Gas. Geosci*. **26**, 1, 118-127, DOI: 10.11764/j.issn.1672-1926.2015.01.0118 (in Chinese).

Sun, Q., S.Y. Zhu, and L. Xue (2015b), Electrical resistivity variation in uniaxial rock compression, *Arab. J. Geosci*. **8**, 4, 1869-1880, DOI: 10.1007/s12517-014-1381-3.

Sundberg, J., P.E. Back, R. Christiansson, H. Hökmark, M. Ländell, and J. Wrafter (2009), Modeling of thermal rock mass properties at the potential sites of a Swedish nuclear waste repository, *Int. J. Rock Mech. Min. Sci*. **46**, 6, 1042-1054, DOI: 10.1016/j.ijrmms.2009.02.004.

Tian, H., M. Ziegler, and T. Kempka (2014), Physical and mechanical behavior of claystone exposed to temperatures up to 1000°C, *Int. J. Rock Mech. Min. Sci*. **70**, 144-153, DOI: 10.1016/j.ijrmms.2014.04.014.

Witherspoon, P.A. (2001), *Geological Challenges in Radioactive Waste Isolation*, Third Word Rev., California, USA.

Wu, J.G., and H.W. Zhou (2008), Dynamic experimental research on phase transformation of Kaoliniteiteunder high temperature within microzone, *Nonmetallic. Min*. **31**, 6, 10-13, DOI: 10.1016/j.clay.2013.07.017 (in Chinese).

Zhang, L.Y. (2012), Research on damage evolution and fracture mechanisms of mudstone under high temperature, Ph.D. Thesis, China. Univ. Min. Tech., Xuzhou (in Chinese).

Zhang, Z.Q., and R.Z. Yuan (1993), Study on dchydroxylation process of Kaolinite and its structural change, *Bull. Chin. Ceramic Soc*. **14**, 37-41 (in Chinese).

Zheng, J.D., B.B. Chang, T.T. Chen, and J. Yin (2010), Study on the high temperature modification of attapulgite, *Appl. Chem. Industry* **39**, 1835-1837 (in Chinese).

Zhu, H.J., X. Yao, and Z.H. Zhang (2008), Optimization of calcined temperature for Kaolinite activation, *J. Build. Mater*. **11**, 621-625 (in Chinese).

Received 19 June 2015
Received in revised form 4 February 2016
Accepted 5 April 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2092-2113
DOI: 10.1515/acgeo-2016-0091

Local Seismic Events in the Area of Poland Based on Data from the PASSEQ 2006-2008 Experiment

Marcin POLKOWSKI[1], Beata PLESIEWICZ[2], Jan WISZNIOWSKI[2], Monika WILDE-PIÓRKO[1], and PASSEQ Working Group

[1]Institute of Geophysics, Faculty of Physics, University of Warsaw, Warsaw, Poland;
e-mail: marcin.polkowski@igf.fuw.edu.pl (corresponding author)

[2]Institute of Geophysics, Polish Academy of Sciences, Warsaw, Poland

Abstract

PASSEQ 2006-2008 (Passive Seismic Experiment in TESZ; Wilde-Piórko *et al.* 2008) was the biggest passive seismic experiment carried out so far in the area of Central Europe (Poland, Germany, the Czech Republic and Lithuania). 196 seismic stations (including 49 broadband seismometers) worked simultaneously for over two years. During the experiment, multiple types of data recorders and seismometers were used, making the analysis more complex and time consuming. The dataset was unified and repaired to start the detection of local seismic events. Two different approaches for detection were applied for stations located in Poland. The first one used standard STA/LTA triggers (Carl Johnson's STA/LTA algorithm) and grid search to classify and locate the events. The result was manually verified. The second approach used Real Time Recurrent Network (RTRN) detection (Wiszniowski *et al.* 2014). Both methods gave similar results, showing four previously unknown seismic events located in the Gulf of Gdańsk area, situated in the southern Baltic Sea. In this paper we discuss both detection methods with their pros and cons (accuracy, efficiency, manual work required, scalability). We also show details of all detected and previously unknown events in the discussed area.

Key words: local seismicity, seismic detection methods, Poland.

Ownership: Institute of Geophysics, Polish Academy of Sciences

1. INTRODUCTION

1.1 PASSEQ 2006-2008

The Transeuropean Suture Zone (TESZ), which extends over 2000 km from Great Britain to the Black Sea, makes a contact zone between Eastern and Western Europe. The eastern border of the TESZ is defined by the Teisseyre–Tornquist Zone (TTZ), which runs across Poland. TTZ is a well-defined 50-100 km wide zone and delimits two tectonic units – the East European Craton and the Eastern and Paleozoic Platform of Western and

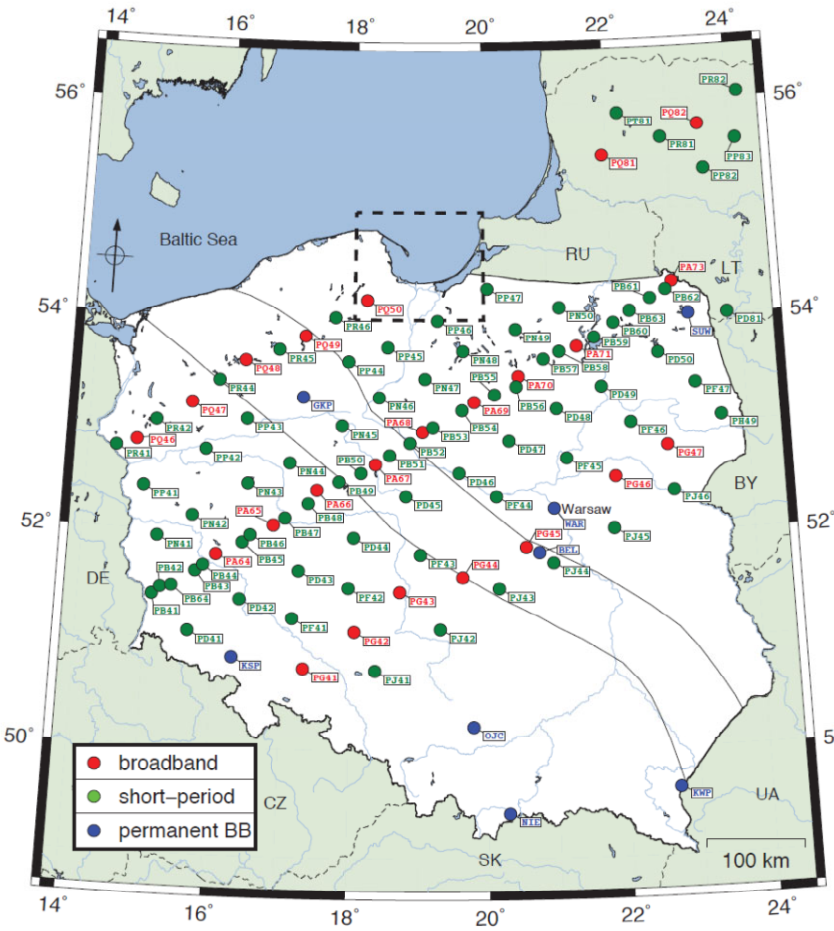


Fig. 1. Location map of PASSEQ 2006-2008 stations used in this study. PASSEQ stations are marked red (broadband) and green (short-period). Permanent stations of the Polish Seismological Network are marked blue. Two thin grey lines indicate the location of the Trans European Suture Zone. The thick dashed line marks the area shown in Fig. 9.

Central Europe (Pharaoh 1999). The East European Craton crust next to TTZ in northeastern Poland is 42-47 km thick, while on the Paleozoic Platform sill of the crust, its depth is 28-34 km (Grad *et al.* 2009). We can find a great variety of lithosphere structures within Polish territory. Recognition of the TTZ deep lithosphere structure is crucial for understanding tectonic processes that are taking place in Central Europe.

The passive seismic experiment (called the Passive Seismic Experiment in the Trans European Suture Zone (PASSEQ)) was carried out from May 2006 till June 2008 (Wilde-Piórko *et al.* 2008). The goal of the project was to recognize deep lithospheric structures in the border zone of the Precambrian Platform in Northern and Eastern Europe, and the Paleozoic Platform in Central and Western Europe. The PASSEQ experiment was concentrated in the central part of TESZ. The PASSEQ experiment was supported by 17 participating scientific institutions from Europe and the United States.

A total of 147 three-component temporary short-period and 49 temporary broadband seismic stations were installed along about 1200 km of seismic profiles, which were used to investigate the structure of the crust and upper mantle in TESZ. The experiment included a 400 km wide array starting in Germany, through the Czech Republic, Poland and ending in Lithuania. The distances between the stations were about 60 km, while in the longest central profile the distance was only about 20 km (Fig. 1) Moreover, the stations of the National Seismic Networks were also included. The stations were mainly located in forester shelters, castles, monasteries and small farms. The sampling frequency of the instruments was 20, 50 or 100 Hz. Finally, all seismic registrations were gathered on the GeoForschungsZentrum Potsdam (GFZ) servers in Potsdam (Germany), where they were exclusive data for 6 years. Since May 2014 these data are open for access and can be found on the GFZ website on request.

1.2 Detection

There are many methods of detection, which were classified into a few categories by Withers *et al.* (1998): methods working in the time domain, methods operating in the frequency domain, methods based on particle motion processing, and pattern matching methods. Detection in the time domain is based on an analysis of the changes of signal amplitude with time. Signal analysis can be performed separately for all signal components and different frequency bands. The simplest possible detection method is based on the amplitude threshold, but since it is extremely sensitive to noise, the commonly used methods involve relations between amplitude averages over moving time windows of different lengths (Allen 1978, 1982; Trnkoczy 2012).

Event detections in the frequency domain are based on Fourier Transform (McEvilly and Majer 1982), band-pass filters (Evans and Allen 1983, Gledhill 1985), Walsh Transform (Goforth and Herrin 1981, Ebel 1996), and autoregressive coefficients (Tarvainen 1991). The pattern recognition methods (PR) (Joswig 1990, 1993) benefit from information both in the time and frequency domains, however it is quite difficult to prepare a good set of patterns.

Artificial neural networks (ANNs) were widely used for seismic event detection. ANNs are taught to detect events; therefore, they require the set of examples of events and noise. Most of ANN methods are based on feed-forward multi-layer-perceptron (MLP) networks. The networks are fed by vectors of momentary parameters of the signal or by the moving window of one parameter. The moving window method (Wang and Teng 1995, 1997; Gentili and Michelini 2006) corresponds to detection in the time domain, whereas a network with a vector of temporal parameters on input corresponds to detection in the frequency domain (Wang and Teng 1995, Tiira 1999). Gentili and Michelini (2006), as well as Madureira and Ruano (2009) fed the ANN by a moving window of vectors. Another solution is to use a network with recurrent neurons (RN) (Williams and Zipser 1989, Elaman 1990) investigating momentary values in the frequency domain as well as any relationships in the time domain (Wiszniowski *et al.* 2014, Tiira 1999). Two methods, namely the detection in time domain and RN, were used for the analysis of data from the experiment PASSEQ.

2. METHODOLOGY – REAL TIME RECURRENT NEURAL NETWORK

2.1 The structure of the Real Time Recurrent Neural Network applied for seismic event detection

The Real Time Recurrent Neural Network (RTRN) belongs to artificial neural networks. The structure of RTRN (Fig. 2) and the learning method were developed by Wiliams and Zipser (1989). The RTRN was used by Wiszniowski *et al.* (2014) for the detection of small natural earthquakes with a magnitude of 0.4 to 2.5, in the Polish part of the Western Carpathians. This is the simplest recurrent artificial neural network that operates in discrete time settings. The *i*-th artificial neuron at a moment *t* has an output value:

$$V_i(t) = \mathrm{g}\left(\sum_{j=0}^{n-1} w_{ij} v_j(t)\right)$$

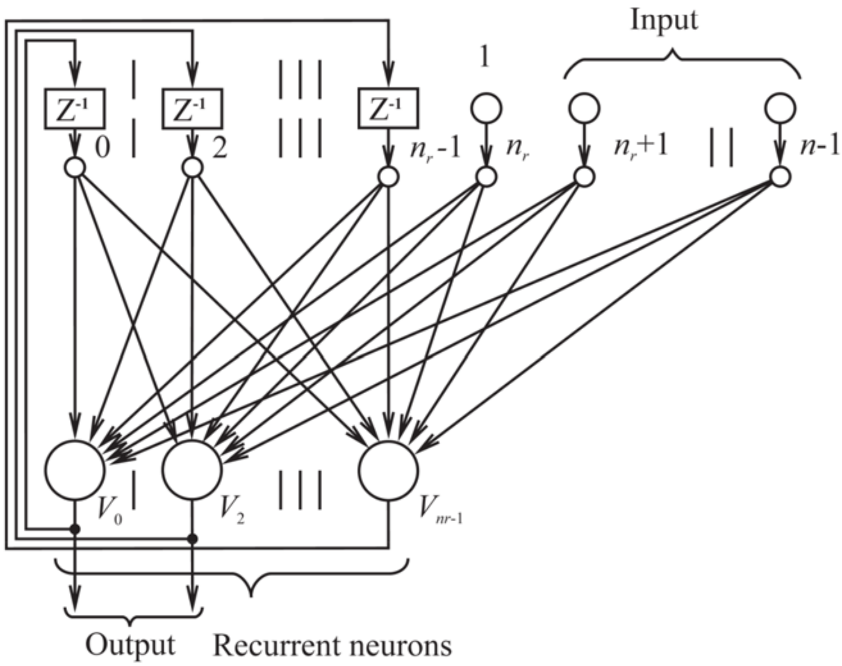


Fig. 2. The model of the RTRN network applied to the detection of small natural earthquakes in Poland. It contains nr recurrent neurons. Outputs of neurons V_0, V_1, and V_2 are outputs of the network. They corresponded to: detection of event (V_0), detection of phase $P(V_1)$ onset, and detection of phase $S(V_2)$ onset. Inputs from n_r+1 to n–1 are inputs of the network, whereas input nr has a constant value 1 (bias). [Z 1] is a delay element with a delay of one step of the network work. RTRN used for detection in PASSEQ has 26 inputs and 16 recurrent neurons. Therefore, each neuron has $n = 43$ inputs.

where g(.) is an activation function, w_{ij} are weight coefficients between inputs and the output of the neuron, $v_j(t)$ are input values:

$$v_j(t) = \begin{cases} V_j(t-1) & j = 0,\ldots,n_r - 1 & \text{recurrent neurons} \\ 1 & j = n_r & \text{constant value egual 1} \\ x_i(t) & j = n_r + i;\, i = 1,\ldots,k & \text{inputs of the RTRN} \end{cases}$$

$V_j(t)$ are output values of neurons, $x_i(t)$ are input values of the RTRN, k is the number of input neurons, and $n = n_r + k + 1$ is the number of inputs of neurons.

RTRN consists of recurrent neurons whose outputs are recurrently connected to their own inputs (Fig. 2). The return signals from recurrent neurons are put to inputs with a delay of one network time step. The purpose of RTRN is to discriminate the seismic events from noise (disturbances), aiming to detect a maximum number of small earthquakes, whose amplitudes are below the noise level, and to produce an acceptable number of false detections. Outputs of the first three recurrent neurons were also outputs of the network and they corresponded to: detection of event (V_0), detection of phase $P(V_1)$ onset and detection of phase $S(V_2)$ onset. Only the output V_0 was taken into consideration for the events detection. Outputs V_1 and V_2, which correspond to P and S waves, are used only during the ANN training.

The data for training the RTRN consists of a section referring to seismic events and another section concerning seismic noise, mainly due to human activity. The training method relies on corrections made to weight coefficients to minimize the discrepancy between real and expected neural network output value according to the formula:

$$\Delta w_{pq} = -\alpha \frac{\partial E(t)}{\partial w_{pq}}$$

where α is a learning rate and the function E is defined as:

$$E(t) = \sum_{i=0}^{2} \eta_i(t) \left[\zeta_i(t) - o_i(t) \right]^2$$

where ζ_i is the expected output of i-th neuron, η_i is a learning coefficient and o_i is the real output of the RTRN, $i = 0,1,2$. Both ζ and η depend on the phases of the seismic event. The values of the expected output of network were:

ζ_0 (event detection): –1 for noise, +1 for a seismic event from the first phase to a few seconds after the signal reached the level of noise,

ζ_1 (phase P): –1 before the onset of the phase, +1 from the beginning of the phase to a few seconds after the end of the P coda, but not less than to the beginning of the phase S,

ζ_2 (phase S): –1 before the onset of the phase, +1 from the beginning of the phase to a few seconds after the end of the S coda.

The values of learning coefficients vary according to the formulas (Wiszniowski *et al.* 2014):

$\eta_0(t) = 0$; $t \in (T_0, T_0+300)$,
$\eta_0(t) = 1$; $t \in (T_0+300, T_P)$,
$\eta_0(t) = 0.001$; $t \in (T_P, T_S)$ or $t \in (T_P, T_P+5)$,
$\eta_0(t) = 1$; $t \in (T_S, T_S+10)$,

$$
\begin{array}{ll}
\eta_0(t) = e^{TS-t+10} & ; t > T_S + 10 , \\
\eta_1(t) = 0 & ; t \in (T_0, T_0+300) , \\
\eta_1(t) = 1 & ; t \in (T_0+300,\ T_S+1) , \\
\eta_1(t) = e^{TS-t+1} & ; t > T_S + 1 , \\
\eta_2(t) = 0 & ; t \in (T_0, T_0+300) , \\
\eta_2(t) = 1 & ; t \in (T_0+300,\ T_S+10) , \\
\eta_2(t) = e^{TS-t+10} & ; t > T_S + 10 ,
\end{array}
$$

where T_P and T_S are times of phases. The T_0 is the time of the beginning of each training record. The value of the learning coefficients is zero after the T_0, because output values of neurons are initially set to zero and neurons need a few time steps to reach their typical operating values.

Preparation of the training data for RTRN was difficult, because of a lack of recordings of local events in the study area. To do this, the seismic records from the Monitoring of Seismic Hazard of the Territory of Poland (MSHTP) project were investigated (Trojanowski *et al.* 2015). Records from 5 days (15-19 January 2009) were analyzed, 170 earthquakes (local – induced in Legnica-Glogow Copper District (LGCD) and the Upper Silesia Coal Basin (USCB), natural from the Podhale region, regional and teleseismic) and about 9000 noises were identified. On this basis, two training sets were prepared − one with seismic events and the other with noises. Parallel computation with the use of OpenMP library was applied to speed up the calculations. 10-fold parallelism on a 16-core computer was obtained. Two methods were used for learning the neural network: real time recurrent learning (Wiliams and Zipser 1989) and back propagation throough time (Werbos 1990).

The RTRN, which was used in MSHTP project, relied on wide range of frequencies of a recorded signal to resolve noise, characterized with, for example, higher frequencies. All stations recorded the seismic signal in a frequency band of 1-40 Hz. The sampling frequency was 100 Hz. Besides, the noise level was lower than in the PASSEQ experiment. Therefore, the RTRN analyzing of only the vertical signal in the MSHTP project was good enough for detection (hereinafter referred to as RTRN1D). In case of the PASSEQ project, the stations were recorded in different a frequency band, mainly depending on the sampling frequency. The lowest sampling frequency used was 20 Hz. This means that maximum registered frequency in all PASSEQ seismic stations was 8 Hz. It resulted in a substantial limitation of the signal information provided to the detection algorithm. Especially, high frequency noises are worse discriminated from seismic waves. To compensate for the limitation of the detection of low frequency data, the vertical signal components were used. This allows for the use of information, such as the polarization of the signal. This solution was prepared for the Podhale data in MSHTP, but it was not applied because of the sufficient effectiveness of the RTRN1D. The new RTRN with three dimensional inputs data (RTRN3D) was extended to process horizontal data. Now, it consists of 16 recurrent neurons and 26 input neurons. The first 11 inputs are supplied with the ratio of short-term averages (STAs) and long term averages (LTAs) of the filter bank of horizontal components of the signal. The filter bank filters the signal in different frequency ranges at the width of 1/3 octave. It consists of 11 filters with a middle frequency in the range of 0.6 to 8 Hz (Fig. 3a). The RTRN3D has less vertical inputs than the RTRN1D, which had 15 inputs, because of a shorter frequency band. The next 11 inputs of the RTRN3D are supplied with the ratio of STAs and LTAs of a root mean square of vertical components of the signal filtered by the same filter bank.

The result of training the RTRN1D shows the quick forgetfulness of a recurrent network. Especially, in the case of regional events, the information of the *P* wave is forgotten when the *S* wave enters. Therefore, to gain more information on time dependencies of the signal, the last 4 inputs are supplied with the signals shifted 18 seconds in time – predating the main signal. The signal is also filtered, but by a smaller number of filters in a wider range

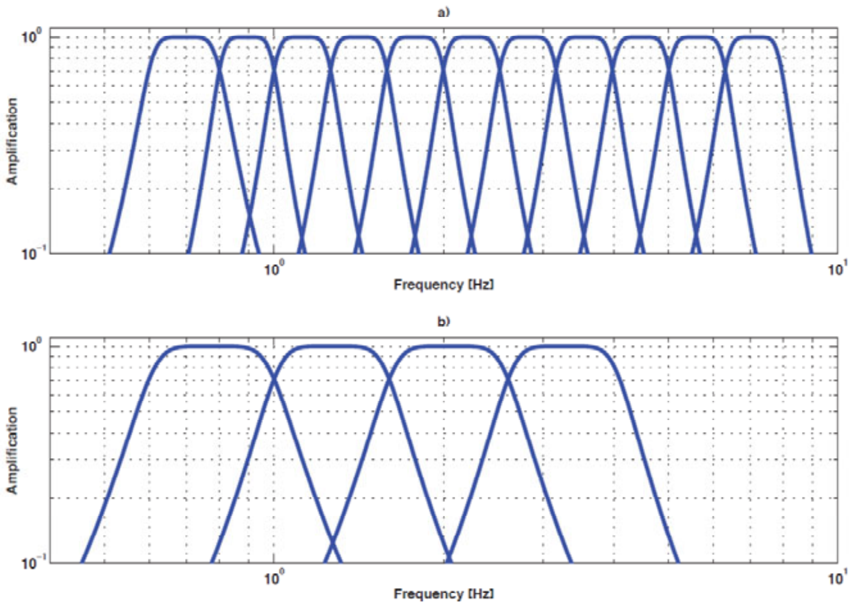


Fig. 3. Amplitude-frequency characteristics of input signals from horizontal and vertical components for RTRN3D detection: (a) filter bank for horizontal and vertical components, (b) filter bank for vertical components supplied to the input of RTRN3D in 18 s advance.

(Fig. 3b). Thanks to shifting the data in time, the future information is available earlier, before the current information is forgotten.

2.2 RTRN3D detection method for local seismic events in Poland based on data from THE PASSEQ project

Data recorded during the PASSEQ 2006-2008 project by seismic stations placed in a radius of 100 km around the town of Jarocin (17 seismic stations, Table 1), was analyzed with the RTRN3D network. The region of Jarocin was selected due to the seismic activity discovered in this region in 2012 (Lizurek *et al.* 2013). The output V_0 of the RTRN3D was taken into consideration for the events detection (Figs. 4 and 5). The detection threshold was set at the level of 0.3. After voting on a least 3 stations, we obtained 3854

Table 1

List of seismic stations of the PASSEQ project from the Jarocin region on which the RTRN3D detection was applied

No.	Station	Place	Latitude [°N]	Longitude [°E]	Recording period
1	PA66	Graby, Poland	52.4092	17.4722	26.07.2006-18.11.2007
2	PB45	Koczury, Poland	51.9093	16.3877	26.07.2006-16.05.2007
3	PB46	Pośmigiel, Poland	51.9817	16.4970	11.09.2007-28.01.2008
4	PB46B	Branikowo, Poland	51.9741	16.4753	26.07.2006-16.05.2007
5	PB47	Niesłabin, Poland	52.1442	17.0076	16.05.2007-28.01.2008
6	PB48	Marianowo, Poland	52.2849	17.3464	27.07.2006-26.06.2008
7	PB49	Ćwierdzin, Poland	52.4918	17.8048	16.05.2007-28.01.2008
8	PB50	Ciencisko, Poland	52.5738	18.1279	27.07.2006-28.01.2008
9	PD43	Zygmuntowo, Poland	51.6553	17.2284	28.07.2006-26.06.2008
10	PD44	Piskory, Poland	51.9710	18.0363	25.07.2006-24.06.2008
11	PD45	Słubin, Poland	52.3633	18.8092	08.07.2006-24.06.2008
12	PF42	Mikstat, Poland	51.5022	17.9694	08.07.2006-24.05.2008
13	PF43	Mianów, Poland	51.8198	19.0269	09.07.2006-27.05.2008
14	PG43	Świerki, Poland	51.4720	18.7240	08.07.2006-27.05.2008
15	PN42	Buków, Poland	52.1417	15.6283	05.07.2006-27.03.2007 15.05.2007-23.06.2008
16	PN43	Duszniki, Poland	52.4581	16.4361	05.05.2007-23.06.2008
17	PN44	Łoskoń Stary, Poland	52.6603	17.0552	15.07.2006-23.06.2008

Table 2

Number of RTRN3D detections on the PASSEQ seismic stations from the regions near Jarocin and the Gulf of Gdańsk

Region	Jarocin	Gulf of Gdańsk
Time period	2006.07.05 - 2008.06.26	2006.06.14 - 2008.06.18
Number of detection	3854	6198
Induced events from LGCB and USCD	59%	20%
Other local and regional events	3%	2%
Teleseismic events	7%	4%
Noises and disturbances	31%	74%

detections (Table 2). They included mostly local induced events from the LGCD and USCB, other local and regional events, teleseismic events, and false detections. The other local and regional events include both natural and induced events from the surrounding Polish area, as well as singular induced events from other regions in Poland, *e.g.*, Bełchatów. Most tremors come from the LGCD (42% of detections). Only one earthquake from 6 May 2007, 7:32 UTC of $M_L = 2.8$ was localized in the region of Jarocin.

Table 3

List of seismic stations of the PASSEQ project from the Gulf of Gdańsk region on which seismic records the RTRN3D detection was applied

No.	Station	Place	Latitude [°N]	Longitude [°E]	Recording period
1	PA70	Wikno, Poland	53.4720	20.5229	27.07.2006-19.11.2007
2	PA71	Kołowinek, Poland	53.7424	21.4321	27.07.2006-28.07.2007
3	PA73	Pobondzie, Poland	54.3191	22.9493	28.07.2006-24.10.2007
4	PB57	Sasek, Poland	53.6294	20.9086	04.10.2006-04.06.2008

to be continued

Table 3 (continuation)

No.	Station	Place	Latitude [°N]	Longitude [°E]	Recording period
5	PB59	Grabówka, Poland	53.8226	21.7022	29.06.2006-24.12.2007
6	PB60	Siemionki, Poland	53.9503	22.0117	29.06.2006-06.11.2007
7	PB61	Garbaś, Poland	54.1635	22.5941	29.06.2006-16.01.2008
8	PD81	Paliepis, Lithuania	54.0056	23.7865	17.06.2006-04.09.2007
9	PN47	Świecie nad Osą, Poland	53.4519	19.0899	18.07.2006-03.06.2008
10	PN48	Sąpy, Poland	53.7092	19.6750	18.07.2006-04.06.2008
11	PN49	Sętal, Poland	53.9103	20.4865	19.07.2006-04.06.2008
12	PP45	Cisowa Góra, Poland	53.7445	18.5186	18.07.2006-18.06.2008
13	PP46	Chartowo, Poland	53.9930	19.2812	18.07.2006-04.06.2008
14	PP47	Kiersiny, Poland	54.2879	20.0599	19.07.2006-12.05.2008
15	PP82	Paliepai, Lithuania	55.3525	23.5543	15.06.2006-10.01.2008
16	PP83	Ustronė, Lithuania	55.6155	24.1015	16.06.2006-09.01.2008
17	PR46	Trzebun, Poland	54.0194	17.7060	16.07.2007-18.06.2008
18	PR81	Lašinskiai, Lithuania	55.6588	22.8915	14.06.2006-09.01.2008
19	PR82	Narteikai, Lithuania	56.0482	24.1764	15.06.2006-09.01.2008
20	PQ50	Jasiowa Huta, Poland	54.1781	18.1912	04.08.2006-08.10.2006
21	PQ81	Radvietis, Lithuania	55.5111	21.9334	02.11.2006-22.09.2007
22	PQ82	Jaugėlai, Lithuania	55.7658	23.4980	03.11.2006-22.09.2007
23	PT81	Žarėnai, Lithuania	55.8361	22.2079	14.06.2006-02.10.2007

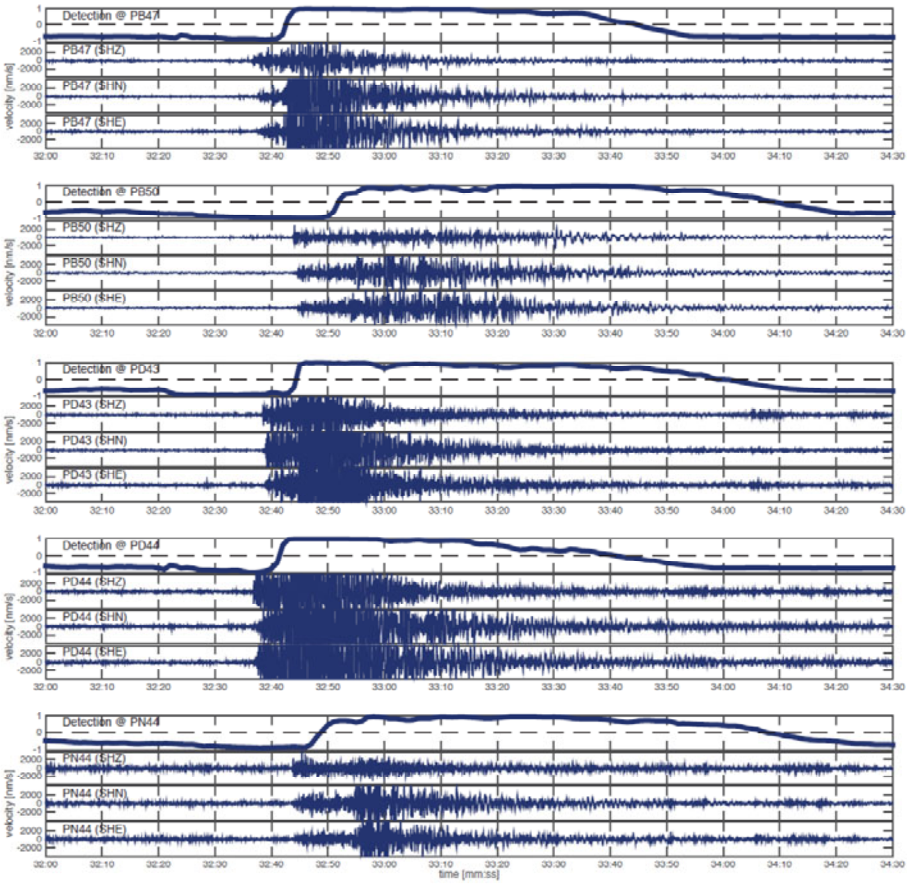


Fig. 4. Example of records and RTRN3D detection − the earthquake of 6 May 2007, 7:32 UTC $M_L = 2.8$ occurred near Jarocin: PB47, PB50, PD43, PD44, and PN44 seismic stations; the signals on top of the records are V_0 (event detection) outputs of RTRN3D.

Figures 4 and 5 show examples of seismograms and RTRN3D network detections of events recorded by PASSEQ seismic stations.

The second region, which was analyzed with the RTRN3D network, was the Gulf of Gdańsk and its neighborhood (23 seismic stations, Table 3). Records from the PASSEQ stations situated within a radius of about 250 km from the shore, including stations located in Poland and Lithuania, were analyzed. The result consisted in 6198 detections (Table 2). They included mostly false detections, generated by sea noise, as well as teleseismic events and regional seismic events from Poland.

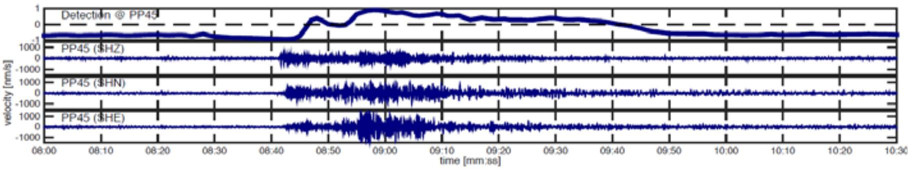


Fig. 5. Output of detection of the RTRN3D (detpp45) together with seismic signal of PASSEQ station PP45 for 2 May 2007 event in the Gulf of Gdańsk area. It is seen that both *P* and *S* waves are used for event detection.

3. METHODOLOGY – STA/LTA DETECTION METHOD

3.1 Detection of seismic signals

A detection goal was to analyze the continuous time series (seismic record) and provide a list of moments where the possible signal (seismic wave) was recorded. The analysis was done after the PASSEQ experiment had ended, so the methods were customized for analyzing large portions of data. The main python function was prepared to return a list of signal detections for a given file name (already prepared one-hour, one station, one component miniseed file). Most of the continuous data analysis and detection were done using ObsPy software (Beyreuther *et al.* 2010, Megies *et al.* 2011).

The analysis was carried out in three steps: data loading, filtration and detection of the signal. The data was filtered with a zero-phase bandpass Butterworth filter (from 4.0 to 9.5 Hz). After filtration, the mean value of the trace was removed and the amplitude normalized to a constant value (the same value for all stations). An example of the filtration of the broadband trace is shown in Fig. 6. A detector has to adapt to current conditions (*e.g.*, noise level). The simplest way to create such a detector is calculating two moving averages over a signal with different windows: short and long. Most of the seismic detectors use a relation between the short-term average (STA) and long term average (LTA) (Withers *et al.* 1998, Trnkoczy 2012). In this paper, Carl Johnson's STA/LTA detection algorithm was used (Johnson 1979). This algorithm calculates four moving averages and takes two parameters:

$$eta = star - (ratio * ltar) - abs(sta - lta) - quiet$$

where: *eta* is the detector response – a value over 0 means detection, *sta* is the short term moving average of signal, *lta* is the long term moving average of signal, star is the short term moving average of absolute value of signal and *lta* difference, *ltar* is the long term moving average of star, ratio, and quiet-sensitivity parameters. Short term average had a 4-second window and

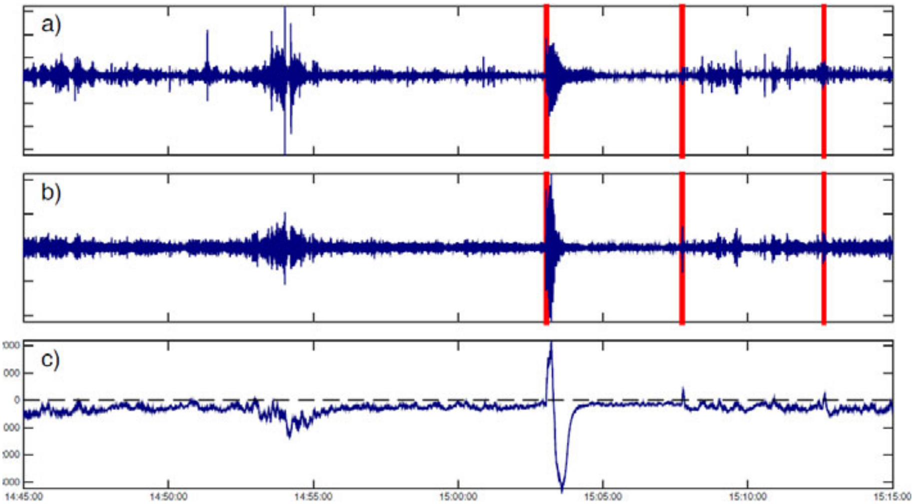


Fig. 6: (a) Example of 30-minute recording (vertical component) from station PA64 recorded on 24 August 2006. Thick red lines indicate the trigger result. (b) Same example filtered with bandpass filter from 4.5 to 9 Hz. (c) Trigger response: values over 0 for longer than 0.1 seconds are taken as a positive trigger and marked thick red lines.

long term average had a 32-second window. Both ratio and quiet parameters had value 2. Example result is shown in Fig. 6.

3.2 Analysis of coinciding detections – grid search method

The detector responds to seismic waves together with noise signals. An additional analysis of detection signals is required to recognize local seismic events. To analyze coinciding detections, the study area was represented as a grid with a cell size of 0.05 by 0.05 arc degrees. The analysis was done one day of recording at the time, which was optimal for the performance and number of I/O operations. An analysis of longer periods would have required more operating memory, exceeding the capacity of a standard workstation computer. A list of detections from all the stations for that day was taken with their exact time – at this point detection on the list included real seismic signals from local events, a detection of regional and teleseismic events and accidental detections of noise. In the next step, a list of detections was analyzed to check if there was a place in space and time where a seismic event might have occurred causing such a distribution of detections. This allows for reducing the number of detections not correlated with local seismicity. For each grid cell, the number of stations operating within a radius of

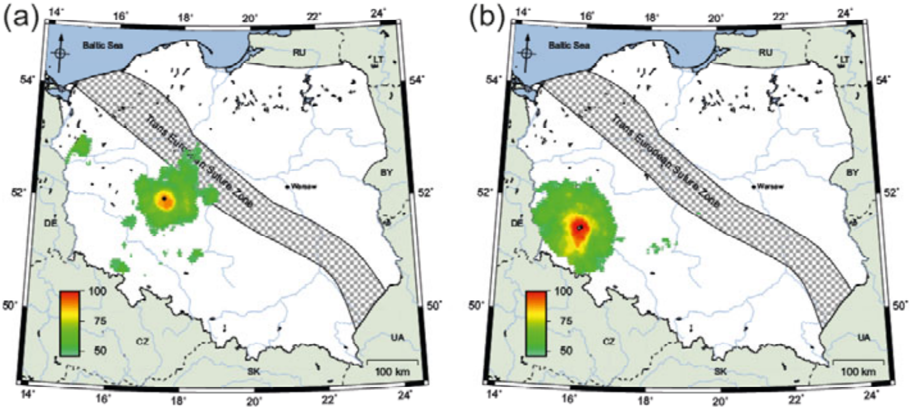


Fig. 7: (a) New seismic event on 6 May 2007 07:32:30 UTC. Color shows percent of stations in 150 km radius with a positive trigger. A black dot shows the estimated earthquake location. See text for more explanation. (b) Example for a known local event. The color shows the percent of stations in a 150 km radius with a positive trigger. A black dot shows the event epicenter from a Local Bulletin of the Institute of Geophysics, Polish Academy of Sciences, and a white dot shows the calculated epicenter.

150 km was determined, and the cells with at least 15 stations working in their neighborhood were selected for the subsequent analysis. For each of the selected cells, all the detections of the surrounding stations (within the 150 km radius) are inspected; for each station, the *P*-wave travel time from a seismic source located in the center of the cell is computed using the 1-D reference iasp'91 model (Kennett and Engdahl 1991, Crotwell and Owens 1999); then, for each particular station detection, the origin time of a corresponding hypothetical event located in the center of the cell is estimated. The hypothetical events identified by more than 50% of the surrounding stations are kept as candidates for real events (differences up to 3 seconds between the origin times provided by individual stations are accepted). This analysis allows us to produce − for each second of studied time period − maps illustrating the probability of a seismic event location; to each cell of the study area, we assign the percentage of (surrounding) stations that identified, at the particular time moment, a seismic event in the center of the cell. Examples of grid search results are shown in Fig. 7 – for the previously known and new local events. Event localization was calculated as a weighted average of locations of grid cells with a probability of over 50%. The value of probability was used as a weight in the averaging process. Figure 8 shows a histogram of the event location accuracy: the difference between the bulletin and grid search location for all 53 events in the LGCB that occurred from 1 August

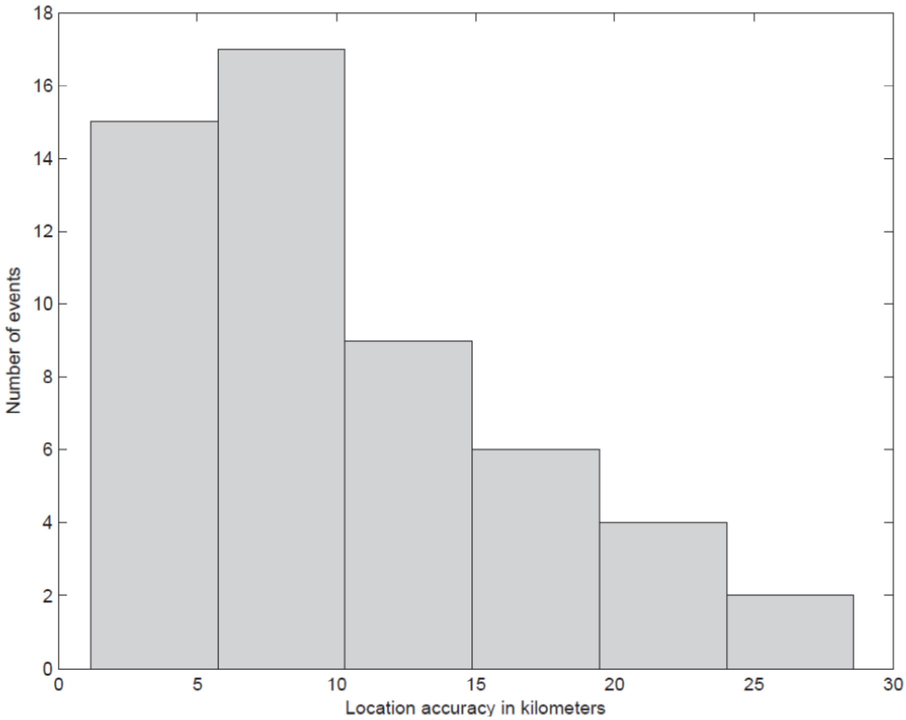


Fig. 8. Histogram of event location accuracy (difference between bulletin and grid search location). Total of 53 events in the LGCB were analyzed from 1 August 2006 to 31 October 2006.

2006 to 31 October 2006. On average, the location accuracy is better than 15 km, however this analysis is not entirely correct due to the unknown event location quality in the bulletin, which is based on fewer seismic stations than available during PASSEQ. The detection accuracy could be improved by use of a more detailed seismic wave velocity model for the studied area (Grad and Polkowski 2012, Polkowski and Grad 2015). The advantage of the grid search method is its ability to detect multiple seismic events occurring at the same time in different parts of the study area.

3.3 Application of STA/LTA for PASSEQ data

After analyzing the whole PASSEQ 2006-2008 data set for Poland, over 3500 events were automatically detected. This list was then filtered to eliminate noise-only detections. Filtration was done by eliminating the events visible in less than 20 grid cells.

The new list consisted of 1206 events. Events in the areas of the LGCD and the USCB were not analyzed due to the known high level of the seismicity in these zones. These events were used for determining the method accuracy (detection and location). Only 46 events from other areas of Poland were on the list and have been manually analyzed: 30 events were accidental detections from known regional and teleseismic events, 12 were accidental coincidences of noise and 4 were newly localized seismic events: one in central Poland, and three in the Gulf of Gdansk.

4. RESULTS

Both the 3D RTRN and STA/LTA methods provided similar results. 3D RTRN detected two additional events in the Gulf of Gdańsk. Detection of these events was difficult, because of the station distribution only to the south of the Gulf of Gdańsk. Table 4 presents statistics of RTRN3D detections on PASSEQ data. LocSAT application with the IASP91 travel-time tables provides estimates of the origin time, epicentral location, and the depth from an iterative least-squares inversion of travel time, slowness, and/or azimuth (Bratt and Bache 1988, Bratt and Nagy 1991). The result of analysis of detections of STA/LTA in the area of Poland, except in the areas of known seismicity (the LGCD and the USCB), is 4 local seismic events, which are presented in Table 4. STA/LTA with grid search provided information only on the origin time and location of the event epicenter (and not

Table 4

Local earthquakes in Poland, detected using the records of the PASSEQ project

No.	M_L	STA/LTA method				RTRN3D method			
		Time [UTC]	Latitude [°N]	Longitude [°E]	Depth [km]	Time [UTC]	Latitude [°N]	Longitude [°E]	Depth [km]
A	2.3	-	-	-	-	2007-06-14 00:10:10	54.48	18.99	10
B	3.0	2007-03-10 23:08:51	54.60	18.75	-	2007-03-20 23:08:51	54.62	18.73	4
C	3.2	2007-05-02 07:08:23	54.69	19.17	-	2007-05-02 07:08:23	54.72	19.14	10
D	-	2006-09-12 15:12:14	54.55	19.32	-	2006-09-12 15:12:14	-	-	-
E	2.6	-	-	-	-	2007-05-22 16:29:30	54.85	19.82	12
F	2.8	2007-05-06 07:32:30	52.02	17.48	-	2007-05-06 07:32:31	52.01	17.50	4

on the event depth and magnitude). The localization of all events from the Gulf of Gdańsk is shown on the map in Fig. 9. The localization of the event near Jarocin is shown on the map in Fig. 7a.

The event near Jarocin in central Poland (6 May 2007 07:32:30 UTC) was a natural, tectonic seismic event. In the same area, a bigger seismic event (*M*3.8) occurred on 6 January 2012 at 15:37:56 UTC (Lizurek *et al.* 2012).

The origin of events in the Gulf of Gdańsk is disputable. While in this area natural earthquakes have occurred in the past (Meyer and Kulhanek 1980, Wiejacz *et al.* 2001, Wiejacz 2006), the Gulf of Gdańsk is also used by the Polish Navy for destroying old explosives from the Second World War. The navy operation is confidential and we were unable to fully confirm if the detected events were or were not a result of military activity. The data quality and distribution of stations does not allow us to confirm the origin of those events. Additionally, in March and April 2015 the Polish Navy has destroyed newly discovered mines in the Gulf of Gdańsk. Seismic waves from these explosions were recorded on the stations of the "13 BB star" experiment located in northern Poland (Grad *et al.* 2015).

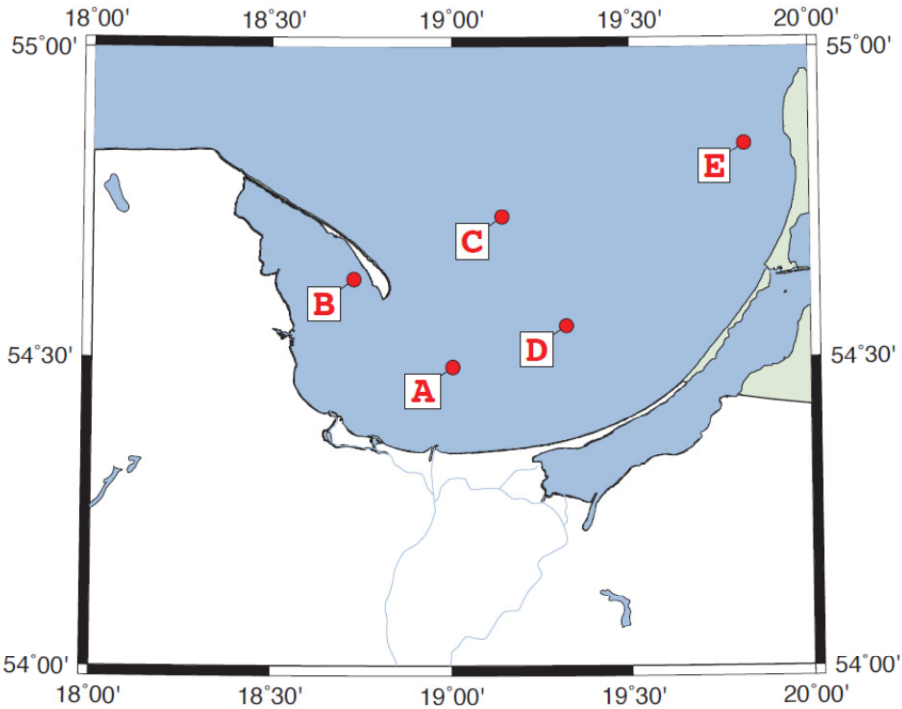


Fig. 9. Location of events in the Gulf of Gdańsk region detected with LTA/STA and RTRN3D methods.

A c k n o w l e d g m e n t s. The public domain GMT software (Wessel *et al.* 2013) has been used to produce maps. The Polish Seismological Network provided data from permanent broadband stations. We also would like to thank Professor Timo Tiira for his advice.

We also would like to thank the members of PASSEQ Working Group: Monika Wilde-Piórko, Wolfram H. Geissler, Jaroslava Plomerová, Marek Grad, Vladislav Babuška, Ewald Brückl, Jolanta Cyziene, Wojciech Czuba, Richard England, Edward Gaczyński, Renata Gazdova, Soren Gregersen, Aleksander Guterch, Winfried Hanka, Endre Hegedűs, Barbara Heuer, Petr Jedlička, Jurga Lazauskiene, G. Randy Keller, Rainer Kind, Klaus Klinge, Petr Kolinsky, Kari Komminaho, Elena Kozlovskaya, Frank Krüger, Tine Larsen, Mariusz Majdański, Jiří Malek, Gediminas Motuza, Oldřich Novotný, Robert Pietrasiak, Thomas Plenefisch, Bohuslav Růžek, Saulius Sliaupa, Piotr Środa, Marzena Świeczak, Timo Tiira, Peter Voss, and Paweł Wiejacz for they support and cooperation.

This work was partially supported within statutory activities No. 3841/E-41/S/2015 of the Ministry of Science and Higher Education of Poland. The National Science Centre Poland provided financial support for this work via an NCN grant DEC-2011/02/A/ST10/00284.

References

Allen, R. (1978), Automatic earthquake recognition and timing from single traces, *Bull. Seismol. Soc. Am.* **68**, 5, 1521-1532.

Allen, R. (1982), Automatic phase pickers: their present use and future prospects, *Bull. Seismol. Soc. Am.* **72**, 6B, S225-S242.

Beyreuther, M., R. Barsch, L. Krischer, T. Megies, Y. Behr, and J. Wassermann (2010), ObsPy: A Python toolbox for seismology, *SRL* **81**, 530-533, DOI: 10.1785/gssrl.81.3.530.

Bratt, S.R., and T.C. Bache (1988), Locating events with a sparse network of regional arrays, *Bull. Seismol. Soc. Am.* **78**, 2, 780-798.

Bratt, S.R., and W. Nagy (1991), *The LocSAT Program*, Science Applications International Corporation.

Crotwell, H.P., T.J. Owens, and J. Ritsema (1999), The taup toolkit: Flexible seismic travel-time and ray-path utilities, *Seismol. Res. Lett.* **70**, 154-160, DOI: 10.1785/gssrl.70.2.154.

Ebel, J. (1996), Development of a seismic event detection and identification algorithm based on wavelet transforms, *Seismol. Res. Lett.* **67**, 1-37.

Elman, J.L. (1990), Finding structure in time, *Cognitive Sci.* **14**, 2, 179-211, DOI: 10.1016/0364-0213(90)90002-E.

Evans, J., and S. Allen (1983), A teleseismic-specific detection algorithm for single short period traces, *Bull. Seismol. Soc. Am.* **73**, 4, 1173-1186.

Gentili, S., and A. Michelini (2006), Automatic picking of P and S phases using a neural tree, *J. Seismol.* **10**, 1, 39-63, DOI: 10.1007/s10950-006-2296-6.

Gledhill, K.R. (1985), An earthquake detector employing frequency domain techniques, *Bull. Seismol. Soc. Am.* **75**, 6, 1827-1835.

Goforth, T., and E. Herrin (1981), An automatic seismic signal detection algorithm based on the Walsh transform, *Bull. Seismol. Soc. Am.* **71**, 4, 1351-1360.

Grad, M., and M. Polkowski (2012), Seismic wave velocities in the sedimentary cover of Poland: Borehole data compilation, *Acta Geophys.* **60**, 4, 958-1006, DOI: 10.2478/s11600-012-0022-z.

Grad, M., T. Tiira, and ESC Working Group (2009), The Moho depth map of the European Plate, *Geophys. J. Int.* **176**, 1, 279-292, DOI: 10.1111/j.1365-246X.2008.03919.x.

Grad, M., M. Polkowski, M. Wilde-Piórko, J. Suchcicki, and T. Arant (2015), Passive Seismic Experiment "13 BB Star" in the Margin of the East European Craton, Northern Poland, *Acta Geophys.* **63**, 2, 352-373, DOI: 10.1515/acgeo-2015-0006.

Johnson, C.E. (1979), CEDAR: An approach to the computer automation of short-period local seismic networks, Ph.D. Thesis, California Institute of Technology, Pasadena, USA.

Joswig, M. (1990), Pattern recognition for earthquake detection, *Bull. Seismol. Soc. Am.* **80**, 1, 170-186.

Joswig, M. (1993), Single-trace detection and array-wide coincidence association of local earthquakes and explosions, *Comput. Geosci.* **19**, 2, 207-221, DOI: 10.1016/0098-3004(93)90119-P.

Kennett, B.L.N., and E.R. Engdahl (1991), Traveltimes for global earthquake location and phase identification, *Geophys. J. Int.* **105**, 2, 429-465, DOI: 10.1111/j.1365-246X.1991.tb06724.x.

Lizurek, G., B. Plesiewicz, P. Wiejacz, J. Wiszniowski, and J. Trojanowski (2013), Seismic event near Jarocin (Poland), *Acta Geophys.* **61**, 1, 26-36, DOI: 10.2478/s11600-012-0052-6.

Madureira, G., and A.E. Ruano (2009), A neural network seismic detector, *Acta Tech. Jaurinesis* **2**, 2, 159-170, DOI: 10.3182/20090921-3-TR-3005.00054.

McEvilly, T.V., and E.L. Majer (1982), ASP: An automated seismic processor for microearthquake networks, *Bull. Seismol. Soc. Am.* **72**, 1, 303-325.

Megies, T., M. Beyreuther, R. Barsch, L. Krischer, and J. Wassermann (2011), Obspy – what can it do for data centers and observatories?, *Ann. Geophys.* **54**, 1, 47-58, DOI: 10.4401/ag-4838.

Meyer, K., and O. Kulhanek (1980), The gulf of Gdańsk seismic event sequence of June 25 – July 3, *Acta Geophys. Pol.* **29**, 4, 315-320.

Pharaoh, T.C. (1999), Palaeozoic terranes and their lithospheric boundaries within the trans-european suture zone (tesz): a review, *Tectonophysics* **314**, 1-3, 17-41, DOI: 10.1016/S0040-1951(99)00235-8.

Polkowski, M., and M. Grad (2015), Seismic wave velocities in deep sediments in Poland: Borehole and refraction data compilation, *Acta Geophys*. **63**, 698-714, DOI: 10.1515/acgeo-2015-0019.

Tarvainen, M. (1991), Detection local and regional seismic events using the data-adaptive method at the VAF seismograph station in Finland, *Bull. Seismol. Soc. Am.* **81**, 4, 1373-1379.

Tiira, T. (1999), Detecting teleseismic events using artificial neural networks, *Comput. Geosci.* **25**, 8, 929-938, DOI: 10.1016/S0098-3004(99)00056-4.

Trnkoczy, A. (2012), Understanding and parameter setting of STA/LTA trigger algorithm, New Manual of Seismological Observatory Practice 2.

Trojanowski, J., B. Plesiewicz, and J. Wiszniowski (2015), Seismic monitoring of Poland – temporary Seismic Project with mobile Seismic Network, *Acta Geophys.* **63**, 1, 17-44, DOI: 10.2478/s11600-014-0255-0.

Wang, J., and T.-L. Teng (1995), Artificial neural network-based seismic detector, *Bull. Seismol. Soc. Am.* **85**, 1, 308-319.

Wang, J., and T.-L. Teng (1997), Identification and picking of S phase using an artificial neural network, *Bull. Seismol. Soc. Am.* **87**, 5, 1140-1149.

Werbos, P. (1990), Backpropagation through time: what it does and how to do it, *Proc. IEEE* **78**, 10, 1550-1560, DOI: 10.1109/5.58337.

Wessel, P., W.H. F. Smith, R. Scharroo, J.F. Luis, and F. Wobbe (2013), Generic Mapping Tools: Improved version released, *EOS Trans. AGU* **94**, 45, 409-410.

Wiejacz, P. (2006), The Kaliningrad earthquakes of September 21, 2004, *Acta Geodyn. Geomat.* **3**, 7-16.

Wiejacz, P., and W. Dębski (2001), New observations of gulf of Gdansk seismic events, *Phys. Earth Planet. Int.* **123**, 2-4, 233-245, DOI: 10.1016/S0031-9201(00)00212-0.

Wilde-Piórko, M., W.H. Geissler, J. Plomerová, M. Grad, V. Babuška, E. Brückl, J. Cyziene, W. Czuba, R. England, E. Gaczyński, R. Gazdova, S. Gregersen, A. Guterch, W. Hanka, E. Hegedűs, B. Heuer, P. Jedlička, J. Lazauskiene, G. Randy Keller, R. Kind, K. Klinge, P. Kolinsky, K. Komminaho, E. Kozlovskaya, F. Krüger, T. Larsen, M. Majdański, J. Málek, G. Motuza, O. Novotný, R. Pietrasiak, Th. Plenefisch, B. Růžek, S. Sliaupa, P. Środa, M. Świeczak, T. Tiira, P. Voss, and P. Wiejacz (2008), PASSEQ 2006-2008: Passive Seismic Experimentin Trans – European Suture Zone, *Stud. Geophys. Geod.* **52**, 439-448.

Williams, R., and D. Zipser (1989), A learning algorithm for continually running fully recurrent neural networks, *Neural Comput.* **1**, 2, 270-280, DOI: 10.1162/neco.1989.1.2.270.

Wiszniowski, J., B. Plesiewicz, and J. Trojanowski (2014), Application of Real Time Recurrent Neural Network for detection of small natural earthquakes in Poland, *Acta Geophys*. **62**, 3, 469-485, DOI: 10.2478/s11600-013-0140-2.

Withers, M., R. Aster, C. Young, J. Beiriger, M. Harris, S. Moore, and J. Trujillo (1998), A comparison of select trigger algorithms for automated global seismic phase and event detection, *Bull. Seismol. Soc. Am.* **88**, 1, 95-106.

Received 22 November 2015
Received in revised from 5 April, 2016
Accepted 5 May 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2114-2135
DOI: 10.1515/acgeo-2016-0107

Surface-based Internal Multiple Elimination in the CMP Domain − Theory and Application Strategies on Land Seismic Data

Shiguang DENG, Wenjin ZHAO, and Zhiwei LIU

Chinese Academy of Geological Sciences, Beijing, People's Repulic of China;
e-mails: dengshiguang@live.com, zhaowj@cae.cn,
zwliu007@sina.com (corresponding author)

Abstract

The data-driven internal multiple elimination (IME) method based on feedback model, which includes CFP-based, surface-based and inversion-based methods, are successfully applied to marine datasets. However, these methods are computationally expensive and not always straightforward on land datasets. In this paper, we first proved that the surface-based IME method, which is the most computationally efficient method among the three methods, can be derived from the CFP theory. Then we extend it to CMP domain under the assumption of locally lateral invariance of the earth, which makes it more computationally efficient. In addition, we proposed applying a time-variant taper based on the first Fresnel zone to predict the multiples more percisely. Besides, the improved S/N ratio and dense offset distribution can be obtained by using the CMP supergather, which makes the CMP-oriented method more suitable for land data. Some practical processing strategies are proposed via case study. The effectiveness of the proposed method is demonstrated with the application to synthetic and field data.

Key words: internal multiple, feedback model, surface-based, CMP-oriented, land seismic data.

Ownership: Institute of Geophysics, Polish Academy of Sciences;

1. INTRODUCTION

Multiple reflections affect seismic imaging quality, especially when strong subsurface reflectors exist. The interference of multiple energy with primary events could result in interpretation uncertainties. It is necessary to remove multiples before subsequent processing. In general, multiples consist of free-surface multiples and internal multiples. Free-surface multiples are multiples that have experienced at least one downward reflection at the air-water "free-surface"; internal multiples are multiples that have all of their downward reflections below the free surface. Land seismic data is mainly affected by internal multiples. Internal multiples have experienced reflectors that are in general more remote and harder to precisely define (in comparison with free-surface multiples); hence, internal multiples are more difficult to predict and attenuate (Weglein 1999).

Two major internal multiple elimination (IME) methods, based on wave theory, are the inverse-scattering series (ISS) and feedback methods. The inverse scattering series (ISS) method for internal multiple elimination is discussed specifically by Araujo *et al.* (1994), Coates and Weglein (1996), Weglein *et al.* (1997). The ISS method is fully data-driven and does not require any subsurface information. However, the cost of the ISS approach is considerably greater than the feedback method (Verschuur and Prein 1999). In practice, the feedback method would be a more effective choice. Berkhout and Verschuur firstly proposed the feedback method for the surface-related multiple elimination (SRME) (Berkhout 1982, Verschuur 1991, Verschuur *et al.* 1992, Berkhout and Verschuur 1997, Verschuur and Berkhout 1997). Berkhout and Verschuur (1997) extended the algorithm from surface to internal multiples by replacing shot records with common-focus-point (CFP) gathers (Berkhout 1997, Thorbecke 1997). Berkhout and Verschuur (2005) illustrated the internal-multiple-removal algorithm with numerical examples. This algorithm can be formulated in terms of boundary-related and layer-related versions. Verschuur and Berkhout (2005) demonstrated the strategy for applying the two versions of internal-multiple-removal algorithm on physical-model and field data. The boundary-related approach requires the construction of CFP gathers, using focusing operators with correct traveltimes, while the layer-related approach allows traveltime errors. From a cost perspective, the layer-related method costs twice as much computation time as the boundary-related method, even though the boundary-related method involves more user interactions. Despite the extra calculation cost, the ease of use and the robustness of the layer-related approach make it preferable to the boundary-related approach in most situations. Jakubowicz (1998) proposed the surface-based IME method, in which the need for CFP gathers in the boundary-related method is avoided, and internal multiples can be estimated directly from the data measured on the surface. In van Borselen (2002), an extension of this procedure is illustrated to remove internal multiples that have crossed a pseudo-boundary that is chosen to lie between two internal reflectors. Thus, the data-driven internal multiple elimination based on surface data can be applied either in boundary-related version or layer-related version. Recently, Ypma and Verschuur (2013) redefined internal multiple elimination as a full waveform inversion process, following the principles of estimating primaries by sparse inversion (EPSI) (van Groenestijn and Verschuur 2009a, b). Song *et al.* (2013) compared three feedback IME methods: the CFP-based method, the surface-based method and the inversion-based method. Internal multiples estimated by the inversion-based generalized EPSI method are clearer and more spatially continuous. However, from a cost perspective, the surface-based method is more computationally efficient. The CFP-based method is twice expensive as the surface-based method, and the generalized EPSI is several tens of times more expensive than the surface-based method.

For land data, internal multiple elimination is more difficult than in the marine case. Complex near-surface condition, irregular geometry patterns, poor S/N ratios and source/receiver coupling issues, are key obstacles that deteriorate the performance of multiple attenuation algorithms (Kelamis *et al.* 2006, Luo *et al.* 2011). In this situation, the multiple elimination method in shot domain becomes cumbersome. An alternative choice is to eliminate the internal multiples in the CMP domain if the approximation of a low-relief structure is valid. Yuan *et al.* (2009) described a processing strategy which combines stacking and *f-k* filtering in CMP domain for multiple elimination on land seismic data. Kelamis *et al.* (2002) and Alá'i and Verschuur (2006) apply the CFP-based method in CMP domain for the internal multiple elimination on land data. For the feedback model IME methods, trace interpolation and offset regulation are requried. Such a processing can be more suitable to perform in CMP domain. With the use of CMP supergather, the S/N ratios and spatial sampling density can be improved considerably. In addition, the CMP-oriented method can be applied to both 2D and 3D data sets.

In this paper, we first derived the surface-based IME method based on the CFP theory, showing the relationship between CFP-based and surface-based IME methods. Then we extend the surface-based IME method to CMP domain under the assumption of locally lateral invariance of the earth, which makes it more computationally efficient. In addition, we proposed, by applying a time-variant taper based on the first Fresnel zone, to predict the multiples more percisely. The effectiveness of the proposed CMP-oriented method is demonstrated with synthetic and field land data. Finally, we end up with some discussion and conclusions.

2. THEORY AND METHODOLOGY

Using the detail-hiding operator notation (Berkhout 1982), which uses data matrix (Fig. 1) to represent discrete prestack data volumes for a single frequency, the surface-related multiple prediction can be formulated as

$$\mathbf{M} = A(f)\mathbf{P}_0\mathbf{P} \ , \tag{1}$$

where **M** represents the predicted surface-related multiple, A(*f*) compensates for source and detector properties. $\mathbf{P}_0$ represents the data without surface-related multiples and **P** represents the data with surface-related multiples. Figure 2 illustrates the process of the surface-related multiple prediction in *t*-*x* domain. The multiple removal process can be written in an iterative manner (Berkhout and Verschuur 1997):

$$\mathbf{P}_0^{(i+1)} = \mathbf{P} - A(f)\mathbf{P}_0^{(i)}\mathbf{P} \tag{2}$$

with $\mathbf{P}_0^{(i+1)}$ being the estimated primaries in the (i+1)-th iteration and initial estimate $\mathbf{P}_0^{(0)} = \mathbf{P}$.

For internal multiple elimination, the concept of CFP gather is introduced. A CFP gather represents focused data with one source in the subsurface and all receivers at the surface (or, *vice versa*, for a receiver gather). Thus the prediction of internal multiples will be as same as the prediction of surface-related multiples (Fig. 3).

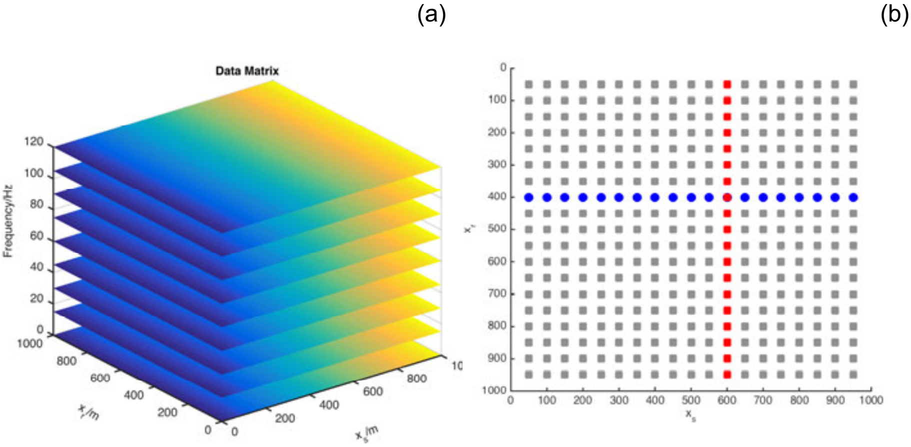


Fig. 1. Data matrices for 2-D seismic data volumes: (a) each slice represents a monochromatic data matrix, (b) in the matrix, the columns contain monochromatic shot records and the rows contain monochromatic common-receiver gathers.

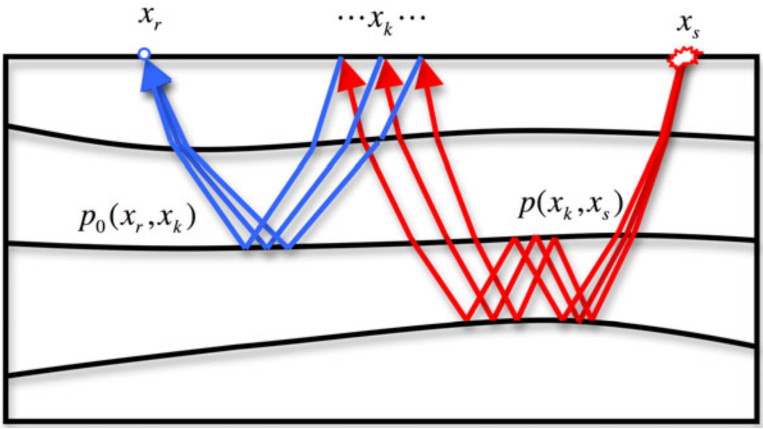


Fig. 2. Surface-related multiple from source x_s to receiver x_r can be seen as the sum of the convolution of $p_0(x_r, x_k)$ and $p(x_k, x_s)$ for each trace in *t-x* domain, where $p_0(x_r, x_k)$ is the data without multiple and $p(x_k, x_s)$ is the data with multiple (k is variable).

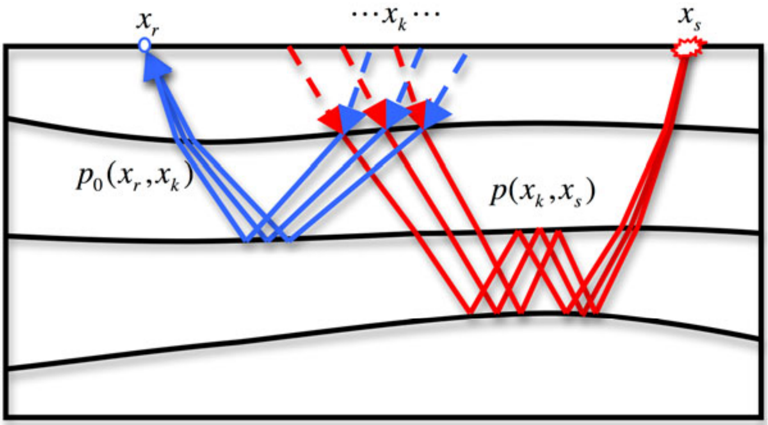


Fig. 3. Internal multiple prediction $p_0(x_r, x_k)$ and $p(x_k, x_s)$ are CFP gathers with sources and receivers at the internal multiple-generating surface, respectively.

Focusing in detection can be formulated as

$$\mathbf{P}(z_n, z_0) = \left[\mathbf{F}(z_n, z_0)\right]^* \mathbf{P}(z_0, z_0) \ . \tag{3}$$

In Eq. 3, z_0 represents the surface and z_n represents the internal multiple-generating surface. $\mathbf{P}(z_n, z_0)$ represents the focused gather whose source is at the surface and receivers are at z_n. $\mathbf{P}(z_0, z_0)$ represents the shot gather whose source and receivers are both at the surface. $\mathbf{F}(z_n, z_0)$ represents the focusing

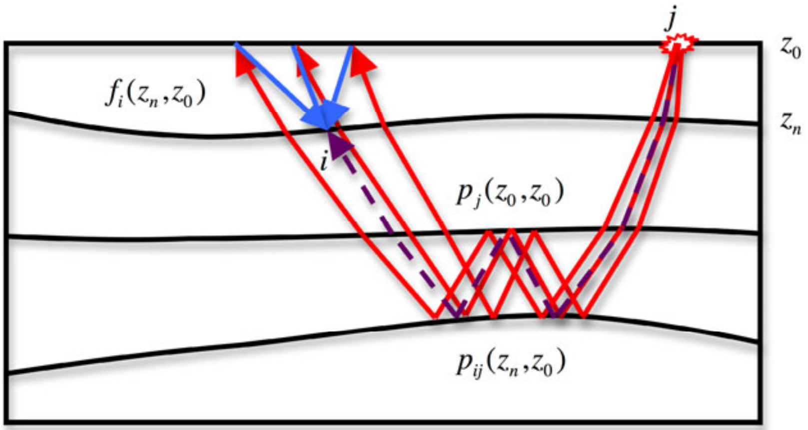


Fig. 4. Construction of one CFP trace for focusing in detection. $p_j(z_0, z_0)$ is the shot gather with its source positioned at (x_j, z_0) and the receivers positioned at the surface. $f_i(z_n, z_0)$ is the Green's function of the virtual receiver. CFP trace $p_{ij}(z_n, z_0)$, whose source positioned at (x_j, z_0) and receiver positioned at (x_i, z_n), is obtained by summing the time convolution of the shot gather and the time revered Green's function along the spatial axes.

operator. Each row of **F** contains the Green's function of the virtual receiver measured at the surface. The focusing process is actually an inverse extrapolation process, thus the time reversed Green's function should be used. The complex conjugate of the Green's function in frequency domain, which is denoted with the superscript * in Eq. 3, stands for time reversal in time domain. In Fig. 4, the process of focusing in receivers is illustrated with ray paths. Similarly, focusing in emission can be formulated as

$$\mathbf{P}(z_0, z_n) = \mathbf{P}(z_0, z_0)\left[\mathbf{F}(z_0, z_n)\right]^* \tag{4}$$

where $\mathbf{P}(z_0, z_n)$ represents the focused gather whose receivers are at the surface and source is at z_n. Each column of the focusing operator **F** contains the Green's function of the virtual source measured at the surface. According to the feedback algorithm, the internal multiples related to boundary z_n removal process can be written as

$$\mathbf{P}_n^{(i)}(z_0, z_0) = \mathbf{P}_{n-1}(z_0, z_0) - A_n(f)\overline{\mathbf{P}}_n^{(i-1)}(z_0, z_n)\overline{\mathbf{P}}_{n-1}(z_n, z_0) \tag{5}$$

In Eq. 5, $\mathbf{P}_n(z_0, z_0)$ represents the seismic data with all primary reflections and internal multiples for $z > z_n$ only. $A_n(f)$ is the match filter between the predicted and actual multiples. $\overline{\mathbf{P}}_n(z_0, z_n)$ represents the seismic data focused in emission with its source positioned at depth level z_n, with all reflections (primaries and multiples) up to level z_n removed. $\overline{\mathbf{P}}_{n-1}(z_n, z_0)$ represents the seismic data focused in detection with its receivers positioned at depth level z_n, with all the primaries up to level z_n removed and all the multiples up to level z_{n-1} removed. Iteration number is represented by i. Internal multiples are generally weaker than surface multiples, so one iteration is often sufficient. Note that the internal multiple elimination should be performed from shallow to deep.

Jakubowicz (1998) proposed the surface-based IME method based on the geometry analysis by Keydar *et al.* (1997). In the surface-based IME method, CFP gather is not needed anymore; the internal multiples can be estimated directly from measured data at the surface. According to the CFP theory discussed above, we can derive the surface-based IME method based on CFP theory. In this paper, we give a mathematic derivation for surface-based IME method based on CFP theory. According to Eq. 4, the internal multiples are given by

$$\mathbf{M}_n^{(i)}(z_0, z_0) = \mathbf{P}_n^{(i-1)}(z_0, z_n)\mathbf{P}_{n-1}(z_n, z_0). \tag{6}$$

Using Eqs. 3 and 4, we can also write Eq. 6 as

$$\mathbf{M}_n^{(i)}(z_0, z_0) = \overline{\mathbf{P}}_n^{(i-1)}(z_0, z_0)\left[\Delta\mathbf{P}_n(z_0, z_0)\right]^* \overline{\mathbf{P}}_{n-1}(z_0, z_0) \tag{7a}$$

with

$$\left[\Delta\mathbf{P}_n(z_0, z_0)\right]^* = \left[\mathbf{F}(z_0, z_n)\right]^* \left[\mathbf{F}(z_n, z_0)\right]^* \tag{7b}$$

where $\mathbf{\Delta P}_{\boldsymbol{n}}(z_0, z_0)$ represents the primary reflection from the multiple-generating interface z_n (the yellow dashed line in Fig. 5). Jakubowicz names $\overline{\mathbf{P}}_n^{(i-1)}(z_0, z_0)$, $\mathbf{\Delta P}_{\boldsymbol{n}}(z_0, z_0)$, and $\overline{\mathbf{P}}_{n-1}(z_0, z_0)$ as "source term", "interbed primary", and "response term", respectively. According to the derivation, we can find that the combination of the two focusing operators is equal to the time-reversed reflection of the corresponding reflector. In addition, if we make the "interbed primary" also including the primaries above this boundary, then this method becomes a layer-related version.

Through the theoretical derivation, we proved that the CFP-based and the surface-based IME methods are based on the same core principles but implemented in different forms. Obviously, the surface-based IME method is much more efficient for implementation.

If the approximation of a low-relief structure is valid, the CMP-oriented method can be applied as well. This method is applied on pre-stack data in the CMP domain under the assumption of locally lateral invariance of the earth. This means the multiple prediction process can be carried out on indi-

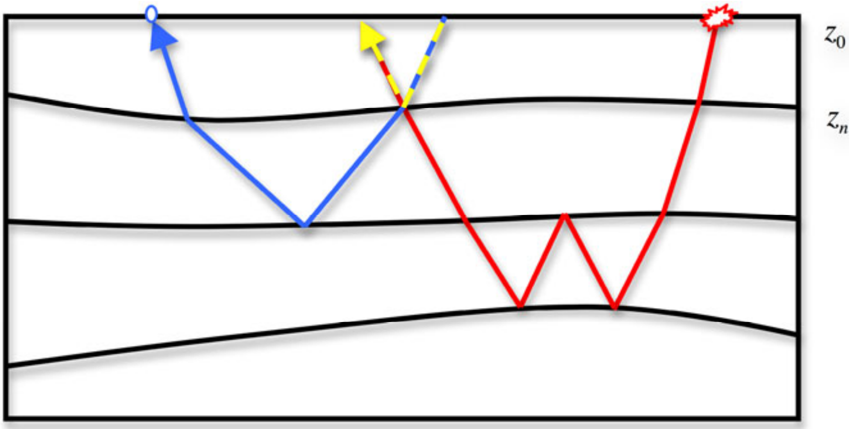


Fig. 5. Internal multiple prediction with the data measured at the surface.

vidual CMP gathers after Fourier transform to the *f-k* domain. In this case, matrix multiplications in *f-x* domain reduced to scalar multiplications in *f-k* domain, and the CMP-oriented internal multiple prediction can be formulated as

$$M_n^{(i)}\left(k_x,\omega\right)=\bar{P}_n^{(i-1)}\left(k_x,\omega\right)\left[\Delta P_n\left(k_x,\omega\right)\right]^{\#}\bar{P}_{n-1}\left(k_x,\omega\right) . \tag{8}$$

In Eq. 8, $\bar{P}_n(k_x,\omega)$ represents the seismic data with all primary reflections and internal multiples for $z > z_n$ only. $\left[\Delta P_n(k_x,\omega)\right]^{\#}$ denotes the time reversed primary reflection of the internal multiple-generating surface in *f-k* domain. Note that the time-reversed data in *f-k* domain cannot be directly acquired. Two steps should be taken: first, the complex conjugation of the data is calculated in *f-x* domain, and then the spatial Fourier transform is performed; this process is denoted with the superscript #. The CMP-oriented method for internal multiple elimination can be a practical alternative to shot-oriented method, especially when the low-relief structure is valid.

The CMP-oriented method is ideally suited for land data. Land seismic data generally have poor S/N ratio; however, the feedback model IME method uses the reflection data itself to predict multiples. That means any none-reflection signals, such as direct wave, surface wave and random noise, will affect the predicted result. Besides, regular offset of the input data is required as well. In CMP domain, a group of neighboring CMP gathers can be merged into a supergather. The CMP supergather has a dense offset distribution and less noise. Offset regulation and trace interpolation could be much easier to perform in a supergather. With the high S/N ratio and well-sampled supergather, multiples can be better predicted.

3. SYNTHETIC DATA EXAMPLE

For the demonstration of CMP-oriented internal multiple elimination, both 1D and 2D synthetic data sets are simulated. The time-variant taper based on the first Fresnel zone, which can reduce edge effect more efficiently in the predicted multiples, is proposed and discussed in detail via 1D synthetic data. The 2D anticline structure model is designed to test the effectiveness of the CMP-oriented IME method in a relatively complex structure. In order to focus on the internal multiple elimination, surface multiples are not generated in the synthetic data.

3.1 1D model

The 1D model is shown in Fig. 6. The CMP gather contains 201 traces in a split-spread configuration with offsets from –500 to 500 m relative to the source position and the trace interval is 5 m. The synthetic CMP gather is displayed in Fig. 7. Besides the three primaries, internal multiples are clearly visible. The internal multiples have been indicated in the figure by a sequence of bounces (upward-downward-upward, *etc*.).

Multiple elimination is an iterative process; however, one iteration is sufficient for internal multiple elimination in practice. Thus, only the interbed primary and the response term are required for the prediction (Fig. 8). The two data sets can be obtained directly by applying inside and (or) outside mute to the raw data, but it is more practical to mute the NMO corrected data, and the inverse NMO correction should be applied after the mute.

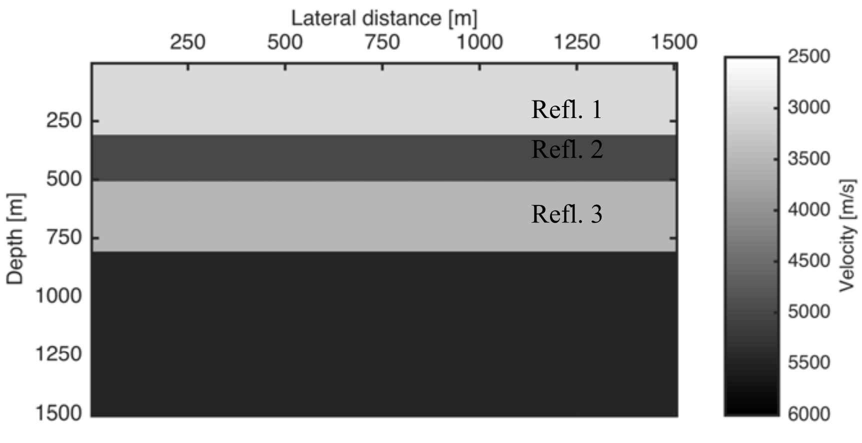


Fig. 6. The 1D subsurface model; a relative high-velocity layer occurs at the depth around 300 to 500 m.

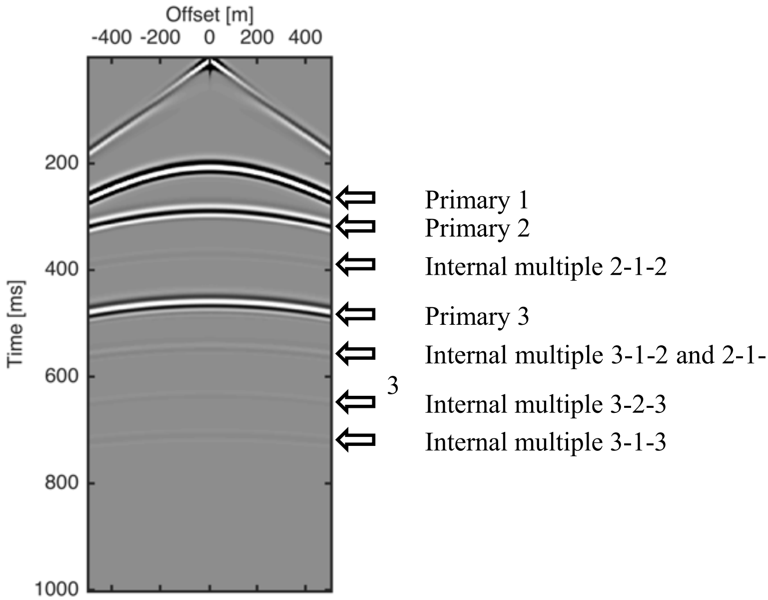


Fig. 7. The synthetic CMP gather.

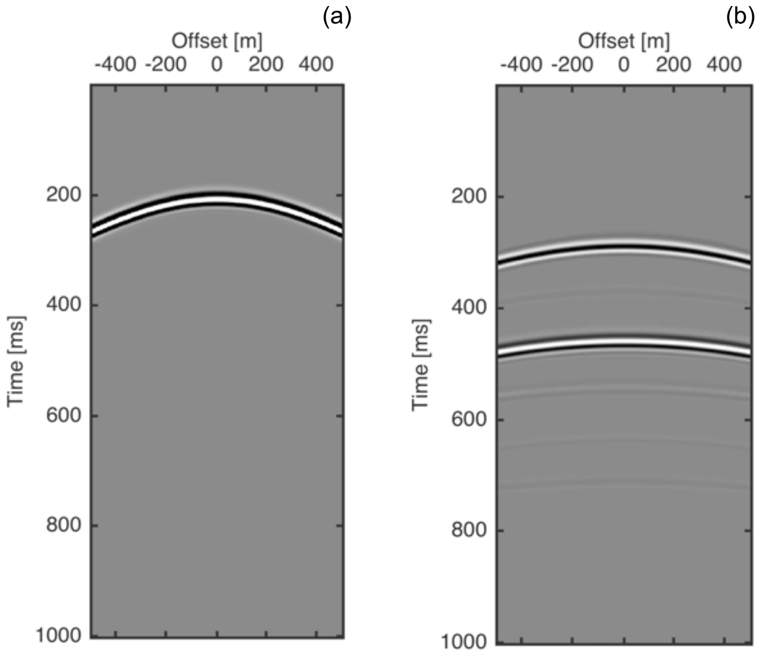


Fig. 8. The wavefield terms used in predicting the internal multiples generated by the first reflector of the data. The interbed primary (a) and response term (b) can be obtained directly by muting the initial data.

Multiple elimination involves prediction and subtraction. First, apply multiple prediction by multiplying the prepared data in the *f-k* domain. In practical prediction, the limited aperture of the data may cause edge effects in the predicted multiples, with genuine multiple events appearing to reflect back into the data (Fig. 9a). The first Fresnel zone plays an important role in multiple prediction; however, the energy outside the first Fresnel zone will actually introduce artifacts in the predicted results when the aperture is limited. Thus, tapering the amplitudes outside the first Fresnel zone would be an efficient way to reduce these edge effects. Generally, the radius of the first Fresnel zone will increase as the depth goes deeper, so a time variant taper would be more proper for the reduction of the edge effects. In this paper, a time variant taper function is designed as follows:

$$\begin{cases} w = 1, x \le R(t) \\ w = \cos\left(\frac{\pi}{2}\left(\frac{x - R(t)}{x_{\max} - R(t)}\right)\right)^2, x > R(t) \end{cases} \tag{9a}$$

with

$$R(t) = \sqrt{\left(v_t t + \frac{v_t}{4f}\right)^2 - \left(v_t t\right)^2} \tag{9b}$$

where w represents the weighting coefficient of the sample point positioned at time t and offset x. $R(t)$ represents the radius of the First Fresnel zone at time t. $x_{\max}$ is the max offset distance in the gather. We can estimate the approximate radius of the first Fresnel zone from Eq. 9a. In Eq. 9a, v_t represents the root mean square velocity at time t, f represents the main frequency of the seismic data. Only approximate velocity and frequency are required as the results are not very sensitive to these parameters. Figure 9b shows the predicted multiples with the use of a constant taper. The edge effects are reduced significantly; however, the multiples at far offset become weak, especially at greater depths. With the use of the time variant taper, the energy of the multiples at far offset is preserved (Fig. 9c).

The next step is to subtract the predicted multiples from the input data. As amplitude and phase errors often existed in the predicted multiples, a two-step adaptive subtraction (Verschuur and Berkhout 1997) is applied in this procedure. First, estimate a long filter (*i.e.*, typical 21 to 31 points) for optimization of the whole gather, and then use that result for a second adaptation step within local time and/or offset windows with smaller filters (*i.e.*, typicaly 5 points). The result of this subtraction is shown in Fig. 10: the input CMP gather with internal multiples (Fig. 10a), the result after multiple removal (Fig. 10b), and the estimated internal multiples (Fig. 10c).

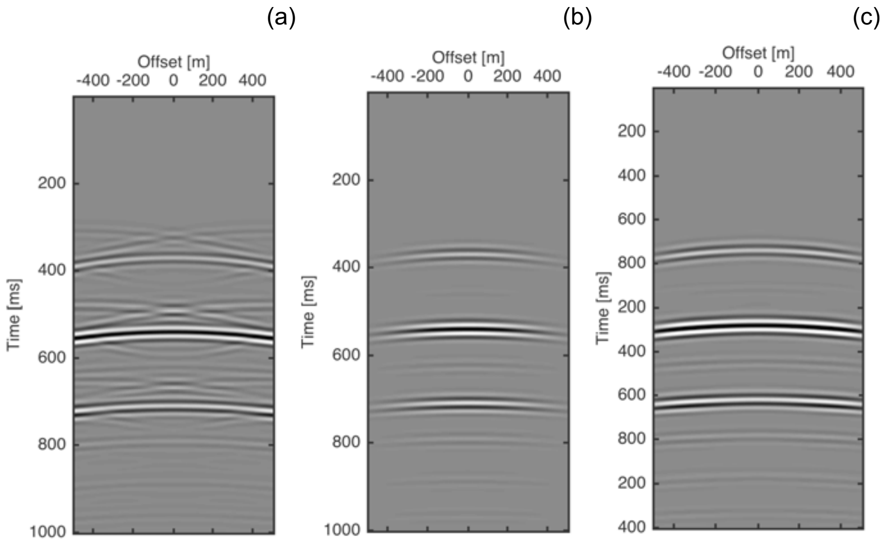


Fig. 9. The predicted multiples (a) are affected by the edge effects severely. With the use of a constant taper, the edge effects in the predicted multiples (b) are reduced significantly, whereas the multiples at far offset are also reduced. Using the time variant taper can preserve the energy of multiples at far offset (c).

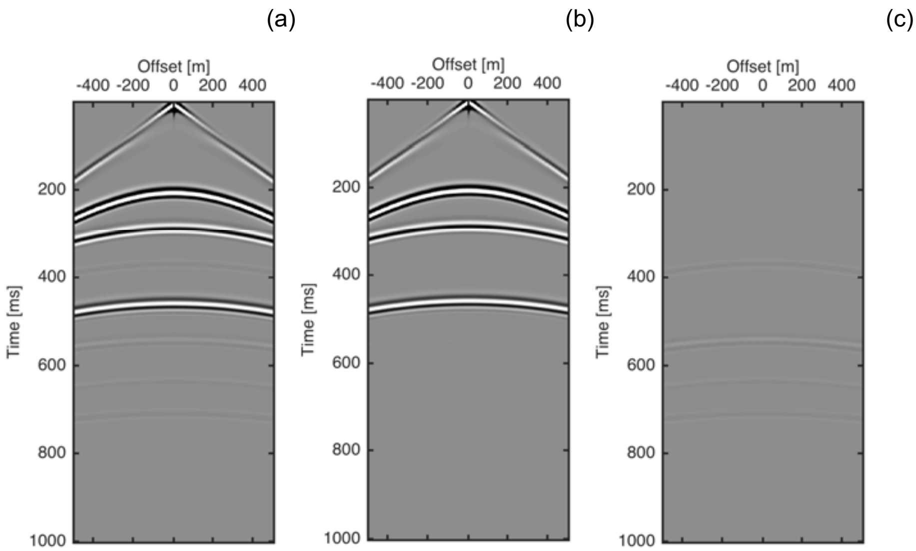


Fig. 10. CMP-oriented internal multiple removal for the 1D data set: (a) CMP gather with internal multiples, (b) CMP gather after internal multiple elimination, (c) internal multiples removed.

3.2 2D model

The anticline structure model, shown in Fig. 11, is designed to test the effectiveness of the CMP-oriented method in 2D situation. The same procedure is applied on this data set. Figure 12 shows the multiple elimination results of 3

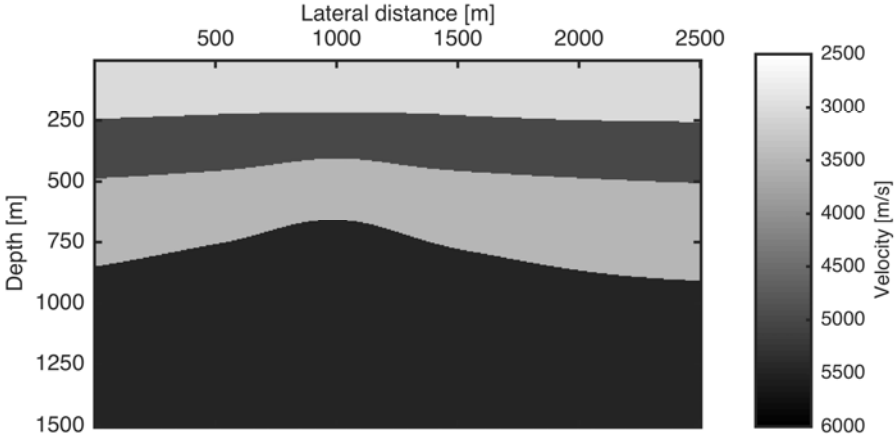


Fig. 11. The anticline structure model.

(a)

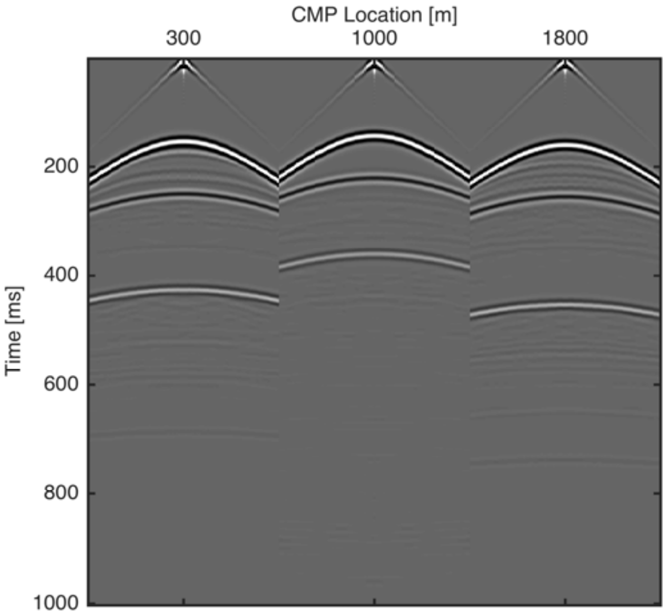
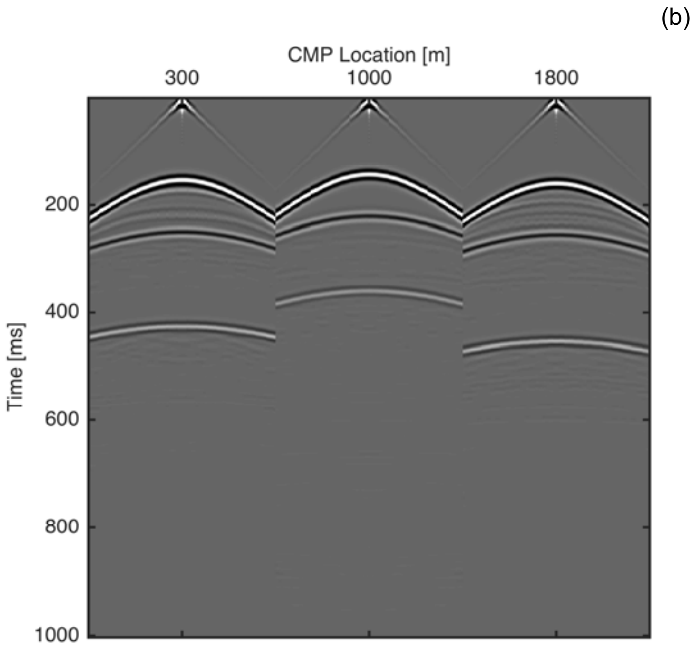


Fig. 12. Continued on next page.



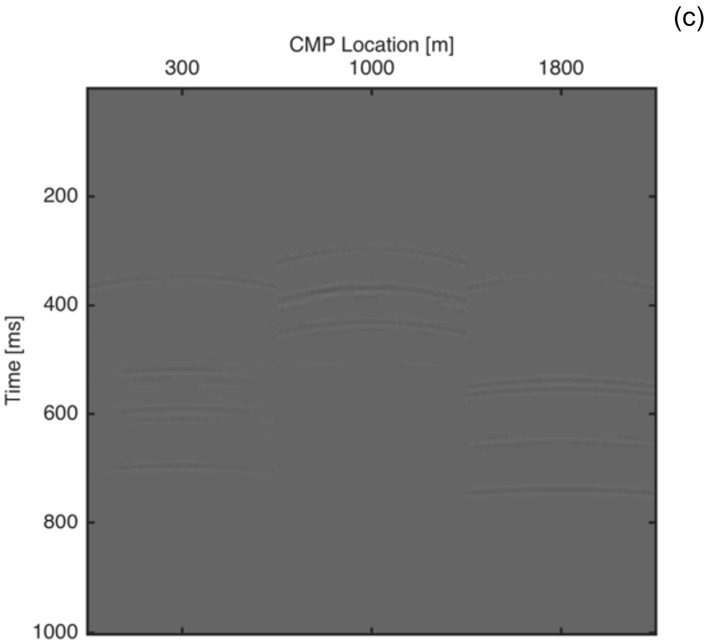


Fig. 12. CMP gathers at different locations, before (a) and after (b) internal multiple attenuation, and the adaptively subtracted internal multiples (c).

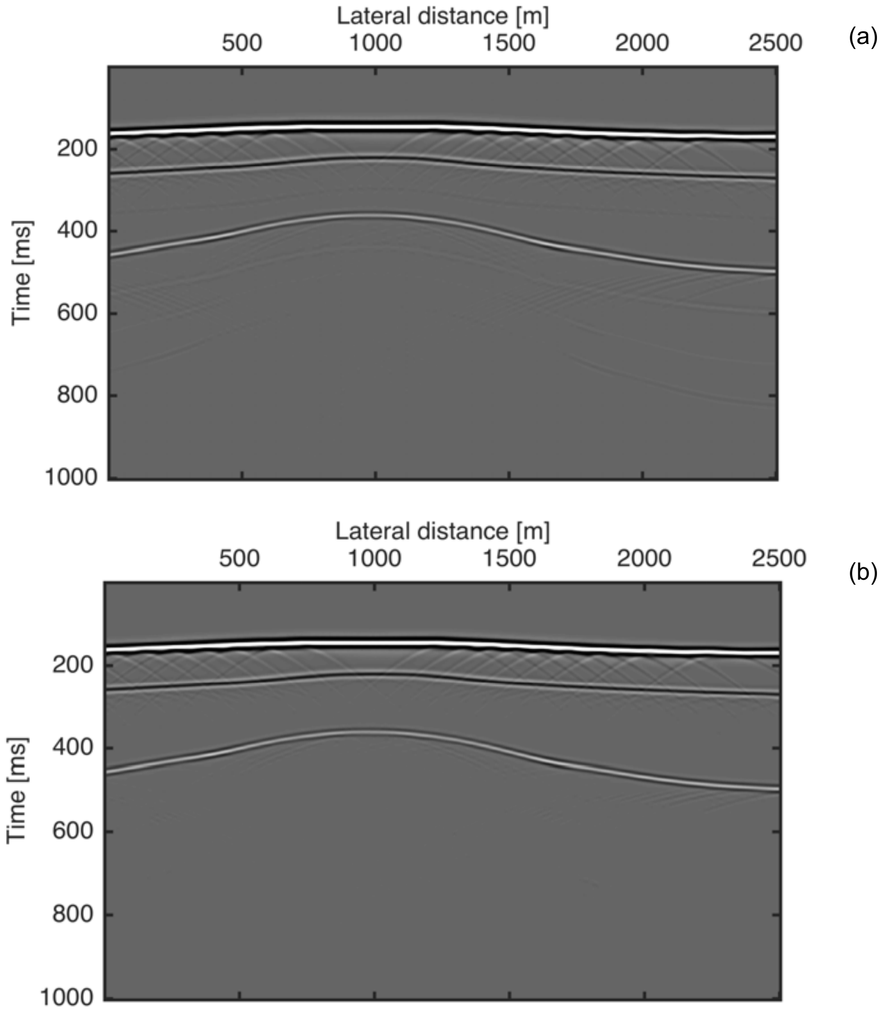


Fig. 13. Stack results before (a) and after (b) the internal multiple elimination.

CMP gathers at different locations. Note that the diffractions have also existed in the input data (Fig. 12a). The CMP gathers after multiple removal is shown in Fig. 12b, and the removed multiples are shown in Fig. 12c. After applying the CMP-oriented internal multiple suppression, the CMP gathers are stacked (Fig. 13). The internal multiples can be observed in the input stacked section (Fig. 13a) clearly. Figure 13b shows the primaries-only stack; the internal multiples are cleanly removed. From this example, we can see the CMP-oriented method can still produce good results when the medium is not totally 1D.

4. FIELD DATA EXAMPLE

When dealing with the land seismic data, internal multiple elimination becomes much more difficult. Low S/N ratio, irregular offset and missing traces impose additional constrains in the prediction of multiples. Some key techniques could be applied to solve these problems effectively in the CMP domain; therefore, the CMP-oriented IME method would be more appropriate for land seismic data. The detailed procedure of preprocessing is demonstrated as follows. In addition, surface-related multiples are proved to be not existing in the filed land seismic data we used in this paper via SRME method.

Traditional preprocessing steps include static correction, noise reduction and offset regulation. High frequency static corrections should be applied to make the reflections more coherent. Ground roll could be suppressed with an *f-k* filter. Linear noise, such as refractions, can be removed with the radial trace filter. Then the CMP gathers are regularized via NMO correction and trace repositioning. The CMP gather after traditional preprocessing is shown in Fig. 14a, but it is still not good enough to obtain a satisfactory result. By merging a group of neighboring CMP gathers (in this case 7) to one supergather, a dense offset distribution is simulated and the S/N ratio of the gather is improved as well (Fig. 14b). For more complex structures, a smaller number of neighboring CMP gathers, such as 5 or 3, is recommended; for merging the supergather, we should balance the tradeoff between noise attenuation and detail preservation. Interpolation is performed after applying an approximate NMO correction to the supergather, and then we apply a lateral smoothing to reduce the random noise. At last, an inverse NMO correction is applied. Figure 14c shows the well prepared CMP supergather for the multiple prediction.

In addition, before the multiple prediction, the deep reflection energy of the response term should be attenuated to avoid the extra predicted multiples emerging at the top of the predicted results due to the limitation of the acquisition time (Fig. 15).

Note that, in this field data, weak reflectors (above 1000 ms) did not induce significant internal multiples; several strong reflectors generated most of the internal multiples. Thus, even though the layer-related version can predict the multiples related to more boundaries at one time, taking all boundaries into consideration in the multiple prediction process is unnecessary and sometime it could even decrease the effectiveness of the adaptive subtraction because the adaptive filter has to take care of both significant multiples generated by strong reflectors and insignificant multiples generated by weak reflectors at the same time. In this situation, the boundary-related version is more targeted to remove the specific internals generated by a cer-

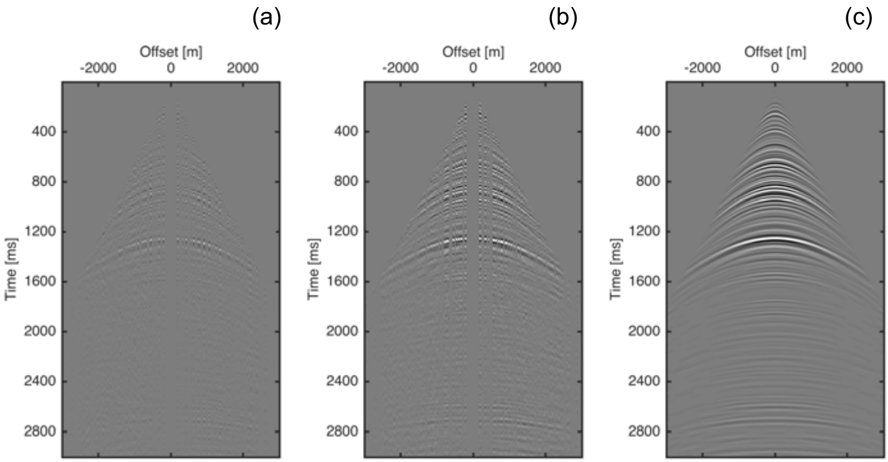


Fig. 14: (a) A CMP gather after regular preprocessing, plotted on true offset; (b) a CMP supergather after merging 7 CMP gathers; (c) supergather after interpolation and random noise reduction.

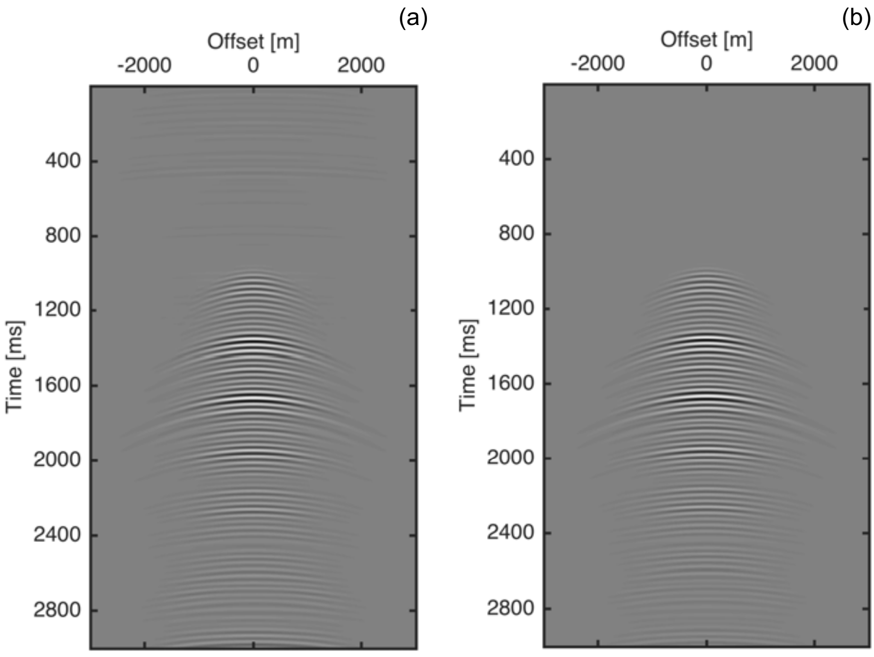


Fig. 15. Predicted results with extra multiples at the top (a), after applying the attenuation, the predicted multiples are free from the interference of the extra multiples (b).

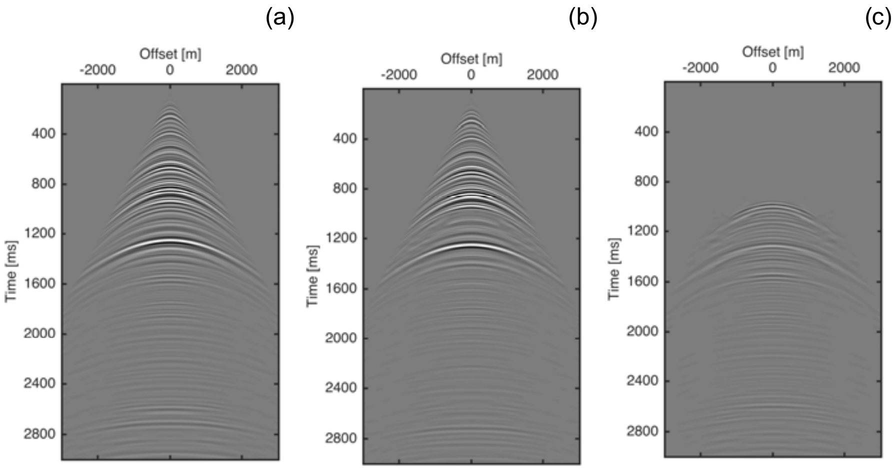


Fig. 16. CMP-oriented multiple removal for land seismic data: (a) CMP gather with multiples; (b) CMP gather after internal multiple suppression; (c) internal multiples removed.

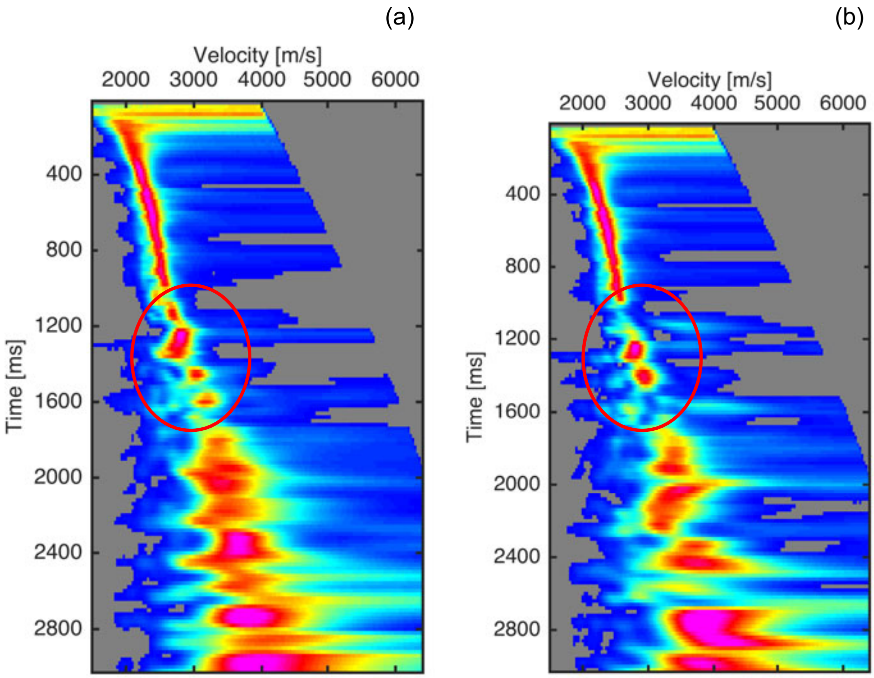


Fig. 17. The velocity spectrums of the CMP gathers before (a) and after (b) internal multiple elimination.

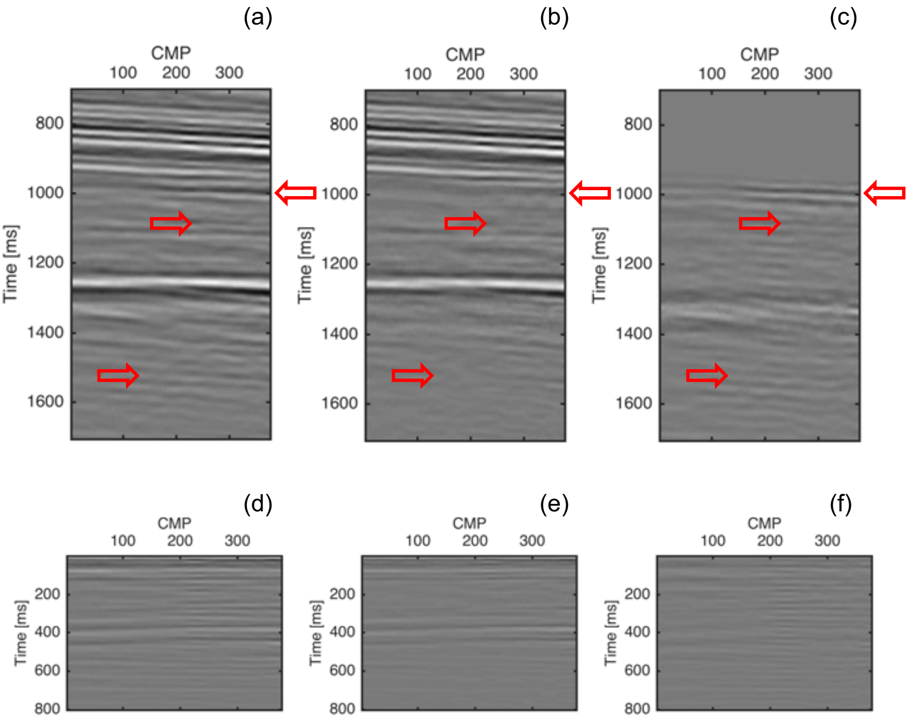


Fig. 18. Stacked land section before (a) and after (b) internal multiple elimination, and (c) the adaptively subtracted internal multiples. The autocorrelations of these stack results are shown in (d), (e), and (f), respectively.

tain strong reflector. In this example, three strong internal multiple-generating surfaces are selected to predict the internal multiples. The result of the internal multiple elimination is shown in Fig. 16, and the velocity spectrums of two CMP gathers (before and after multiple elimination) are shown in Fig. 17. The relatively low velocity energy, for example the energy occurred at 1100 and 1300 ms in Fig. 17a, generally implies the existence of internal multiples, and, sometimes, internal multiples also experience similar (or even higher) velocities to primaries in their vicinity, such as the energy occurred at 1600 ms. In Fig. 17b, the multiple energy is significantly removed, thus reducing a lot of the picking ambiguity. Figure 18a shows a local stack result of this land data set before internal multiple elimination. Several strong reflectors, which occurred at 850 and 1250 ms, particularly induced the generation of the internal multiples. The autocorrelation of this stack section is shown in Fig. 18d, periodic energy implies the existence of the multiples. The primaries-only stack is shown in Fig. 18b, the internal multiples are significantly removed. For the convienience of comparison, red arrows are added to the stack sections (Fig. 18). Periodic energy is apparently reduced in its autocorrelation, as shown in Fig. 18e. Figure 18c shows the removed internal multiples, and its autocorrelation is shown in Fig. 18f.

5. CONCLUSIONS

We have successfully extended the data-driven IME method to the CMP domain. For low-relief structures, the CMP-oriented method is more suitable for land seismic data with more challenging problems than marine data.

Proper preprocessing of data is essential in order to obtain satisfactory results. The raw data should be preprocessed with high-frequency static corrections, noise reduction, offset regulation, CMP supergather composition, interpolation and random noise attenuation sequentially to improve the S/N ratio and densify the offset distribution. When the predicted results are plagued with edge effects, attenuating the far offset energy with the time variant taper will effectively reduce the edge effects, without attenuating the predicted multiples. In addition, the multiples caused by deep reflectors may occur at the top of the predicted results due to the limitation of acquisition time, and this could be avoided by simply attenuating the deep reflection energy of the response term. These techniques improved the accuracy of the predicted multiples and therefore the multiples can be adaptively subtracted more effectively. The boundary-related method is more suitable and targeted for internal multiple elimination of the seismic data most of whose internal multiples are generated only by several strong reflectors. The field data example shows that satisfactory results can be obtained by just taking several strong subsurface reflectors into consideration.

The CMP-oriented IME method is applied to synthetic data sets and a land field data set; the application examples demonstrate the effectiveness of the proposed methodology.

Acknowledgments. We greatly appreciated the constructive suggestions from editors and anonymous reviewers. This work is supported by the Geologic Survey Project of China (Grant No. DD20160161), and the Fundamental Research Funds for the Central Universities of Chian (Grant No. YYWF201506).

References

Alá'i, R., and D.J. Verschuur (2006), Case study of surface-ralated and internal multiple elimination on land data. **In:** *76th Annual International Meeting, SEG*, Expanded abstracts.

Araujo, F.V., A.B. Weglein, P.M. Carvalho, and R.H. Stolt (1994), Inverse scattering series for multiple attenuation: An example with surface and internal multiples. **In:** *64th Annual International Meeting, SEG*, Expanded abstracts, 1039-1041.

Berkhout, A.J. (1982), *Seismic Migration: Imaging of Acoustic Energy by Wavefield Extrapolation. A. Theoretical Aspects*, 2nd ed., Elsevier Science Publ. Co., Amsterdam.

Berkhout, A.J. (1997), Pushing the limits of seismic imaging. Part I: Prestack migration in terms of double dynamic focusing, *Geophysics* **62**, 3, 937-953, DOI: 10.1190/1.1444201.

Berkhout, A.J., and D.J. Verschuur (1997), Estimation of multiple scattering by iterative inversion. Part I: Theoretical considerations, *Geophysics* **62**, 5, 1586-1595, DOI: 10.1190/1.1444261.

Berkhout, A.J., and D.J. Verschuur (2005), Removal of internal multiples with the common-focus- point (CFP) approach – Part 1: Explanation of the theory, *Geophysics* **70**, 3, V45-V60, DOI: 10.1190/1.1925753.

Coates, R.T., and A.B.Weglein (1996), Internal multiple attenuation using inverse scattering: Results from prestack 1 and 2D acustic and elastic synthetics. **In:** *66th Annual International Meeting, SEG*, Expanded abstracts, 1522-1525.

Jakubowicz, H. (1998), Wave equation prediction and removal of interbed multiples. **In:** *68th Annual International Meeting, SEG*, Expanded abstracts, 1527-1530, DOI: 10.3997/2214-4609.201408173.

Kelamis, P.G., D.J. Verschuur, K.E. Erickson, R.L. Clark, and R.M. Burnstad (2002), Data-driven internal multiple-attenuation applications and issues on land data. **In:** *72nd Annual International Meeting, SEG*, Expanded abstracts, 2035-2038, DOI: 10.1190/1.1817099.

Kelamis, P.G., W.H. Zhu, K.O. Rufaii, and Y. Luo (2006), Land multiple attenuation – the future is bright. **In:** *76th Annual International Meeting, SEG*, Expanded abstracts.

Keydar, S., E. Landa, B. Gurevich, and B. Gelchinsky (1997), Multiple prediction using wavefront characteristics of primary reflections, **In:** *59th Conference, EAGE, Geneva, Switzerland*, Extended abstracts, Paper A016.

Luo, Y., P.G. Kelamis, Q. Fu, S.G. Huo, G. Sindi, S. Aramco, S.Y. Hsu, and A.B. Weglein (2011), Elimination of land internal multiples based on the inverse scattering series, *The Leading Edge* **30**, 8, 884-889, DOI: 10.1190/ 1.3626496.

Song, J.W., E.Verschuur, and X.Chen (2013), Comparing three feedback internal multiple elimination methods, *J. Appl. Geophys*. **95**, 66-76, DOI: 10.1016/j.jappgeo.2013.05.010

Thorbecke, J.W. (1997), Common focus point technology, Ph.D. Thesis, Delft University of Technology, Delft, The Netherlands.

van Borselen, R.G. (2002), Fast-track, Data-driven interbed multiple removal – a North Sea data example. **In:** *64th EAGE Conference and Exhibition 2002*, Extended abstracts.

van Groenestijn, G.J.A., and D.J. Verschuur (2009a), Estimating primaries by sparse inversion and application to near-offset data reconstruction, *Geophysics* **74**, 3, A23-A28, DOI: 10.1190/1.3111115.

van Groenestijn, G.J.A., and D.J. Verschuur (2009b), Estimation of primaries and near offsets by sparse inversion: marine data applications, *Geophysics* **74**, 6, R119-R128, DOI: 10.1190/1.3213532.

Verschuur, D.J. (1991), Surface-related multiple elimination, an inversion approach, Ph.D. Thesis, Delft University of Technology, Delft, The Netherlands.

Verschuur, D.J., and A.J. Berkhout (1997), Estimation of multiple scattering by iterative inversion. Part II: Practical aspects and examples, *Geophysics* **62**, 5, 1596-1611, DOI: 10.1190/1.1444262.

Verschuur, D.J., and A.J. Berkhout (2005), Removal of multiples with the common-focus-point (CFP) approach – Part 2: Application strategies and data examples, *Geophysics* **70**, 3, V61-V72, DOI: 10.1190/1.1925754.

Verschuur, D.J., and R.J. Prein (1999), Multiple removal results from Delft University, *The Leading Edge* **18**, 1, 86-91, DOI: 10.1190/1.1438164.

Verschuur, D.J., A.J. Berkhout, and C. Wapenarr (1992), Adaptive surface-related multiple elimination, *Geophysics* **57**, 9, 1166-1177, DOI: 10.1190/1.1443330.

Weglein, A.B. (1999), Multiple attenuation: an overview of recent advances and the road ahead, *The Leading Edge* **18**, 1, 40-44, DOI: 10.1190/1.1438150.

Weglein, A.B., F. Gasparotto, P. Carvalho, and R.H. Stolt (1997), An inverse-scattering series method for attenuating multiples in seismic reflection data, *Geophysics* **62**, 6, 1975-1989, DOI: 10.1190/1.1444298.

Ypma, F.H.C., and D.J.Verschuur (2013), Estimating primaries by sparse inversion, a generalized approach, *Geophys. Prospect.* **61**, s1, 94-108, DOI: 10.1111/j.1365-2478.2012.01095.x

Yuan, Y., D. Han, and R. Zhang (2009), Multiple suppression on land seismic data – case history. **In:** *79th Annual International Meeting, SEG*, Expanded abstracts, The Society of Exploration Geophysicists, Tulsa, USA, 3093-3097.

Received 16 November 2015
Received in revised form 1 March 2016
Accepted 26 July 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2136-2150
DOI: 10.1515/acgeo-2016-0106

Analysis of Coseismic Fault Slip Models of the 2012 Indian Ocean Earthquake: Importance of GPS Data for Crustal Deformation Studies

Endra GUNAWAN[1], Putra MAULIDA[2], Irwan MEILANO[2,3], Masyhur IRSYAM[3], and Joni EFENDI[4]

[1]Graduate Research on Earthquake and Active Tectonics, Faculty of Earth Science and Technology, Bandung Institute of Technology, Bandung, Indonesia
e-mail: endra@lppm.itb.ac.id

[2]Geodesy and Geomatics Engineering, Faculty of Earth Science and Technology, Bandung Institute of Technology, Bandung, Indonesia

[3]Research Center for Disaster Mitigation, Bandung Institute of Technology, Bandung, Indonesia

[4]Geospatial Information Agency, Cibinong, Indonesia

Abstract

Based on continuous GPS data, we analyze coseismic deformation due to the 2012 Indian Ocean earthquake. We use the available coseismic slip models of the 2012 earthquake, derived from geodetic and/or seismic waveform inversion, to calculate the coseismic displacements in the Andaman-Nicobar, Sumatra and Java. In our analysis, we employ a spherical, layered model of the Earth and we find that Java Island experienced coseismic displacements up to 8 mm, as also observed by our GPS network. Compared to coseismic offsets measured from GPS data, a coseismic slip model derived from multiple observations produced better results than a model based on a single type of observation.

Key words: the 2012 Indian Ocean earthquake, coseismic deformation, GPS displacement.

1. INTRODUCTION

On 11 April 2012, a M_w8.6 Indian Ocean earthquake (IOE) occurred 300 km west of the Sunda trench (Meng *et al.* 2012). It was part of an earthquake sequence which involved a preshock of M_w7.3 on 10 January 2012, a mainshock of M_w8.6 on 11 April 2012, and an aftershock of M_w8.2 two hours after the mainshock (Duputel *et al.* 2012). Considered to be the largest known strike-slip intraplate earthquake event, it has been suggested that the 2012 IOE occurred as a result of stress transfer after the 2004 M_w9.2 Sumatra-Andaman earthquake and the 2005 M_w8.7 Nias earthquake (Delescluse *et al.* 2012). The Indonesian Agency for Meteorology, Climatology and Geophysics (BMKG) reported that aftershock of $M > 3$ occurred practically continuously, with up to 400 events registered for six months after the 2012 IOE (Fig. 1).

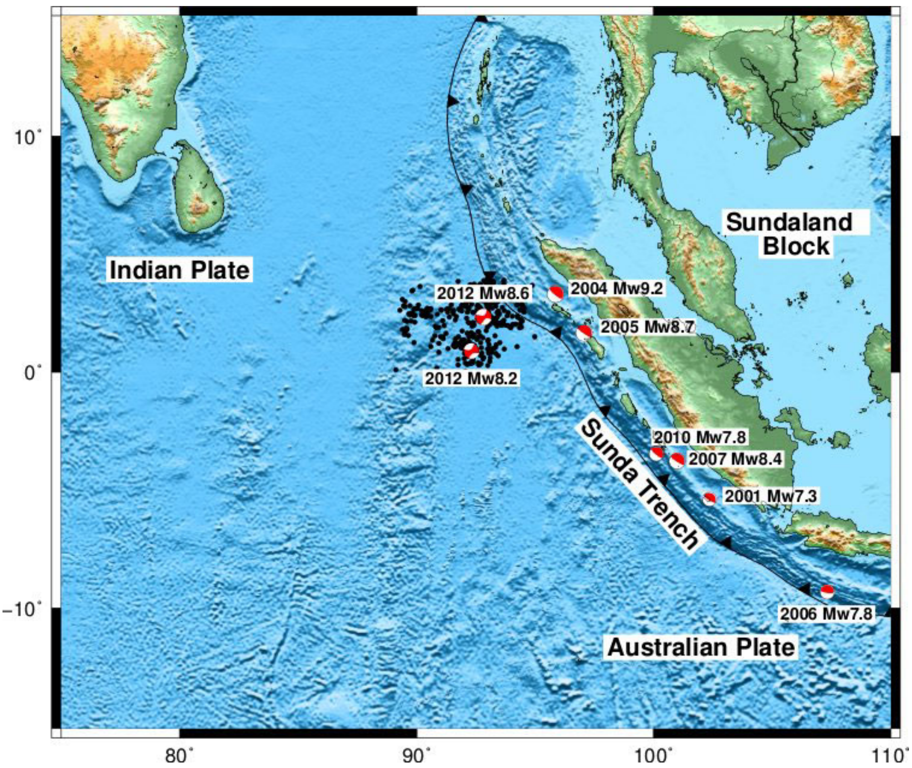


Fig. 1. Tectonic background of this study. Beach balls represent earthquake occurrences offshore the western side of northern Sumatra over the last decade from the Global CMT catalogue. Black dots indicate six-month aftershocks of the 2012 IOE based on the Indonesian Agency for Meteorology, Climatology and Geophysics (BMKG) catalogue.

Global Positioning System (GPS) data are an indispensable tool for analyzing the crustal deformation of the Earth. Previous crustal deformation studies, such as those of fault slip rate (Ito *et al.* 2012), fault coupling (Hanifa *et al.* 2014, Ohkura *et al.* 2015), coseismic and postseismic deformation (Gusman *et al.* 2015, Ito *et al.* 2016, Gunawan *et al.* 2016a, Alif *et al.* 2016), and aseismic slip after earthquake occurrences (Ardika *et al.* 2015, Raharja *et al.* 2016), inferring rheological models (Gunawan *et al.* 2014) and use of deformation for seismic hazard analysis (Meilano *et al.* 2015), have shown that GPS can be used comprehensively for crustal deformation analysis.

We utilize and compile all available GPS data sets associated with the coseismic deformation of the 2012 IOE, which have been previously reported in studies of Andaman-Nicobar (Yadav *et al.* 2013) and Sumatra (Feng *et al.* 2015). Also, we report additional GPS data from stations located in Java and along Sumatra that have not been reported so far.

Using these complete GPS sets, this study investigates the response of various available finite fault models of the 2012 IOE, namely by Hayes (http://earthquake.usgs.gov/earthquakes/eqinthenews/2012/usc000905e/finite_fault.php; accessed on December 2015), hereinafter referred to as the H12 model, by Yue *et al.* (2012), hereinafter referred to as the Y12 model; by Wei *et al.* (2013), hereinafter referred to as the W13 model; and by Hill *et al.* (2015), hereinafter referred to as the H15 model. We investigate how well these models fit to the observed coseismic offsets as derived from GPS data. Investigation of the best coseismic slip model is very crucial, especially in analyzing postseismic deformation after earthquake occurrences. In addition, understanding the distribution of coseismic slip during earthquake occurrences will greatly improve further hazard analysis.

2. COSEISMIC SLIP MODEL

Similar to other earthquakes of magnitude 8 or more, the 2012 M_w8.6 IOE has attracted investigation of crustal deformation associated with this intraplate earthquake. In order to invert the coseismic slip of the 2012 IOE, previous studies have used geodetic and/or seismic data. The H12 model used 38 teleseismic broadband *P* waveforms, 13 broadband *SH* waveforms, and 56 long-period surface waves to analyze the coseismic slip of the 2012 IOE. It was inferred that multiple faults were involved in the rupture. This suggested that a NNE-SSW fault plane with a strike of 199° and a WNW-ESE fault plane with a strike of 108° ruptured during this earthquake, with a maximum slip of 70 m (Fig. 2).

The Y12 model analyzed the 2012 IOE using short-period body waves and long-period surface waves recorded at teleseismic stations; it was suggested that rupture involved a very complex network of faults that had never

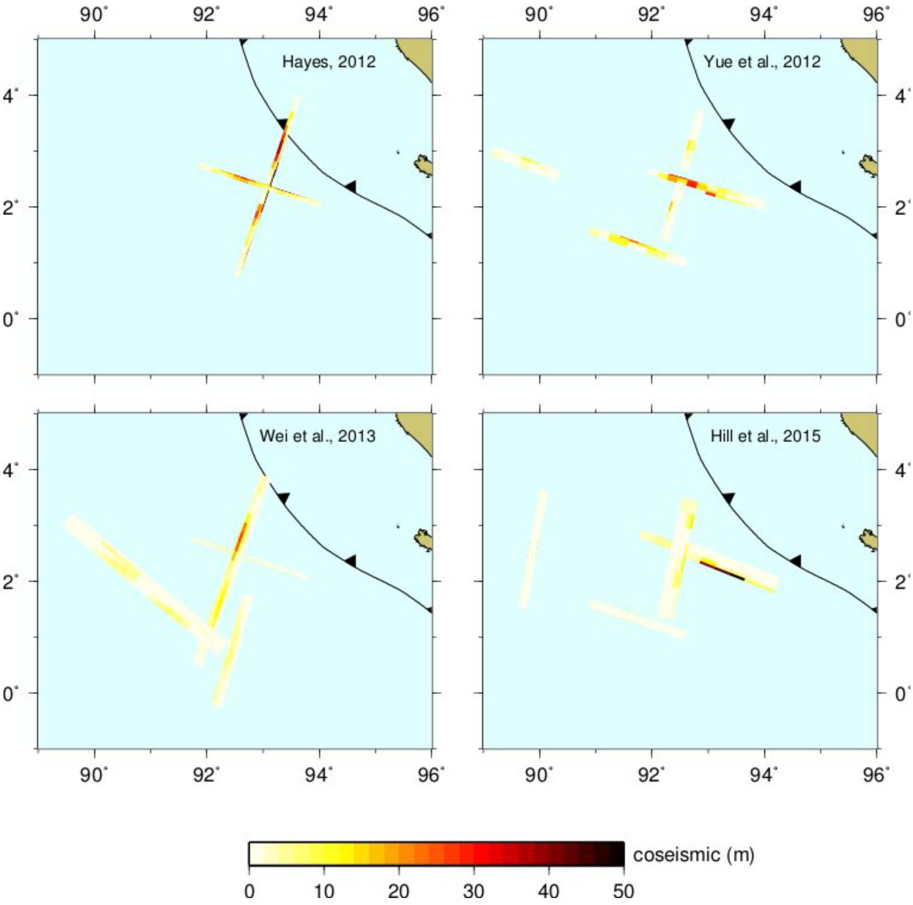


Fig. 2. Coseismic slip models of the 2012 IOE used in this study based on the results of Hayes (USGS), Yue *et al.* (2012), Wei *et al.* (2013), and Hill *et al.* (2015).

been documented or mapped before. They showed that the first fault, with WNW-ESE orientation, ruptured with seismic moment of ~M_w8.5, triggering the cross-cutting, orthogonal second fault in the NNE-SSW direction with seismic moment of ~M_w7.9. Then, the third fault in the direction of WNW-ESE ruptured with seismic moment of ~M_w8.3, at about 150 km southwest of the first fault. Finally, the fourth fault in the WNW-ESE direction, at about 330 km from the epicenter, ruptured with seismic moment of ~M_w7.8. The total rupture at these faults corresponds to a seismic moment of ~M_w8.7 (Fig. 2).

The W13 model analyzed coseismic slip of the 2012 IOE earthquakes using joint inversion of regional and teleseismic waveform data, showing that the rupture started on an ESE-WNW fault with a strike of 289°, which produced a maximum slip of 12 m. Then, after 10 s, rupture continued on a 20° strike fault with maximum slip of 24 m. The third rupture occurred 70 s later, with a maximum slip of 8 m on a 310° strike fault. They showed that two hours after the mainshock, a M_w8.2 earthquake occurred along a 16° strike fault with a maximum slip of 6.5 m (Fig. 2).

The H15 model used joint inversion of high-rate GPS data, far-field static GPS displacements, teleseismic observations and source time functions derived from broadband surface waves to analyze the 2012 IOE. They showed that most of the moment release was along the WNW-ESE right-lateral fault plane with a maximum slip of 48 m. Their result suggests that multiple faults ruptured during the event, including a NNE-SSW fault at 400 km west of the hypocenter (Fig. 2).

Because these studies used different methodologies and data, their results yield very different coseismic slip models. For example, the coseismic slip models of H12 and W13 considered the main rupture as occurring on a NNE-SSW fault, while the Y12 and H15 models considered the WNW-ESE fault to be the main rupture. Maximum slips obtained by these studies also differ, with the H12 coseismic fault model producing the highest maximum slip, followed by the H15, Y12, and W13 models.

3. GPS DATA AND METHOD

This study utilizes GPS data from stations surrounding the epicenter of the 2012 IOE, located in the Andaman-Nicobar region, Sumatra, and Java (Fig. 3). Some of these GPS data are obtained from previous studies, namely those reported by Yadav *et al.* (2013) and Feng *et al.* (2015). In addition to these available GPS data, we also use new GPS data sets from stations located in Sumatra and Java as part of the Indonesian Continuously Operating Reference Stations (InaCORS). Table 1 summarizes the GPS sites used in this study.

Yadav *et al.* (2013) reported the coseismic offsets of the 2012 IOE observed by the Andaman-Nicobar Permanent GPS Network (ANPGN). They showed that GPS sites in the Andaman Islands, which are located to the north of 2012 IOE, experienced southward movement with coseismic displacements up to 3 cm. The GPS site located at Nicobar Island experienced coseismic displacement of 4 cm in the southeast direction (Fig. 4).

GPS sites along Sumatra Island also experienced coseismic displacements during the 2012 IOE. Feng *et al.* (2015) pointed out that the SuGAr (Sumatran GPS Array) recorded significant offsets from the 2012 IOE, showing that Simeuleu Island experienced the largest coseismic offsets, up to 28 cm. The GPS network in northern Sumatra, named the Aceh GPS Net-

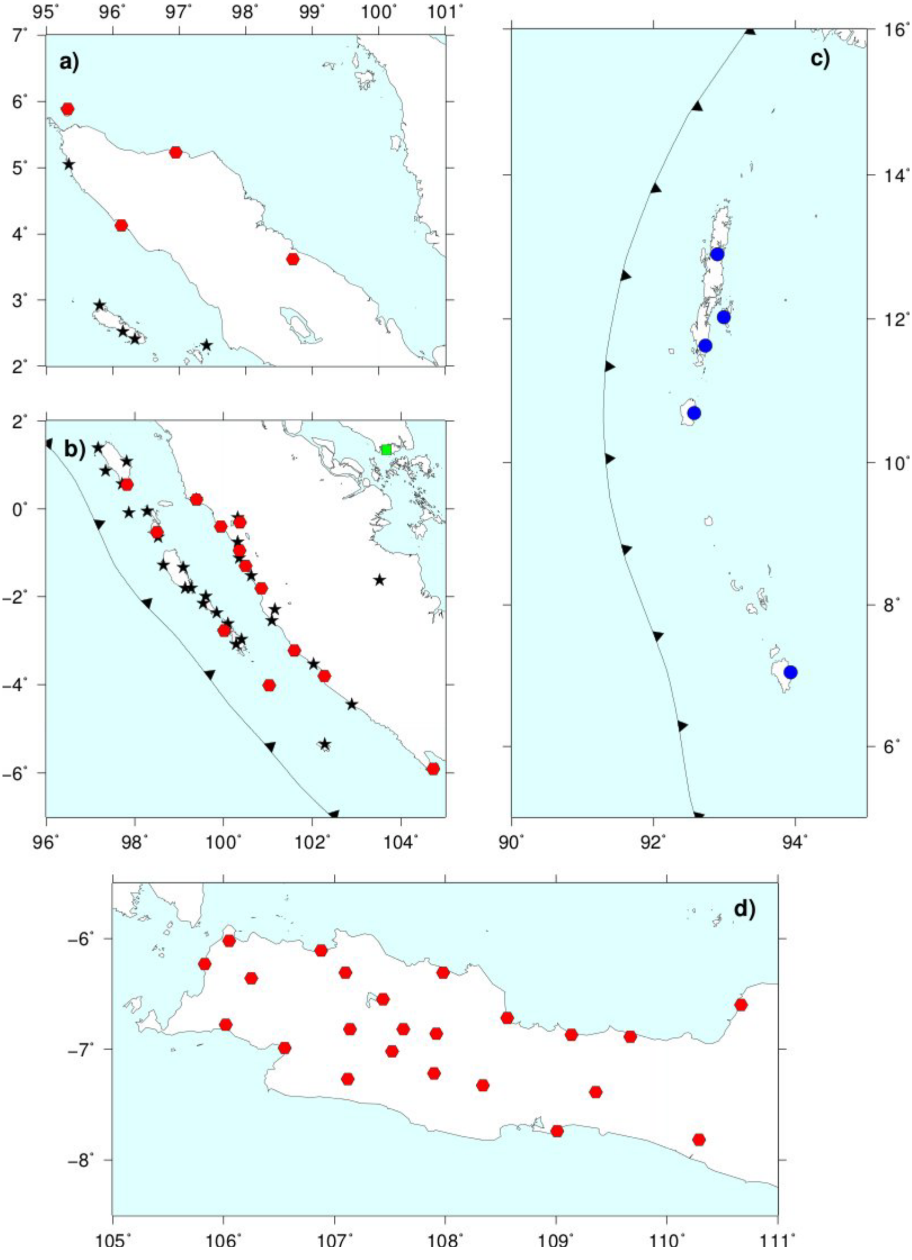


Fig. 3. Locations of GPS sites used in this study. Blue circles indicate GPS sites at Andaman-Nicobar from Yadav *et al.* (2013), black stars denote GPS sites of the SuGAr (Feng *et al.* 2015) and red hexagons and green squares represent GPS sites of BIG and IGS, respectively.

Table 1

Information about the GPS sites used in this study

Site	Long. [°]	Lat. [°]	Network	Site	Long. [°]	Lat. [°]	Network
ABGS	99.39	0.22	SuGAr[1]	MLBU	100.85	–1.81	InaCORS[2]
BITI	97.81	1.08	SuGAr[1]	SAMP	98.71	3.62	InaCORS[2]
BNON	96.15	2.52	SuGAr[1]	SEBL	101.60	–3.22	InaCORS[2]
BSAT	100.28	–3.08	SuGAr[1]	SLBI	100.01	–2.77	InaCORS[2]
BSIM	96.33	2.41	SuGAr[1]	TDAL	97.82	0.55	InaCORS[2]
BTET	98.64	–1.28	SuGAr[1]	TIKU	99.94	–0.40	InaCORS[2]
BTHL	97.71	0.57	SuGAr[1]	TLOK	100.50	–1.30	InaCORS[2]
BUKT	100.32	–0.20	SuGAr[1]	TNBL	98.50	–0.53	InaCORS[2]
HNKO	97.34	0.87	SuGAr[1]	CBAY	93.93	7.05	ANPGN[3]
JMBI	103.52	–1.62	SuGAr[1]	HBAY	92.57	10.69	ANPGN[3]
KTET	99.84	–2.36	SuGAr[1]	PORT	92.73	11.63	ANPGN[3]
LAIS	102.03	–3.53	SuGAr[1]	HAVE	92.99	12.03	ANPGN[3]
LEWK	95.80	2.92	SuGAr[1]	MBDR	92.90	12.90	ANPGN[3]
LHW2	97.17	1.39	SuGAr[1]	PALK	80.70	7.27	IGS[3]
LNNG	101.16	–2.29	SuGAr[1]	COCO	96.83	12.18	IGS[3]
MKMK	101.09	–2.54	SuGAr[1]	BAKO	106.84	6.49	IGS[3]
MLKN	102.28	–5.35	SuGAr[1]	XMIS	105.68	10.45	IGS[3]
MNNA	102.89	–4.45	SuGAr[1]	CUSV	100.53	13.74	IGS[3]
MSAI	99.09	–1.33	SuGAr[1]	DGAR	72.37	7.26	IGS[3]
NGNG	99.27	–1.80	SuGAr[1]	CANG	107.52	–7.02	InaCORS[2]
NTUS	103.67	1.346	IGS[1]	CBTU	107.10	–6.31	InaCORS[2]
PARY	100.32	–0.75	SuGAr[1]	CCIR	108.56	–6.72	InaCORS[2]
PBJO	98.52	–0.64	SuGAr[1]	CCLP	109.01	–7.74	InaCORS[2]
PBLI	97.41	2.31	SuGAr[1]	CGON	106.05	–6.02	InaCORS[2]
PKRT	99.54	–2.15	SuGAr[1]	CJKT	106.88	–6.11	InaCORS[2]
PPNJ	99.60	–1.99	SuGAr[1]	CJPR	110.67	–6.60	InaCORS[2]
PRKB	100.40	–2.97	SuGAr[1]	CJUR	107.14	–6.82	InaCORS[2]
PSKI	100.35	–1.12	SuGAr[1]	CLBG	107.62	–6.82	InaCORS[2]
PSMK	97.86	–0.09	SuGAr[1]	CMIS	108.34	–7.33	InaCORS[2]
PTLO	98.28	–0.05	SuGAr[1]	CMLP	106.02	–6.78	InaCORS[2]
SLBU	100.01	–2.77	SuGAr[1]	CPBL	109.36	–7.39	InaCORS[2]
SMGY	100.10	–2.61	SuGAr[1]	CPKL	109.67	–6.89	InaCORS[2]
TLLU	99.13	–1.80	SuGAr[1]	CPSR	105.83	–6.23	InaCORS[2]

to be continued

Table 1 (continuation)

Site	Long. [°]	Lat. [°]	Network	Site	Long. [°]	Lat. [°]	Network
TRTK	100.62	–1.52	SuGAr[1]	CPTU	106.55	–6.99	InaCORS[2]
UMLH	95.34	5.05	SuGAr[1]	CPWK	107.44	–6.55	InaCORS[2]
CAIR	99.39	0.21	InaCORS[2]	CRKS	106.25	–6.36	InaCORS[2]
CBKL	102.27	–3.80	InaCORS[2]	CROL	107.98	–6.31	InaCORS[2]
CBKT	100.37	–0.31	InaCORS[2]	CRUT	107.90	–7.22	InaCORS[2]
CPDG	100.36	–0.95	InaCORS[2]	CSUM	107.92	–6.86	InaCORS[2]
CSAB	95.32	5.89	InaCORS[2]	CTAN	107.12	–7.27	InaCORS[2]
CTCN	104.73	–5.91	InaCORS[2]	CTGL	109.14	–6.87	InaCORS[2]
LHMI	96.95	5.23	InaCORS[2]	JOGS	110.29	–7.82	InaCORS[2]
MEGO	101.03	–4.01	InaCORS[2]	HYDE	17.42	78.55	IGS[2]
MEUL	96.13	4.13	InaCORS[2]	IISC	13.02	77.57	IGS[2]

[1]Feng *et al.* (2015)
[2]This study
[3]Yadav *et al.* (2013)

work for the Sumatran fault System (AGNeSS; Tabei *et al.* 2015), was also affected by the 2012 IOE (Anugrah *et al.* 2015; Fig. 4).

In addition to the published data sets described above, we use the GPS network in Java and Sumatra installed and maintained by the Geospatial Information Agency of Indonesia (BIG). To analyze the daily coordinate solutions from these GPS sites, we use the GAMIT/GLOBK software package (Herring *et al.* 2010a, b). We process these GPS data together with several International GNSS Service (IGS) sites – DGAR, HYDE, IISC, COCO, XMIS, CUSV, KUNM, PIMO, GUAM, DARW, TOW2, KARR, TNML, and YAR2 – and tied the solutions into the International Terrestrial Reference Frame (ITRF) 2008 (Altamimi *et al.* 2011).

Our analysis indicates that GPS sites in Java experienced coseismic displacements up to 8 mm in the northwest direction (Fig. 4). The standard deviation obtained in our solution for each GPS site was about 4 mm. Table 2 summarizes the coseismic displacements from the 2012 IOE at the GPS sites used in this study.

Using the available coseismic slip models – H12, Y12, W13, and H15 – we calculate the modeled surface coseismic displacements based on a spherical, layered model of the Earth (Pollitz 1996) with elastic thickness of 65 km and Maxwell viscosity of 8.0×10^{18} Pa·s (Gunawan *et al.* 2014). Modeled surface displacements calculated using the coseismic slip models H12, Y12, W13, and H15 are compared to the observed displacements from

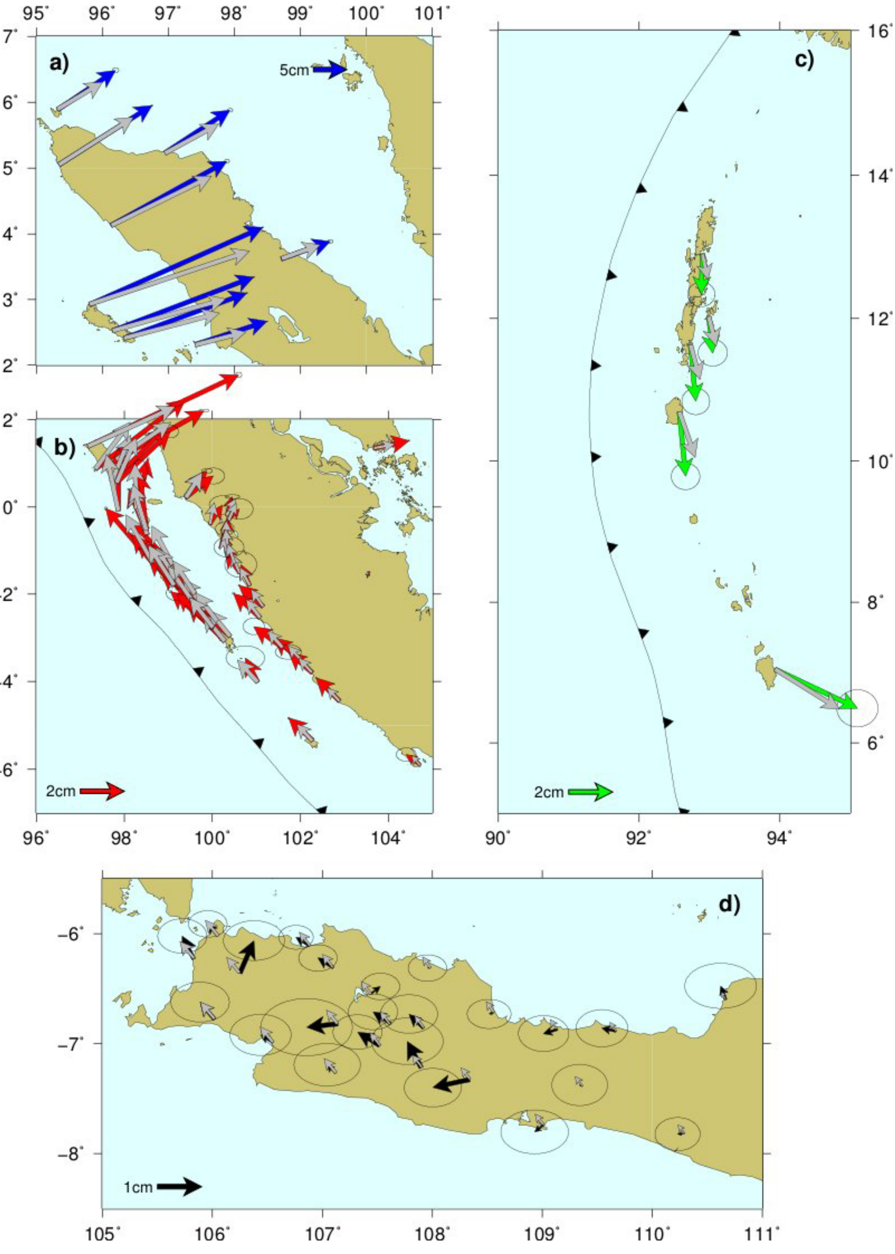


Fig. 4. Modeled displacements (gray arrows) calculated using the coseismic slip model of Hill *et al.* (2015) and observed coseismic displacements of the 2012 IOE at (a) northern Sumatra, indicated by blue arrows; (b) southern Sumatra, indicated by red arrows; (c) Andaman-Nicobar, indicated by green arrows; and (d) Java, indicated by black arrows.

Table 2

Coseismic displacements of the 2012 IOE at GPS sites

Site	E [mm]	N [mm]	σ_E	σ_N	Site	E [mm]	N [mm]	σ_E	σ_N
ABGS	9.8	12.0	1.6	0.7	MLBU	–4.1	9.7	4.8	3.9
BITI	40.0	22.3	1.5	0.4	SAMP	74.4	26.5	2.5	1.9
BNON	214.4	81.6	0.8	0.5	SEBL	–13.2	9.6	3.2	2.5
BSAT	–16.9	17.1	0.4	0.3	SLBI	–16.4	15.8	3.2	2.3
BSIM	185.6	68.4	0.9	0.5	TDAL	24.4	23.3	2.5	1.9
BTET	–21.3	24.7	0.4	0.4	TIKU	5.7	8.8	3.9	2.8
BTHL	24.7	24.8	0.5	0.3	TLOK	–2.5	7.5	4.5	3.1
BUKT	2.7	8.6	2.2	0.6	TNBL	–3.0	26.3	2.2	3.8
HNKO	40.9	31.5	0.6	0.5	CBAY	37.5	–17.8	6.2	5.6
JMBI	0.9	2.5	0.3	0.2	HBAY	2.9	–29.4	4.3	4.0
KTET	–20.3	17.1	0.4	0.4	PORT	2.4	–25.7	4.2	4.0
LAIS	–12.6	10.6	0.3	0.2	HAVE	1.8	–16.6	4.4	4.0
LEWK	262.8	117.4	1.0	0.5	MBDR	–0.5	–18.3	4.1	3.7
LHW2	68.6	32.5	0.9	0.8	PALK	–0.5	–4.9	4.0	3.6
LNNG	–11.0	9.7	0.3	0.2	COCO	–4.9	16.1	4.2	3.8
MKMK	–12.2	11.1	0.3	0.3	BAKO	–4.2	1.8	4.6	4.0
MLKN	–10.9	10.3	0.3	0.3	XMIS	–4.4	5.1	4.1	3.8
MNNA	–10.6	10.6	0.2	0.2	CUSV	3.4	5.0	4.4	4.0
MSAI	–18.8	20.0	0.3	0.2	DGAR	–10.3	–5.3	4.7	4.1
NGNG	–19.9	20.9	0.4	0.3	CANG	–5.0	3.1	3.7	2.6
NTUS	16.0	3.4	0.2	0.2	CBTU	–3.5	2.2	2.9	2.0
PARY	0.5	10.0	0.4	0.3	CCIR	–1.1	–0.3	3.3	2.2
PBJO	–8.4	20.7	0.4	0.3	CCLP	–2.0	–1.6	5.1	3.3
PBLI	108.5	35.3	0.6	0.4	CGON	–2.3	2.7	2.9	2.0
PKRT	–20.0	19.6	0.3	0.3	CJKT	–3.0	1.9	2.6	1.8
PPNJ	–19.7	19.3	0.3	0.3	CJPR	–1.2	3.1	5.4	3.4
PRKB	–24.1	21.8	0.4	0.4	CJUR	–7.1	–0.8	6.9	4.3
PSKI	–2.6	10.3	0.3	0.3	CLBG	–4.0	2.7	3.7	2.7
PSMK	0.1	28.1	0.3	0.3	CMIS	–8.4	–1.8	4.3	2.9
PTLO	5.1	22.8	0.3	0.3	CMLP	–3.3	3.9	4.4	3.0
SLBU	–19.2	16.5	0.4	0.3	CPBL	–0.5	0.3	4.2	3.1
SMGY	–17.0	17.6	0.4	0.3	CPKL	–3.2	0.7	3.9	2.8
TLLU	–19.3	23.2	0.3	0.3	CPSR	–2.6	5.2	3.7	2.6

to be continued

Table 2 (continuation)

Site	E [mm]	N [mm]	σ_E	σ_N	Site	E [mm]	N [mm]	σ_E	σ_N
TRTK	–6.6	9.5	0.3	0.3	CPTU	–2.8	1.7	4.6	3.1
UMLH	141.1	90.0	0.7	0.8	CPWK	2.2	1.7	2.9	2.0
CAIR	12.6	9.9	3.4	2.4	CRKS	3.2	7.4	4.6	3.0
CBKL	–10.9	9.5	3.9	2.2	CROL	–0.6	–0.1	2.9	2.0
CBKT	4.9	5.2	4.1	3.1	CRUT	–3.2	6.0	5.4	3.5
CPDG	–1.6	8.9	2.8	2.1	CSUM	–3.3	3.2	4.3	2.9
CSAB	87.9	59.8	3.2	2.5	CTAN	–1.9	1.9	5.0	3.2
CTCN	–6.5	4.9	3.0	2.0	CTGL	–3.2	–1.0	3.8	2.7
LHMI	98.5	64.9	2.6	1.9	JOGS	–1.5	–0.1	3.4	2.5
MEGO	–5.6	11.0	5.8	3.6	HYDE	2.1	–7.8	2.8	2.3
MEUL	175.6	97.3	2.5	1.9	IISC	3.3	–4.8	2.1	1.6

Table 3

Misfit of data displacements to model

Coseismic slip source	Data used for inversion	RMS
Hayes (USGS)	Seismic	28.10
Yue *et al.* (2012)	Seismic	20.51
Wei *et al.* (2013)	Seismic	17.47
Hill *et al.* (2015)	GPS + seismic	12.33

the GPS sites. We calculate the root mean square (RMS) between modeled and observed displacements to find the optimum coseismic model (Table 3).

4. DISCUSSION

Our study employs a spherical Earth model to calculate surface displacements at each GPS site. Other studies applied an elastic half-space model (Okada 1992), because it is much faster in terms of required calculation time (Piersanti *et al.* 1997). Unfortunately, the half-space model has limitations in areas far from the source, as the sphericity of the Earth starts playing role. To check the impact of using a half-space model for the far field region in Java, we calculate the H15 model using an elastic half-space model, yielding RMS of 22.17, a much higher misfit than that resulting from using spherical Earth analysis (RMS = 12.33). This indicates that using a spherical Earth model to analyze far field deformation is much more appropriate than using a half-space model.

This study suggests that the H15 model generates smaller misfit than the other models (Table 3), most likely because the H15 model uses joint inversion of seismic and geodetic data, while the other considered models use seismic data only. In another earthquake case, such as that of the 2004 Sumatra-Andaman earthquake, Poisson *et al.* (2011) showed that a coseismic slip model derived from geodetic data better explains the recorded tsunami observed by satellite altimetry than a coseismic slip model derived from geological observation, tide gauge records, or other data.

In another example, considering the 2006 Java tsunami earthquake, Gunawan *et al.* (2016b) suggested that the coseismic fault slip model obtained from tsunami data better agrees with the postseismic deformation observed by GPS than with the other models obtained from seismic data. Although only one GPS site (BAKO), located at a remote distance, was active during the 2006 Java tsunami earthquake, modeled coseismic displacements from data BAKO detected during the 2006 event were about 2.5 mm, which fits to the tsunami data better than the other models.

To summarize our investigation of the 2012 Indian Ocean earthquake, we showed that analysis of coseismic slip during earthquake occurrences obtained from a single type of observation, namely seismic data, poorly fits the GPS measurements. Similar results have been found in various other studies which only use seismic data with coseismic slip models, such as the study of the 2004 Sumatra-Andaman earthquake by Poisson *et al.* (2011). Instead, using multiple observations, such as seismic, geodetic, or tsunami data, as also shown by Hill *et al.* (2015) and Rhie *et al.* (2007), generates better results than using a single type of observation.

5. CONCLUSIONS

We analyzed the coseismic slip of the 2012 Indian Ocean earthquake based on the H12, Y12, W13, and H15 models. We used coseismic offsets obtained from GPS measurements in Andaman-Nicobar, Sumatra and Java, comparing these offsets to the modeled coseismic displacements from the available slip models using a spherical, layered model of the Earth. We find that Java Island experienced coseismic deformation of up to 8 mm, as observed by our GPS network. Our results suggest that a coseismic slip model obtained from multiple observations of seismic and geodetic data – the H15 model – produced less misfit than the coseismic slip models obtained using a single type of observation (from a seismic network).

Acknowledgments. We thank Vineet K. Gahalaut, two anonymous reviewers, and the Editor for the constructive suggestions, which helped to improve the quality of this manuscript. This research was partially funded by the Australian Department of Foreign Affairs and Trade (DFAT) for Graduate Research on Earthquake and Active Tectonics at the Bandung Institute of Technology and the World Class University Research Fund from the Bandung Institute of Technology. We thank Fred Pollitz for making STATIC1D code freely available. Figures were generated using the Generic Mapping Tool software (Wessel and Smith 1998).

References

Alif, S.M., I. Meilano, E. Gunawan, and J. Efendi (2016), Evidence of postseismic deformation signal of the 2007 M8.5 Bengkulu earthquake and the 2012 M8.6 Indian Ocean earthquake in Southern Sumatra, Indonesia, based on GPS data, *J. Appl. Geod.* **10**, 2, 103-108, DOI: 10.1515/jag-2015-0019.

Altamimi, Z., X. Collilieux, and L. Métivier (2011), ITRF2008: an improved solution of the international terrestrial reference frame, *J. Geod.* **85**, 8, 457-473, DOI: 10.1007/s00190-011-0444-4.

Anugrah, B., I. Meilano, E. Gunawan, and J. Efendi (2015), Estimation of postseismic deformation parameters from continuous GPS data in northern Sumatra after the 2004 Sumatra–Andaman earthquake, *Earthq. Sci.* **28**, 5-6, 347-352, DOI: 10.1007/s11589-015-0136-x.

Ardika, M., I. Meilano, and E. Gunawan (2015), Postseismic deformation parameters of the 2010 M7.8 Mentawai, Indonesia, earthquake inferred from continuous GPS observations, *Asian. J. Earth Sci.* **8**, 4, 127-133, DOI: 10.3923/ajes.2015.127.133.

Delescluse, M., N. Chamot-Rooke, R. Cattin, L. Fleitout, O. Trubienko, and C. Vigny (2012), April 2012 intra-oceanic seismicity off Sumatra boosted by the Banda-Aceh megathrust, *Nature* **490**, 7419, 240-244, DOI: 10.1038/nature11520.

Duputel, Z., H. Kanamori, V.C. Tsai, L. Rivera, L. Meng, J.P. Ampuero, and J.M. Stock (2012), The 2012 Sumatra great earthquake sequence, *Earth Planet. Sci. Lett.* **351**, 247-257, DOI: 10.1016/j.epsl.2012.07.017.

Feng, L., E.M. Hill, P. Banerjee, I. Hermawan, L.L. Tsang, D.H. Natawidjaja, B.W. Suwargadi, and K Sieh (2015), A unified GPS-based earthquake catalog for the Sumatran plate boundary between 2002 and 2013, *J. Geophys. Res. Solid Earth* **120**, 5, 3566-3598, DOI: 10.1002/2014JB011661.

Gunawan, E., T. Sagiya, T. Ito, F. Kimata, T. Tabei, Y. Ohta, I. Meilano, H.Z. Abidin, I. Agustan, I. Nurdin, and D. Sugiyanto (2014), A comprehensive model of postseismic deformation of the 2004 Sumatra-Andaman earthquake deduced from GPS observations in northern Sumatra, *J. Asian Earth Sci.* **88**, 218-229, DOI: 10.1016/j.jseaes.2014.03.016.

Gunawan, E., M. Kholil, and I. Meilano (2016a), Splay-fault rupture during the 2014 Mw7. 1 Molucca Sea, Indonesia, earthquake determined from GPS measurements, *Phys. Earth Planet. Int.* **259**, 29-33, DOI: 10.1016/j.pepi.2016.08.009.

Gunawan, E., I. Meilano, H.Z. Abidin, N.R. Hanifa, and Susilo (2016b), Investigation of the best coseismic fault model of the 2006 Java tsunami earthquake based on mechanisms of postseismic deformation, *J. Asian Earth Sci.* **117**, 64-72, DOI: 10.1016/j.jseaes.2015.12.003.

Gusman, A.R., S. Murotani, K. Satake, M. Heidarzadeh, E. Gunawan, S. Watada, and B. Schurr (2015), Fault slip distribution of the 2014 Iquique, Chile, earthquake estimated from ocean-wide tsunami waveforms and GPS data, *Geophys. Res. Lett.* **42**, 4, 1053-1060, DOI: 10.1002/2014GL062604.

Hanifa, N.R., T. Sagiya, F. Kimata, J. Efendi, H.Z. Abidin, and I. Meilano (2014), Interplate coupling model off the southwestern coast of Java, Indonesia, based on continuous GPS data in 2008-2010, *Earth Planet. Sci. Lett.* **401**, 159-171, DOI: 10.1016/j.epsl.2014.06.010.

Herring, T.A., R.W. King, and S.C. McClusky (2010a), GAMIT reference manual release 10.4, Report, Massachusetts Institute Technology, Cambridge, USA, 171 pp.

Herring, T.A., R.W. King, and S.C. McClusky (2010b), GLOBK reference manual: global Kalman filter VLBI and GPS analysis program release 10.4, Report, Massachusetts Institute Technology, Cambridge, USA, 95 pp.

Hill, E.M., H. Yue, S. Barbot, T. Lay, P. Tapponnier, I. Hermawan, J. Hubbard, P. Banerjee, L. Feng, D. Natawidjaja, and K. Sieh (2015), The 2012 Mw 8.6 Wharton Basin sequence: A cascade of great earthquakes generated by near-orthogonal, young, oceanic mantle faults, *J. Geophys. Res. Solid Earth* **120**, 5, 3723-3747, DOI: 10.1002/2014JB011703.

Ito, T., E. Gunawan, F. Kimata, T. Tabei, M. Simons, I. Meilano, I. Agustan, Y. Ohta, I. Nurdin, and D. Sugiyanto (2012), Isolating along-strike variations in the depth extent of shallow creep and fault locking on the northern Great Sumatran Fault, *J. Geophys. Res. Solid Earth* **117**, B6, B06409, DOI: 10.1029/2011JB008940.

Ito, T., E. Gunawan, F. Kimata, T. Tabei, I. Meilano, I. Agustan, Y. Ohta, N. Ismail, I. Nurdin, and D. Sugiyanto (2016), Co-seismic offsets due to two earthquakes (Mw6.1) along the Sumatran fault system derived from GNSS measurements, *Earth Planet. Space* **68**, 57, DOI: 10.1186/s40623-016-0427-z.

Meilano, I., E. Gunawan, D. Sarsito, K. Prijatna, H.Z. Abidin, and J. Efendi (2015), Preliminary deformation model for National Seismic Hazard map of Indonesia, *AIP Conf. Proc.* **1658**, 030003, DOI: 10.1063/1.4915011.

Meng, L., J.P. Ampuero, J. Stock, Z. Duputel, Y. Luo, and V.C. Tsai (2012), Earthquake in a maze: Compressional rupture branching during the 2012 Mw 8.6 Sumatra earthquake, *Science* **337**, 6095, 724-726, DOI: 10.1126/science.1224030.

Ohkura, T., T. Tabei, F. Kimata, T.C. Bacolcol, Y. Nakamura, A.C. Luis, A. Pelicano, R. Jorgio, M. Tabique, M. Abrahan, E. Jorgio, and E. Gunawan (2015), Plate convergence and block motions in Mindanao Island, Philippine as derived from campaign GPS observations, *J. Disaster Res.* **10**, 1, 59-66, DOI: 10.20965/jdr.2015.p0059.

Okada, Y. (1992), Internal deformation due to shear and tensile faults in a half-space, *Bull. Seismol. Soc. Am.* **82**, 2, 1018-1040.

Piersanti, A., G. Spada, and R. Sabadini (1997), Global postseismic rebound of a viscoelastic Earth: Theory for finite faults and application to the 1964 Alaska earthquake, *J. Geophys. Res.* **102**, B1, 477-492, DOI: 10.1029/96JB01909.

Poisson, B., C. Oliveros, and R. Pedreros (2011), Is there a best source model of the Sumatra 2004 earthquake for simulating the consecutive tsunami?, *Geophys. J. Int.* **185**, 3, 1365-1378, DOI: 10.1111/j.1365-246X.2011.05009.x.

Pollitz, F.F. (1996), Coseismic deformation from earthquake faulting on a layered spherical Earth, *Geophys. J. Int.* **125**, 1, 1-14, DOI: 10.1111/j.1365-246X.1996.tb06530.x.

Raharja, R., E. Gunawan, I. Meilano, H.Z. Abidin, and J. Efendi (2016), Long aseismic slip duration of the 2006 Java tsunami earthquake based on GPS data, *Earthq. Sci.*, DOI: 10.1007/s11589-016-0167-y.

Rhie, J., D. Dreger, R. Bürgmann, and B. Romanowicz (2007), Slip of the 2004 Sumatra–Andaman earthquake from joint inversion of long-period global seismic waveforms and GPS static offsets, *Bull. Seismol. Soc. Am.* **97**, 1, S115-S127, DOI: 10.1785/0120050620.

Tabei, T., F. Kimata, T. Ito, E. Gunawan, H. Tsutsumi, Y. Ohta, T. Yamashina, Y. Soeda, N. Ismail, I. Nurdin, D. Sugiyanto, and I. Meilano (2015), Geodetic and geomorphic evaluations of earthquake generation potential of the northern Sumatran fault, Indonesia. **In:** *International Association of Geodesy Symposia*, Springer, Berlin Heidelberg, DOI: 10.1007/1345_2015_200, 1-8.

Wei, S., D. Helmberger, and J.P. Avouac (2013), Modeling the 2012 Wharton basin earthquakes off-Sumatra: Complete lithospheric failure, *J. Geophys. Res. Solid Earth* **118**, 7, 3592-3609, DOI: 10.1002/jgrb.50267.

Wessel, P., and W.H.F. Smith (1998), New, improved version of the generic mapping tools released, *Eos Trans. AGU* **79**, 47, 579, DOI: 10.1029/98EO00426.

Yadav, R.K., B. Kundu, K. Gahalaut, J. Catherine, V.K. Gahalaut, A. Ambikapthy, and M.S. Naidu (2013), Coseismic offsets due to the 11 April 2012 Indian Ocean earthquakes (Mw 8.6 and 8.2) derived from GPS measurements, *Geophys. Res. Lett.* **40**, 13, 3389-3393, DOI: 10.1002/grl.50601.

Yue, H., T. Lay, and K.D. Koper (2012), En echelon and orthogonal fault ruptures of the 11 April 2012 great intraplate earthquakes, *Nature* **490**, 7419, 245-249, DOI: 10.1038/nature11492.

Received 18 January 2016
Received in revised form 24 June 2016
Accepted 26 July 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2151-2170
DOI: 10.1515/acgeo-2016-0105

Site Specific Ground Motion Modeling and Seismic Response Analysis for Microzonation of Baku, Azerbaijan

Gulam BABAYEV[1] and Luciano TELESCA[2]

[1]Geology and Geophysics Institute, Azerbaijan National Academy of Sciences, Baku, Azerbaijan

[2]Consiglio Nazionale delle Ricerche, Istituto di Metodologie per l'Analisi Ambientale, Tito (PZ), Italy;
e-mail: luciano.telesca@imaa.cnr.it

Abstract

We investigated ground response for Baku (Azerbaijan) from two earthquakes of magnitude *M*6.3 occurred in Caspian Sea (characterized as a near event) and *M*7.5 in Shamakhi (characterized as a remote extreme event). *S*-wave velocity with the average shear wave velocity over the topmost 30 m of soil is obtained by experimental method from the V_P values measured for the soils. The downtown part of Baku city is characterized by low V_{S30} values (< 250 m/s), related to sand, water-saturated sand, gravel-pebble, and limestone with clay. High surface PGA of 240 gal for the *M*7.5 event and of about 190 gal for the *M*6.3 event, and hence a high ground motion amplification, is observed in the shoreline area, through downtown, in the north-west, and in the east parts of Baku city with soft clays, loamy sands, gravel, sediments.

Key words: Baku, Azerbaijan, peak ground acceleration, site amplification, strong ground motion.

Ownership: Institute of Geophysics, Polish Academy of Sciences;

1. INTRODUCTION

Densely populated cities situated in seismic areas characterized by the occurrence of earthquakes with moderate intensity always raise big issues in terms of seismic hazard and risk, making the earthquake disaster preparedness really challenging for decision makers. This is the reason why seismic microzonation has been becoming a widely used methodology for predicting the seismic hazard of an area in order to mitigate the earthquake disaster and assess the seismic risk. One recent example is given by Shiuly *et al.* (2014), who computed the ground motion amplification scenario of sedimentary deposits for the seismic microzonation of Kolkata Megacity (India), located on the world's largest delta island with very soft soil deposit.

Baku is the capital of Azerbaijan, and is one of the fastest growing cities of the country. It lies in one of the seismic zones of Azerbaijan, Absheron peninsula, which is situated on the NW part of the South Caspian region (Babayev *et al.* 2010). Recently, Telesca *et al.* (2012) performed a detailed study of temporal properties of its historical and instrumental seismic catalogue. Although earthquakes of very low intensity from its own focal zone occurred in the peninsula, a potential seismic hazard can arise from the active focal zones of the Caspian Sea and Shamakhi-Gobustan (Fig. 1). We can identify two sources of earthquake activity in the area: the subduction zone (Jackson *et al.* 2002) and shallow crustal faults.

Baku is situated in the trough representing a fan-shaped depression, as a result of the presence of north-western dislocations which occur from the western and eastern fault systems (Shikhalibeyli 1996, Babayev *et al.* 2010), and the whole Absheron zone with Baku city is mostly compressed with the thrust and reverse faults (Babayev *et al.* 2010). The shallow substrate of the city is mainly represented by deposits of Quaternary age composed by sands, gravel-pebble, and limestone with clay and intercalated layers of sand and rubble and water-saturated sand (Babayev *et al.* 2010). Since soft sediments, responsible for the great amplification of seismic waves, represent a very large fraction of the soil, it is crucial to investigate their site effects on the region (Subhadra *et al.* 2015). Furthermore, Baku city (the investigated area is indicated by the black rectangle in Fig. 2) is subjected to a continuous increase of urbanization that should be taken into account along with the large seismic potential of the region around, struck by rather intense earthquakes during 1842 (*Ms*5), 1902 (*Ms*7.5), 1910 (*Ms*4.9), 1922 (*Ms*5), 1935 (*Ms*3.5), 1937 (*Ms*5), 1946 (*Ms*5.1), 1958 (*Ms*2.9), 1971-1973 (*Ms*5.1-5.6), 1979 (*Ms*4.4), 1983 (*Ms*5), 1992 (*Ms*4.5) (Babayev 2010, Babayev *et al.* 2010, Babayev and Telesca 2014, Gasanov 2003). In this study, for the earthquake scenario we use macroseismic parameters of 1902 Shamakhi earthquake (*Ms*7.5) and those of the South Caspian 2000 earthquake (*Ms*6.3), because

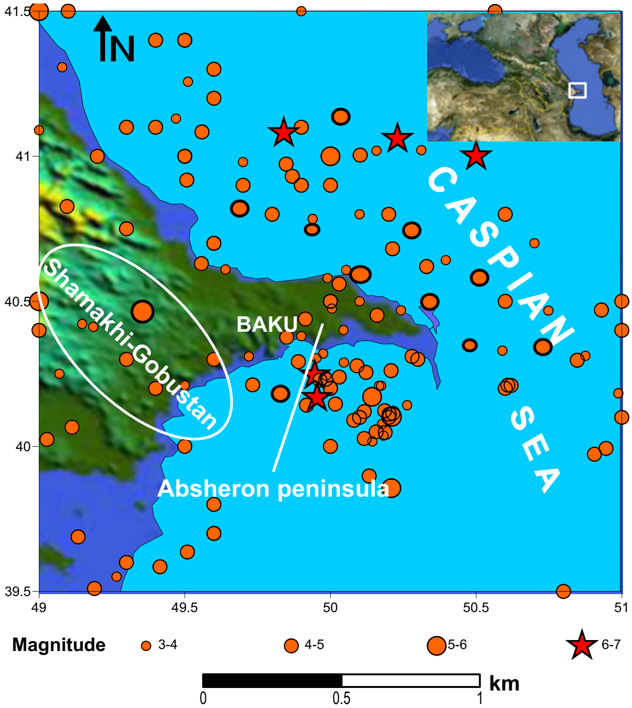


Fig. 1. Absheron peninsula with earthquakes' distribution for the period 1842-2014 (Babayev and Telesca 2014).

Fig. 2. Map of the Absheron peninsula. The investigated area is indicated by the black rectangle.

Absheron peninsula experienced these earthquakes with intensity effects of V-VI (1902 Shamakhi earthquake) and of VI-VII (2000 Caspian earthquake) on its area, and because of the regional fault system. In particular, during the strong 2000 Caspian earthquake 35 people were killed and more than one thousand buildings were severely damaged (Babayev *et al.* 2010). So, there is a high possibility that these earthquakes might re-occur in the future with a similar or higher magnitude level.

The city of Shamakhi (situated about 110 km west of Baku, Fig. 3) was struck by strong seismic events in the past: in 1191, 1667, and 1859 (Veber 1904). On 13 February 1902, a catastrophic earthquake (Shamakhi earthquake) occurred in the region (intensity = $X_{(\text{Rossi-Forel scale})}$) (Boghdanovitch 1904). The event destroyed completely all the buildings located around the epicenter, and several rockfalls, landslides and eruptions of mud volcanoes occurred; also moderate ground shaking was felt in Baku city. Its surface magnitude was estimated as 6.9 ± 0.2 (Kondorskaya and Shebalin 1982); however, considering that the ground shaking lasted for approximately 30-40 s and the ruptured area was about 80 km long, its magnitude could be estimated as 7.5 (Levitski 1902, Boghdanovitch 1904, Babayev *et al.* 2010).

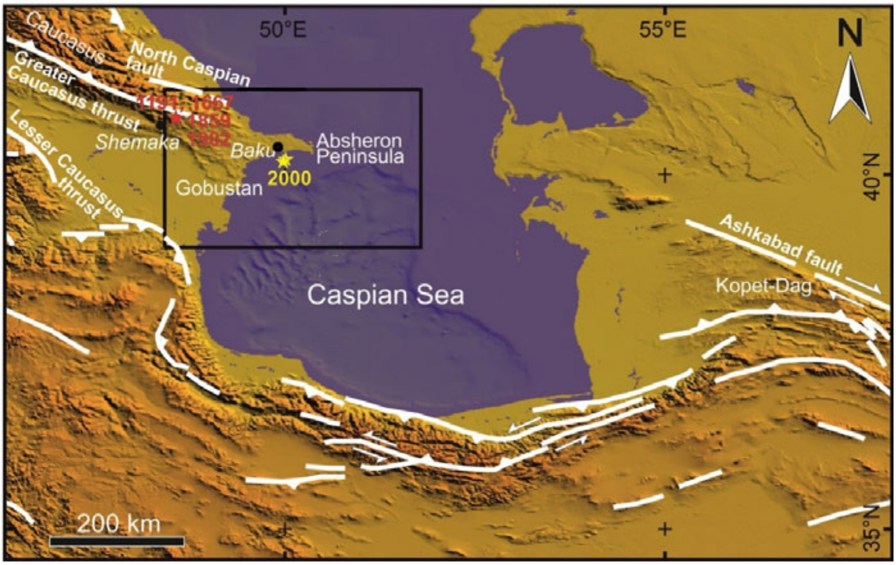


Fig. 3. Topographic map of South Caspian Region. The white lines indicate the main Quaternary faults of the area. The map shows the epicenters of the two seismic events that occurred in 2000 (yellow stars). The red stars indicate the epicenters of other large events that struck Absheron peninsula in 1191, 1667, 1859, and 1902. The marked box indicates the investigated area (tectonics modified after Jackson *et al.* 2002, Babayev *et al.* 2010, 2014).

Thus, even though Baku city, Absheron peninsula, and adjacent Caspian Sea are characterized by moderate seismicity, with intensity of VIII on the MSK-64 scale, the region is considered to be at a high seismic risk due to infrastructures' high vulnerability.

Subsurface ground conditions can be considered as being mainly responsible for the effects of potential earthquakes (Subhadra *et al.* 2015). Secondary wave velocity with the average thickness of 30 m of soil (V_{S30}) is a parameter that significantly influences the amplification factor of the site and the level of ground shaking. The values of the near-surface *S*-wave velocity are generally used to assess seismic hazard of the site. On the basis of many earthquake scenarios it is well known that it is the amplification of ground motion, especially that generated by soft soil layers, that is mainly responsible for the damage on buildings and structures, as it was observed from earthquakes that occurred in Mexico, Japan, USA, and Turkey (Subhadra *et al.* 2015). Therefore, it is highly recommended to estimate the site effects and understand the soil performances during strong shaking, in order to mitigate the level of earthquake disaster. In this research, we calculated the *S*-wave velocity, performed the site response assessment, and calculated the Peak Ground Acceleration (PGA) for scenarios of two earthquakes which were felt in Baku city and we analyzed their distribution through the city both at bedrock and surface ground levels. Furthermore, we used the software SHAKE2000 (Ordonez 2010) to carry out the 1-D response of the site at depths ranging between 20 and 50 m below ground level. We determined the amplification factor and calculated the accelerograms in a subsurface layer.

2. METHODOLOGY

2.1 Modeling parameters for earthquake scenarios

The determination of acceleration of the surface motion in the investigated area was based on the analysis of local site effects, geological settings of the surface, amplification factor, and the seismic response of subsurface soil. In our models, the areal extension of the examined area was 2800×2000, so it was gridded into 28×20 square cells (with side size of 100 m) (Fig. 2).

Since the seismic data are few and the strong earthquakes are generally associated with local irregularities, we employed a scenario-based deterministic approach (Babayev and Telesca 2014). For this strong motion simulation, we select near-field and remote (extreme) target earthquake scenarios. We selected the target earthquakes taking into account their distance from Baku, their magnitude, their effects on the investigated area, the event location in relation to the regional fault system, and their re-occurrence within a certain time interval (Babayev *et al.* 2010). We consider the Caspian earthquake occurred in 2000 (with magnitude 6.3 and epicenter at about 35 km from Baku) as a near event (Fig. 3, indicated by yellow star), and a hypothetic event of *M*7.5 occurred in Shamakhi (at about 110 km from Baku) as an extreme (remote) event (Fig. 3, indicated by red stars). The disaster caused by the recent *M*7 Haiti earthquake that occurred on 12 January 2010, is a clear example of the need to consider historical extreme events in this type of studies. In fact, the probabilistic seismic hazard model that was built for Haiti's area did not consider historical large seismic events; therefore, the surface PGA was absolutely underestimated (USGS 2010). In this study we modeled ground motion for each cell of the city in terms of PGA both at bedrock and surface ground levels using the parameters of the abovementioned two target earthquakes. To estimate the expected PGA at bedrock, we used the following relationship (Aptikayev and Kopnichev 1979):

$$\log A = \begin{cases} 0.28M - 0.8\log R + 1.7, & A > 160\,\text{gal} \quad (\text{for near-field events}) \\ 0.28M - 2.3\log R + 0.8, & A < 160\,\text{gal} \quad (\text{for far-field events}) \end{cases} \tag{1}$$

where A is measured in gal and the hypocentral distance R in km. This relationship can be adapted to a wide range of bedrocks from soft to hard, as Baku city is characterized. Comparing Eq. 1 with the following relationship empirically obtained for Japanese intra-plate earthquakes (*e.g.*, Tonouchi and Kaneko 1984)

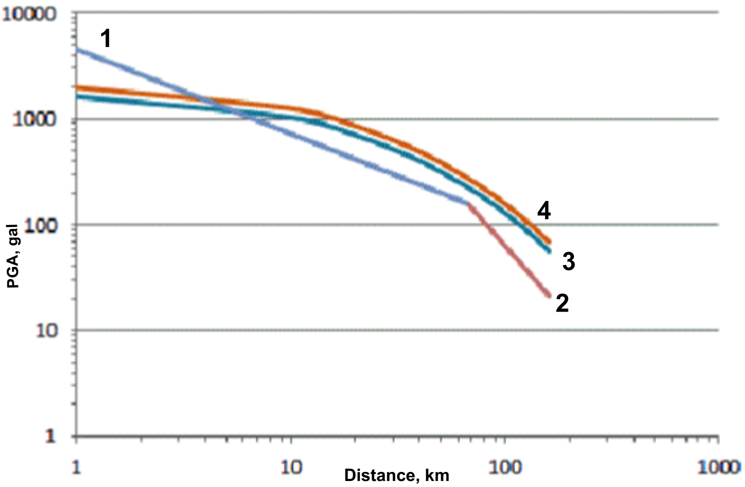


Fig. 4. Comparison of acceleration formula depending on the distance calculated for *M*6.3 (near event) and *M*7.5 (remote event) earthquakes. The curves labeled 1 and 2 indicate the "near" and "remote" earthquakes from Baku, respectively, and are obtained applying relationship 1. The curves labeled as 3 and 4 indicate the Japanese intra-plate earthquakes occurred at the depths of 35 and 55 km, respectively, obtained using Eq. 2.

$$\log A = 0.5M + 0.0043H - \log\left(R + 0.0055 \times 10^{0.5M}\right) - 0.003R + 0.83 \, , \qquad (2)$$

where H is the focal depth, we find a good agreement between the attenuation curves (Fig. 4), thus providing a sound ground for using the relationship 1 (Babayev *et al.* 2010).

2.2 Shear wave velocity distribution for Baku city

Site effects (or the soil seismic response) analysis examines the vertical propagation of shear waves within a one-dimensional soil deposit and calculates the expected response at the surface (Kirtas *et al.* 2016). The soil deposit is considered to consist of homogeneous and isotropic horizontal layers with different elastic properties, laterally extended to infinity (Kirtas *et al.* 2016).

We modeled the subsurface structure down to the seismic bedrock by a horizontally multi-layered structure, with depth-dependent shear-wave velocity, density and thickness (Babayev *et al.* 2010). Using the P-wave velocity values measured in several boreholes, we developed, for each cell of the grid in which Baku city was divided, the subsurface ground model and we identified the type and thickness of sediments, along with the variation of the underlying rock layers. P-wave velocity test was performed in the laboratory (Kuliyev 1986), on-site and on the samples retrieved from the boreholes. The rock samples collected from several stone quarries situated in Absheron peninsula were checked in order to be used as standard testing specimens without macroscopic defects, altered or fractured zones. By means of ultrasonic nondestructive tester that measures the time of propagation of ultrasound pulses, the V_P was measured in the samples.

Figure 5 shows the map of Baku city with the locations of the drilled boreholes, while Fig. 6 shows the calculated accelerograms for the typical subsurface models. The synthesized accelerogram of maximum possible effect on the reference ground in Baku city was calculated on the basis of site effects, geological surface conditions, amplification factor and subsurface soil seismic response (Babayev *et al.* 2010).

The time duration of the strong motion of all models involved is less than 5 s and the whole signal lasts 55 s. The peak value of acceleration of model A1 is around 1 m/s^2. Compared to the other models involved in the current research, the peak value of acceleration of model A1 is smaller. Subsurface thickness of the site of model A1 is not large and soil structure consists of two layers. Although subsurface layers consist of sands and clays (soft sediments), the additional presence of sandstones and tuffs (hard rocks) in the basement does not allow the amplitude to increase and attenuate the waves. Model D4, three layers of which contain strata of soft sediments, has a larger peak value of acceleration, of around 3 m/s^2 of N-S component of

Fig. 5. Map of Baku city; the red dots indicate the sites of drilled boreholes.

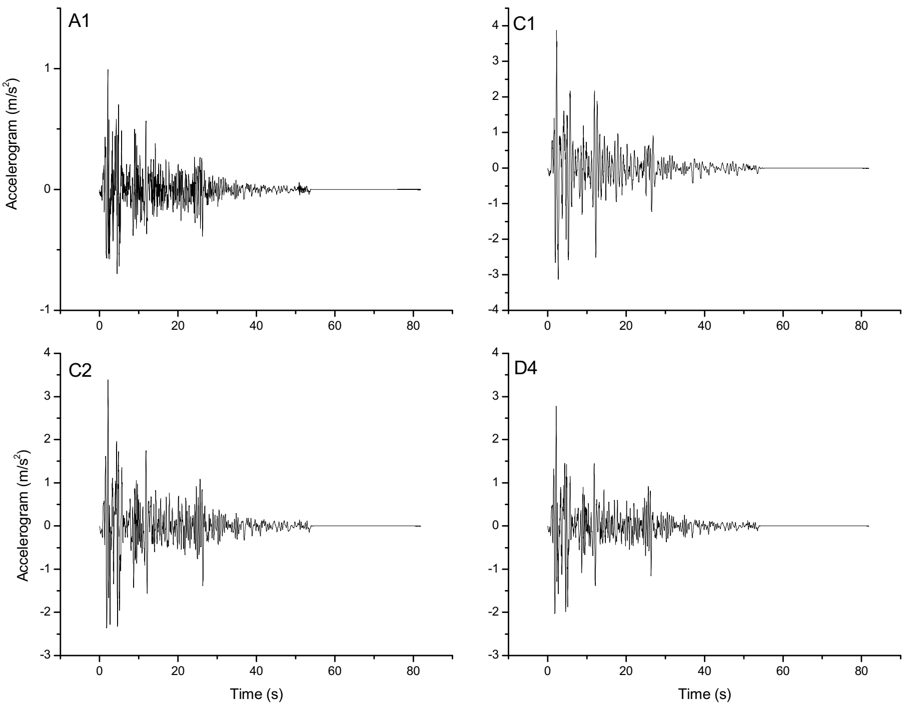


Fig. 6. The calculated accelerograms for typical subsurface models A1, C1, C2, and D4 throughout the area (see Table 1 for the subsurface model description).

Table 1

Subsurface models of soil and sediments for Baku

Model	Thickness of a layer [m]	Lithology
A1	5 1010	sands, clays, sandstones tuff, breccia, shale
C1	4 5 20 3800	sands, gravel-pebble clays and clayey sands clays clays, clayey sandstone, clayey limestone
C2	7 7 23 3200	sands, water-saturated sands clays, pebble, soft-weathered limestone clays organic clays, clayey sandstone
D4	7 20 1390	limestone, sands sands, clays, limestone conglomerate, tuff, sandstone, breccia, shale

synthesized accelerogram. However, the basement of the model consists of hard rock components, such as limestone, conglomerate, tuff, sandstone, and breccia, which also attenuate the wave. The soil structure of models C1 and C2 are assumed to be the four-strata structure, with three subsurface layers with soft sediments. The peak values of acceleration for C1 and C2 are the largest: 4 m/s^2 for C1 and higher than 3 m/s^2 for C2. The strata are inclined to cause a significant amplification of wave amplitude consisting of water-saturated sediments, soft-weathered limestone, organic clay and sands.

Table 1 shows the subsurface models assigned to the model cells. Shear wave velocity V_S [m s^{-1}] averaged over the top 30 m of the soil are estimated from the following empirical relationship

$$V_S = V_P / (4.34 - 0.49 V_P). \tag{3}$$

The V_P value was measured for the specific soils by experimental method (Seed *et al.* 1969). The *S*-wave velocity averaged over the upper 30 m of the soil column and obtained by Eq. 3 was mapped on the city (Fig. 7).

For hard sedimentary rocks, the amplification factor within a layer was calculated by using the following relationship 4 (Midorikawa *et al.* 1992):

$$\log A_{\mathrm{PGA}} = 1.11 - 0.42 \log V_S \ , \tag{4}$$

where A_{PGA} is the amplification factor of PGA between target layer and the layer with $V_S = 440$ m/s. Shear wave velocity (V_S) has the advantage of reflecting the stiffness of soils. Consequently, V_S is an important parameter in

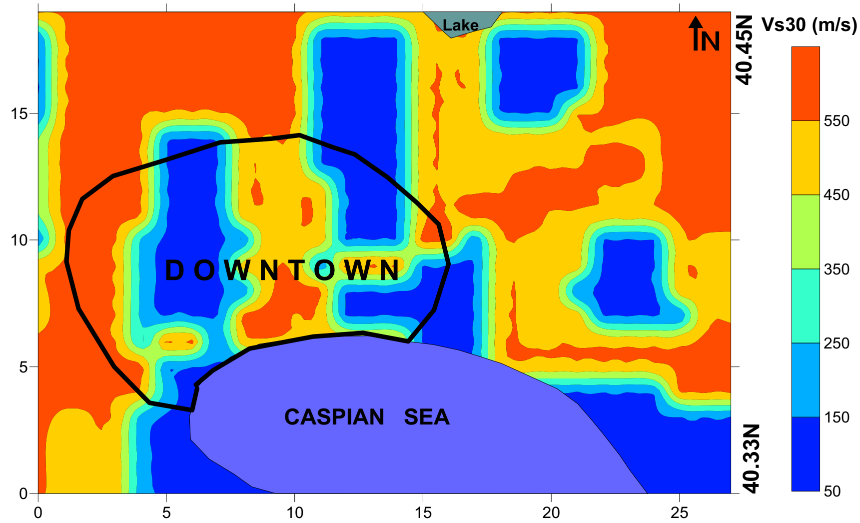


Fig. 7. Spatial distribution of average S-wave velocity V_{S30} [m/s] for Baku estimated by empirical calculations.

subsoil exploration. Shear wave velocity (V_S) by itself is a useful parameter for seismic classification of soils. A widely used seismic soil profile (Dobry *et al.* 2000, Sabetta and Bommer 2002) criteria considers the average V_S in the upper 30 m (V_{S30}). The reference site condition for which the current research is done equals 440 m/s which is represented by dense soil, soft rock (Dobry *et al.* 2000), typical for subsoils in Absheron peninsula. For soft sedimentary rocks and soils, the amplification factor and the relevant accelerogram in a subsurface layer have been determined from shear-wave velocities, density, and thickness of the layer using the SHAKE software (Ordonez 2010). The measured and calculated values of seismic wave velocity and density, and the calculated values of the amplification factor for the principal subsurface units of the model are those used by Babayev *et al.* (2010) (the reader is referred to Table 3 in Babayev *et al.* 2010).

Figure 7 shows the average S-wave velocity (V_{S30}) distribution in the investigated area.

The high V_{S30} values (450-500 m/s) are found in a large part of the city, towards the west-northwest and east-northeast parts of the study area, including some parts of the downtown area. S-wave velocity of >500 m/s is found in 40% of the Baku city area. Some shoreline areas, central part of the city and small spots of the city are featured by comparatively low V_{S30} values (<200 m/s) that could be correlated with eolian-delluvial, water-saturated deposits, soils consisted of clay with inserted layers of sand, pebble and gravel. In general, the decrease of VS30 is observed when the ground motion increases (Subhadra *et al.* 2015).

The determination of the site features for estimating seismic hazard is generally performed on the basis of the near-surface *S*-wave velocity that with the average of over 30 m thickness of soil (V_{S30}) represents one of the site parameters widely utilized for site classification and prediction in terms of the seismic shaking amplification.

2.3 One dimensional (1-D) ground response analysis

The amplification factor for soft soils was obtained by means of SHAKE2000 software (Ordonez 2010), whose basic assumptions are that the soil succession extends infinitely in the horizontal direction and the soil responses are responsible for upward propagation of *S*-waves from the underlying rocks. By means of stress-dependent soil properties, SHAKE2000 performs simulations of the non-linear dynamics of the subsurface soil and sediments, requiring as input of the values of the *S*-wave velocity, density, thickness, shear modulus and damping factor of each layer of the subsurface model. We calculated the 1-D response of a soil column, consisting of a number of horizontal layers, with infinite horizontal extension. The bottom layer is the half-space. For each layer we imposed the condition of homogeneity and isotropy with characteristic of input values. Since the analysis

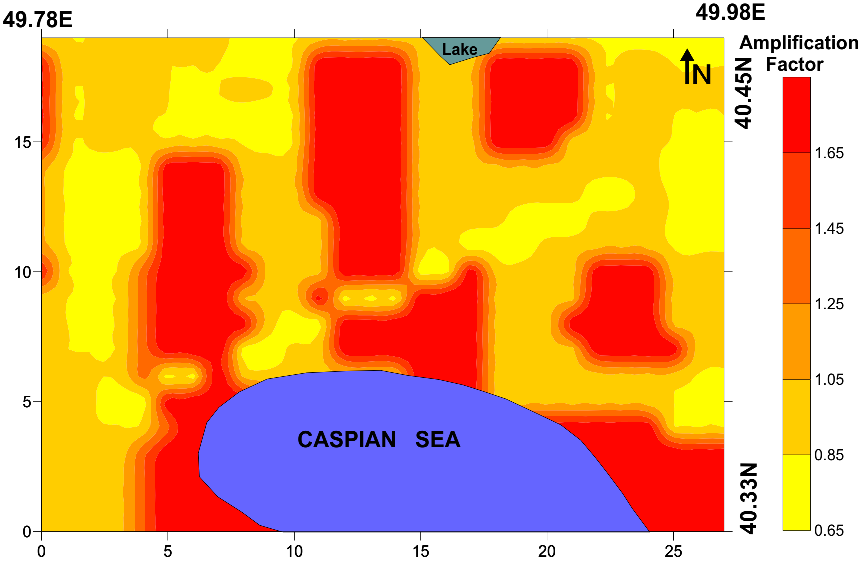


Fig. 8. Map of amplification factor for Baku city for the highest near *M* 6.3 target event.

Table 2

Dynamic soil and rock layer properties

Era	Period		Epoch	Lithology	Charac-teristics	Max. thickness [m]	V_S [m/s]	Density [g/cm3]	G/G_0
Cenozoic	Tertiary	Paleogene	Paleocene-Eocene	Conglomerate	Rock	1500	1.15	2.13	N/A
				Shale	Rock				
				Marlstone	Rock				
			Oligocene	Sandstone	Rock				
		Neogene	Miocene	Sandstone	Rock	1200	0.99	2.04	N/A
				Tuff, brecctia	Rock				
				Shale, tuff, sandstone	Rock				
				Conglomerate	Rock				
			Pliocene	Clayey limestone	Rock	3800	0.59	1.95	N/A
				Clayey sandstone	Rock				
	Quaternary		Pleistocene	Clay, sand-stone, tuff	Rock	2000	0.5	1.89	N/A
			Pleistocene-Holocene	Sand & gravel	Stiff soil	23	0.6	2.1	N/A
				Clayey soil	Medium soil	20	0.25	1.8	C*
				Sandy soil	Stiff soil	7	0.5	2.0	N/A
				Clayey soil	Soft soil	5	0.1	1.6	C
			Holocene	Organic sands	Soft soil	15	0.12	1.7	S*
				Clayey soil	Medium soil	7	0.15	1.7	C
				Sandy soil, gravel	Medium soil	4	0.35	1.9	G*

*) C implies clay, S implies sand, G implies gravel;
G/G_0 – relation of shear modulus ratio *versus* shear strain and the strain dependent damping. Initial shear or low amplitude shear modulus is denoted by G_0.

takes into account the non-linearity of the soils using an iterative procedure, an important role is played by dynamic soil properties. For each soil and rock layer we compiled a table with dynamic properties and the properties have been assigned to each cell, respectively (Table 2).

In this analysis, the necessary input data were the dynamic soil properties and the soil profile. The soil profile is given by the number of layers with the corresponding input values.

This method has been proved to provide well-constrained results and is widely used to perform 1D dynamic soil response analyses (Kirtas *et al.* 2016, Theodoulidis *et al.* 2014). The modification of the soil elastic properties is based on an effective rather than the peak shear strain time-history value (Kirtas *et al.* 2016). The latter refers to a single occurrence with a limited effect on the soil column response during the entire strong ground motion (Theodoulidis *et al.* 2014).

The amplification factor for Baku varies from 0.7 to 1.7 for the highest near target *M*6.3 earthquake (Fig. 8).

3. DISCUSSION

Using all dynamic parameters of subsurface soil, amplification factor and the values of the PGA at bedrock for the two events as seismic input motion parameters of the model, the surface PGA for both target earthquakes was determined (Figs. 9 and 10).

The PGA map for both the bedrock and the ground surface level for the *M*7.5 target earthquake (remote extreme Shamakhi earthquake, 13 February 1902) is shown in Fig. 9. The PGA values vary from 120 gal (0.120 g) to 122 gal (0.122 g) for bedrock level (Fig. 9a) and 80 gal (0.08 g) to 240 gal (0.24 g) for surface level (Fig. 9b).

The PGA values for soft soils derived from SHAKE2000 were used to spatially map the PGA at both the ground surface (Fig. 10a) and the bedrock level (Fig. 10b), with the *M*6.3 earthquake (near event, the Caspian earthquake of 25 November 2000) as the most vulnerable source. The PGA values range from 70 gal (0.07 g) to 120 gal (0.12 g) at the bedrock and 40 gal (0.04 g) to 190 gal (0.19 g) at the surface (Fig. 10). The E-NE sectors of Baku and downtown show PGA values at the surface of about 170-190 gal and at the bedrock of about 120 gal for the *M*6.3 target earthquake, indicating large damage (Fig. 10). The PGA strongly varies at the surface in the eastern section of the city (Fig. 10b).

In Fig. 11, the 5% damped site specific response spectrum at the surface for the *M*6.3 scenario earthquake is plotted. Two peaks, of 0.7 and 0.88 g, are clearly visible in the spectral acceleration, indicating an enhanced amplification for the mesh with clayey sands and sandy clays (areas with C1 typical subsurface model; see Table 1 for description and Fig. 7) with very low V_{S30} value (<200 m/s).

The PGA values for the *M*7.5 earthquake are larger than those for the *M*6.3 earthquake along the shoreline, in downtown area, and northern, eastern and north-eastern parts, which may be caused by the wave propagation through the crust to the top of bedrock underneath the investigated site and ground conditions (Subhadra *et al.* 2015).

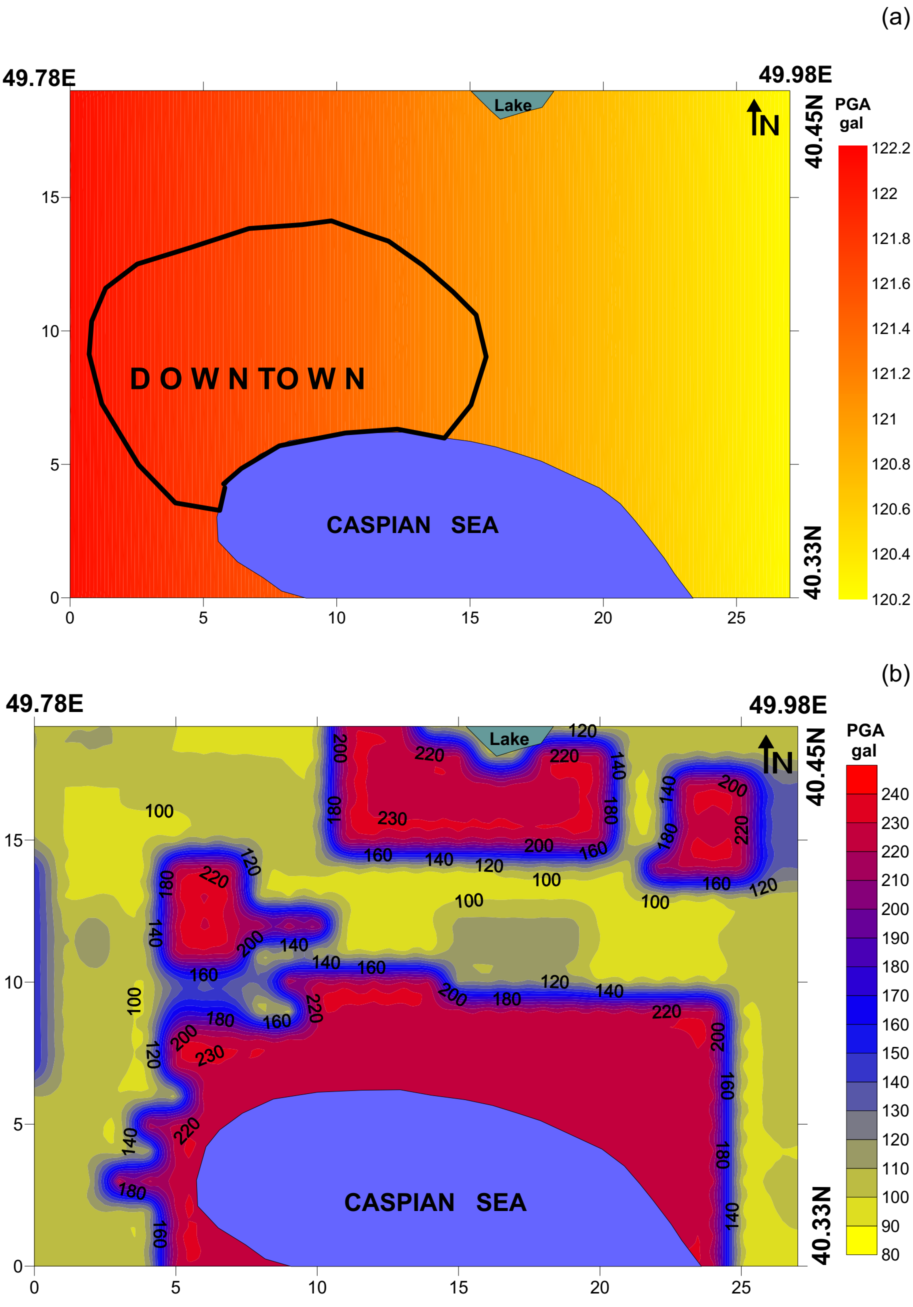
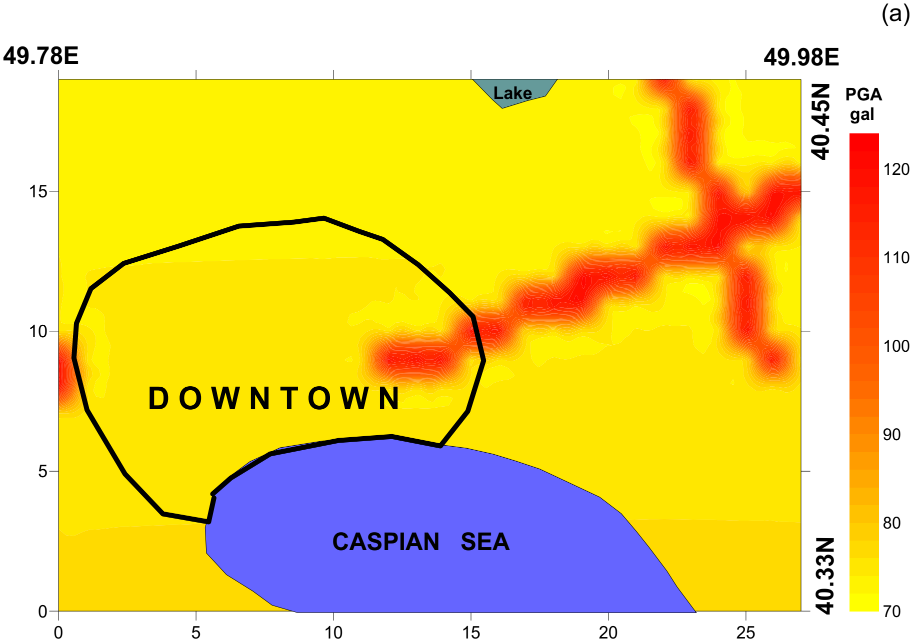


Fig. 9. Distribution of PGA at (a) bedrock level, and (b) surface for the largest *M*7.5 target earthquake (Shamakhi earthquake, 13 February 1902).



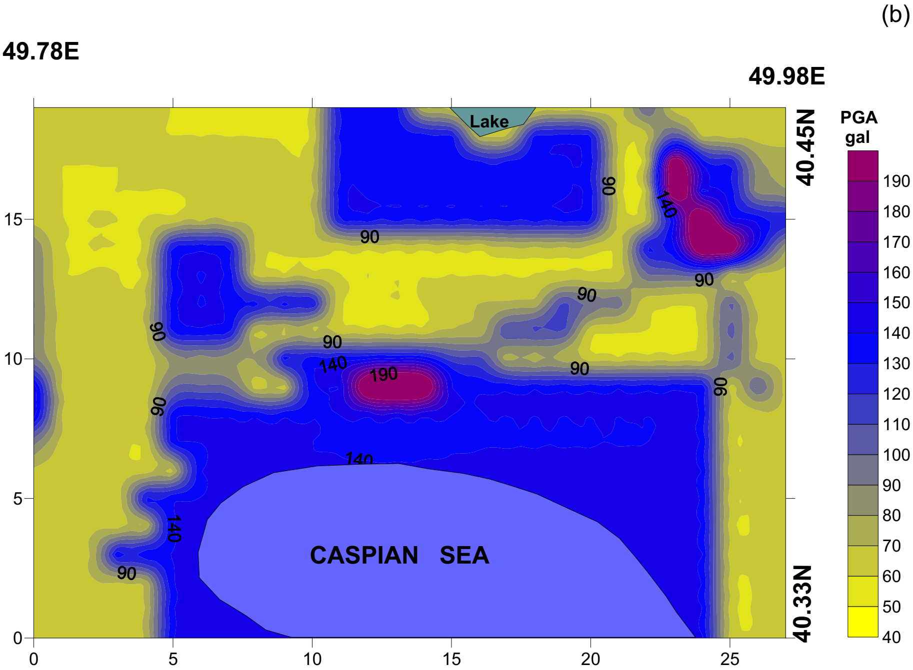


Fig. 10. Distribution of PGA at (a) bedrock level, and (b) surface for the largest *M*6.3 target earthquake (Caspian earthquake, 25 November 2000).

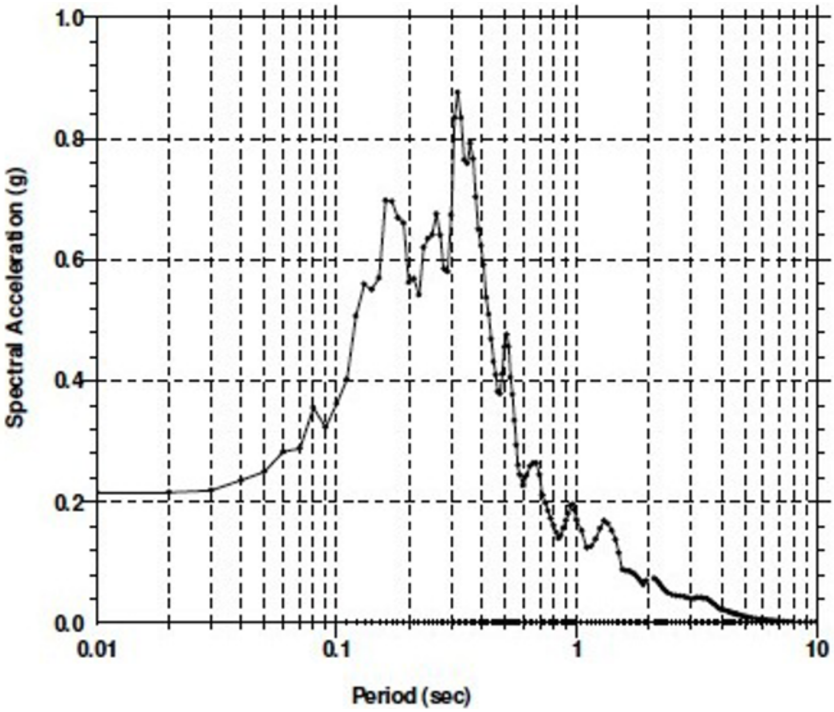


Fig. 11. The 5% damped response spectrum at the surface for $M6.3$ earthquake scenario in Baku.

Also, for both target earthquakes, the calculated PGA values are higher at the surface than at the bedrock level (Figs. 9 and 10); this could be due to shallow overburden between surface and rock that has amplified ground motion. As observed in Subhadra *et al.* (2015), the level of site amplification changes with site characteristics, type, thickness, stiffness and impedance contrast with the underlying bedrock. The PGA variations are maybe due to the typical characteristics of the material. For instance, seismic waves amplify or de-amplify, depending, respectively, on the low or high values of the *S*-wave velocity (V_{S30}) (Subhadra *et al.* 2015). A low *S*-wave velocity (V_{S30}) is in relationship with a high PGA (indicating larger potential damage to buildings) (see Figs. 9 and 10, especially throughout shoreline of the city). Amplification factors become larger in those sections of the city, characterized by lower *S*-wave velocity V_{S30} and higher seismic hazard. Therefore, a better estimation of the high seismic vulnerability of an area is a consequence of a better estimation of the susceptibility to higher ground motion amplification at the surface (Subhadra *et al.* 2015).

Although a quantitative assessment of surface PGA for the same near and remote events was performed for Baku city in 2010 by Babayev *et al.* (2010); in this current research we used more realistic values for the ground conditions, estimating the dynamic soil and rock layer properties of the study area considering also the relationship of shear modulus ratio *versus* both shear strain and strain-dependent damping, and the initial shear modulus. Besides, we used additional typical subsurface model (A1) to see if the surface PGA distribution was different from that obtained by Babayev *et al.* (2010). As a result, we verify that ground condition and additional subsurface model information do not influence significantly the overall results of surface PGA.

Additionally, in this study we calculated also bedrock PGA with the parameters of the same target scenario earthquakes which are important for the investigated area, but without considering amplification factor of the layers to the surface, in order to see the distribution of the bedrock seismic shaking caused by those two events. Such approach allowed us to estimate PGA at bedrock level to reveal the probable trace of the seismogenetic fault in order not to underestimate the peak ground acceleration near the fault at the surface, and consequently the intensity level. According to Fig. 10a, we observe the fault traces with intersection at the bedrock level, as it is revealed by the near event. This might drive to further investigations of the area in order to bring the results of tectonic researches, to investigate the seismic catalogue specifically for those cells where traces and intersection are, and to see the focal mechanism solutions of the events to estimate the fault types. Additionally, a number of models should be compiled for that area to characterize the surface layer on the fault area, to estimate how the maximum acceleration of strong ground motions is affected by the attenuation relationship between the peak acceleration and the distance.

4. CONCLUSION

The capability to build adequate emergency responses in case of strong earthquake occurrence or to draft suitable seismic disaster prevention plans closely rely to the availability of estimates of site effects due to earthquake ground motion. In this context, in the present paper we modeled the site specific ground motion and analysed the seismic response for the city of Baku, situated in Absheron peninsula. We used two earthquake scenarios related to *M*6.3 Caspian Sea earthquake (characterized as near event) and *M*7.5 Shamakhi earthquake (characterized as remote extreme event). We performed 1-D nonlinear analysis in order to understand more clearly the soil succession. We applied an amplification function which can explain nonlinear effect of soil and connect average *S*-wave velocity with the function in order to apply to the area where borehole data are provided to high density. In this study, we showed that the prediction of site amplification depends mainly on the average *S*-wave velocity at the surface. The proposed model enables to spatially map the ground motion of the urban city with boring database. The maps would furnish a solid basis to reinforce the structures in the area, to mitigate the consequences that possible seismic events with similar intensity could produce and, thus, to diminish significantly the probable loss.

Acknowledgements. The present study was supported by the bilateral project CNR-ANAS 2016-2017 and by the Science Foundation of SOCAR (grant No. 2, 2014-2015). The research was performed in the International Laboratory of Geology and Geophysics Institute of Azerbaijan National Academy of Sciences and the Institute of Methodologies for Environmental Analysis National Research Council (Italy) “Earthquake space-time analysis and hazard laboratory (ESTAHL)”.

References

Aptikayev, F., and Y. Kopnichev (1979), Considering focal earthquake mechanism at the prediction of strong motion parameters, *Doklady/Trans. USSR Acad. Sci.* **247**, 822-825 (in Russian).

Babayev, G. (2010), About some aspects of probabilistic seismic hazard assessment of Absheron peninsula. **In:** *Catalogue of Seismoforecasting Research Carried Out in Azerbaijan Territory in 2009*, Republican Seismic Survey Center of Azerbaijan National Academy of Sciences, “Teknur”, Baku, 59-64 (in Russian).

Babayev, G., and L. Telesca (2014), Strong motion scenario of 25th November 2000 earthquake for Absheron peninsula (Azerbaijan), *Nat. Hazards* **73**, 1647-1661, DOI: 10.1007/s11069-014-1159-7.

Babayev, G., A. Ismail-Zadeh, and J.-L. Le Moüel (2010), Scenario-based earthquake hazard and risk assessment for Baku (Azerbaijan), *Nat. Hazards Earth Syst. Sci.* **10**, 2697-2712, DOI: 10.5194/nhess-10-2697-2010.

Babayev, G., A. Tibaldi, F.L. Bonali, and F. Kadirov (2014), Evaluation of earthquake-induced strain in promoting mud eruptions: the case of Shamakhi–Gobustan–Absheron areas, Azerbaijan, *Nat. Hazards* **72**, 789-808, DOI: 10.1007/s11069-014-1035-5.

Boghdanovitch, G. (1904), Remarques sur le tremblement de terre a chemakha le 31 janvier 1902. **In:** *Comptes Rendue des Seances*, Tome I, Academie Imperiale des Sciences, St.-Petersbourg, Russia, 282-291 (in French).

Dobry, R., R.D. Borcherdt, C.B. Crouse, I.M. Idriss, W.B. Joyner, G.R. Martin, M.S. Power, E.E. Rinne, and R.B. Seed (2000), New site coefficients and site classification system used in recent building seismic code provisions, *Earthq. Spectra* **16**, 41-67, DOI: 10.1193/1.1586082.

Gasanov, A. (2003), Earthquakes of Azerbaijan for 1983-2002, Elm, Baku (in Russian).

Jackson, J., K. Priestley, M. Allen, and M. Berberian (2002), Active tectonics of the South Caspian Basin, *Geophys. J. Int.* **148**, 214-245, DOI: 10.1046/j.1365-246X.2002.01588.x.

Kirtas, E., P. Koliopoulos, G. Panagopoulos, E. Mouratidis, I. Sous, A. Kappos, N. Theodoulidis, A. Savvaidis, B. Margaris, and E. Rovithis (2016), Identification of earthquake ground motion using site effects analysis in the case of Serres city, Greece, *Int. J. Civil Eng. Architec.* **2**, 20-27.

Kondorskaya, N.V., and N.V. Shebalin (eds.) (1982), New catalog of strong earthquakes in the USSR from ancient times through 1977,World Data Center for Solid Earth Geophysics, Report SE-31, Boulder, Colorado, 608 pp.

Kuliyev, F.T. (1986), Otchet o seysmicheskom mikrorayonirovanii territorii Bolshogo Baku, Baku, Fond Instituta Geologii AN Azerb. SSR., 56 pp. (in Russian).

Levitski, G. (1902), Bulletin de la Commission Centrale Sismique Permanente (Annee 1902, Janvier-Juin), Academie Imperiale des Sciences, St.-Petersbourg, Russia, 62 pp.

Midorikawa, S., M. Matsuoka, and K. Sakugawa (1992), Evaluation of site effects on peak ground acceleration and velocity observed during the 1987 Chiba-ken-Toho-Oki earthquake, *J. Struct. Constr. Eng. Architec. Inst. Jap.* **442**, 71-78 (in Japanese with English abstract).

Ordónez, G.A. (2010), SHAKE2000: A computer program for the 1-D analysis of geotechnical earthquake engineering problems, User's Manual, GeoMotions, LLC, Lacey, Wash, 262 pp.

Sabetta, F., and J. Bommer (2002), Modification of the spectral shapes and subsoil conditions in Eurocode 8. **In:** *12th Europ. Conf. on Earthquake Engineering*, paper ref. 518.

Seed, H.B., I.M. Idriss, and F.W. Kiefer (1969), Characteristics of rock motions during earthquakes, *J. Soil Mech. Found. Div.* **95**, SM5, 1199-1218.

Shikhalibeyli, E. (1996), *Some Aspects of Geological Structures and Tectonics of Azerbaija*n, Elm, Baku (in Russian).

Shiuly, A., V. Kumar, and J.P. Narayan (2014), Computation of ground motion amplification in Kolkata Megacity (India) using finite-difference method for seismic microzonation, *Acta Geophys.* **62**, 3, 425-450, DOI: 10.2478/s11600-013-0169-2.

Subhadra, N., P. Simanchal, P. Prabhakara Prasad, and T. Seshunarayana (2015), Site-specific ground motion simulation and seismic response analysis for microzonation of Nanded City, India, *Nat. Hazards* **78**, 915-938, DOI: 10.1007/s11069-015-1749-z.

Telesca, L., G. Babayev, and F. Kadirov (2012), Temporal clustering of the seismicity of the Absheron-Prebalkhan region in the Caspian Sea area, *Nat. Hazards Earth Syst. Sci.* **12**, 3279-3285, DOI: 10.5194/nhess-12-3279-2012.

Theodoulidis, N., C. Karakostas, V. Lekidis, K. Makra, B. Margaris, K. Morfidis, C. Papaioannou, E. Rovithis, T. Salonikios, and A. Savvaidis (2014), The Cephalonia (Greece) earthquakes of January 26 and February 3, 2014: Effects on soil and built environment. **In:** *2nd Europ. Conf. on Earthquake Engineering and Seismology (2ECEES), 25-29 August 2014, Istanbul, Turkey*, Paper No. 3008.

Tonouchi, K., and F. Kaneko (1984), Methods of evaluating seismic motion at base layer. OYO Tech. Rep. No. 6, Japan, 12-14.

USGS (2010), U.S. Geological Survey Earthquake Hazards Program: Seismic Hazard Map for Haiti Region, available from: http://neic.usgs.gov/neis/eqdepot/2010/eq100112rja6/neic rja6w.html (accessed: 1 June 2010).

Veber, M.V. (1904), Recherches preliminaries sur le tremblement de terre a Chamakha. **In:** *Comptes Rendue des Seances*, Tome I, Academie Imperiale des Sciences, St.-Petersbourg, Russia, 238-241 (in French).

Received 14 March 2016
Received in revised form 29 July 2016
Accepted 23 August 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2171-2199
DOI: 10.1515/acgeo-2016-0088

Potential Field, Geoelectrical and Reflection Seismic Investigations for Massive Sulphide Exploration in the Skellefte Mining District, Northern Sweden

Saman TAVAKOLI[1], Mahdieh DEHGHANNEJAD[2],
María de los Ángeles GARCÍA JUANATEY[2], Tobias E. BAUER[1],
Pär WEIHED[1], and Sten-Åke ELMING[1]

[1]Division of Geosciences and Environmental Engineering,
Luleå University of Technology, Luleå, Sweden;
e-mail: saman.tavakoli@ltu.se

[2]Department of Earth Sciences, Division of Geophysics, Uppsala University,
Uppsala, Sweden

Abstract

Multi-scale geophysical studies were conducted in the central Skellefte district (CSD) in order to delineate the geometry of the upper crust (down to maximum ~ 4.5 km depth) for prospecting volcanic massive sulphide (VMS) mineralization. These geophysical investigations include potential field, resistivity/induced polarization (IP), reflection seismic and magnetotelluric (MT) data which were collected between 2009 and 2010. The interpretations were divided in two scales: (i) shallow (~ 1.5 km) and (ii) deep (~ 4.5 km). Physical properties of the rocks, including density, magnetic susceptibility, resistivity and chargeability, were also used to improve interpretations. The study result delineates the geometry of the upper crust in the CSD and new models were suggested based on new and joint geophysical interpretation which can benefit VMS prospecting in the area. The result also indicates that a strongly

Ownership: Institute of Geophysics, Polish Academy of Sciences;

conductive zone detected by resistivity/IP data may have been missed using other geophysical data.

Key words: potential field data, seismic reflection, resistivity, induced polarization, magnetotelluric, 3D modeling.

1. INTRODUCTION

The Skellefte mining district (Fig. 1) is regarded as a major ore district in Sweden and one of the main paleoproterozoic arc systems in the world (Weihed 2010) which produces base metals including Zn, Cu, Pb, Ag, and Au, from VMS and organic gold deposits (Bauer 2010). The economical importance of the district, and international increase in metal price, led to numerous exploration activities in the CSD during recent years (see, *e.g.*, Allen *et al.* 1996, Bauer *et al.* 2011, 2013; Dehghannejad *et al.* 2012a, García Juanatey 2012, Hübert *et al.* 2013, Skyttä *et al.* 2012, Tavakoli *et al.* 2012a, b; Weihed 2010). To date, nearly 160 million tons of ore is excavated, containing 1.9 ppm gold, 47 ppm silver, 0.7% copper, 0.3% zinc, 0.4% lead, 0.8% arsenic, and 25% sulphur (Kathol and Weihed 2005).

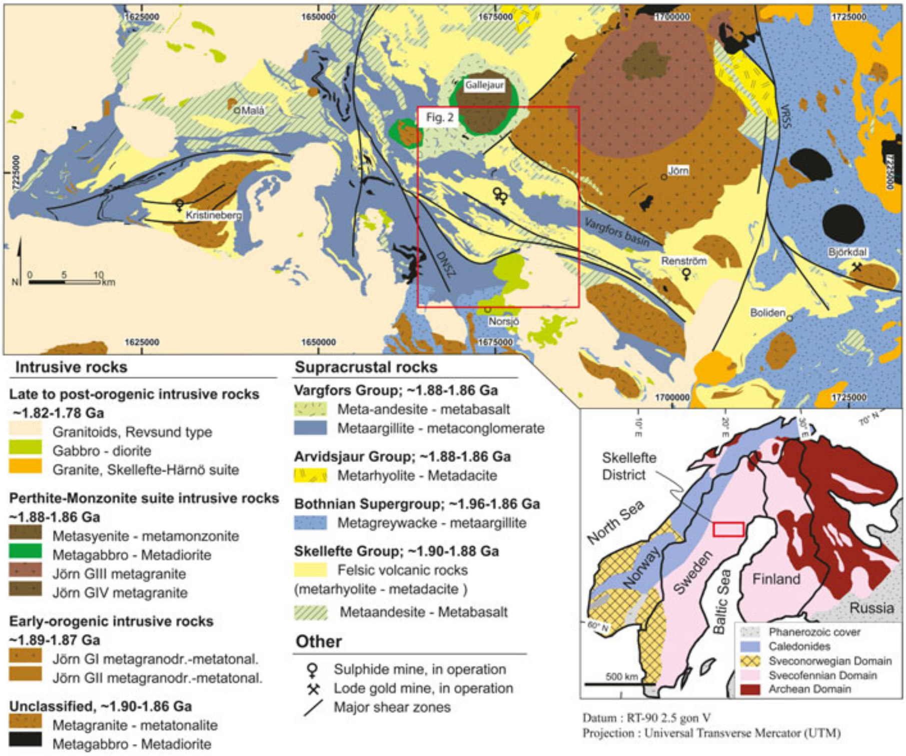


Fig. 1. Overview of the Skellefte district (modified after Kathol and Weihed 2005).

Exploration activities, including geological and geophysical investigations were initiated in 2008 in the CSD within a framework of the VINNOVA 4D project. These efforts were conducted to expand exploration activities using multiple geological and geophysical studies by modeling the crust in 3D, followed by illustrating the dynamic evolution of the crust through 4D-animation (Skyttä 2012). These geophysical investigations were conducted in the field between 2009 and 2010 (Table 1).

Table 1

Characteristic of the geophysical studies and the measured profiles in the CSD

Geophysical method	Year measured	Item	Profiles									
			I	II	E-1	3D	C1	C2	C3	P1	P2	P3
2009-2D resistivity/IP	2009	i	yes	yes	no	no	no	no	no	yes	yes	–
		ii	5.6	6.8	–	–	–	–	–	27	27	–
		iii	0.43	0.43	–	–	–	–	–	10	10	–
		iv	–	–	–	–	–	–	–	–		–
2010-2D resistivity/IP	2010	i	no	no	yes	no	no	no	no	no	no	–
		ii	–	–	10	–	–	–	–	–	–	–
		iii	–	–	1.5	–	–	–	–	–	–	–
		iv	–	–	–	–	–	–	–	–	–	–
3D resistivity/IP	2010	i	no	no	yes	yes	no	no	no	no	–	–
		ii	–	–	1.8	1.8	–	–	–	–	–	–
		iii	–	–	0.5	0.5	–	–	–	–	–	–
		iv	–	–	–	2.16	–	–	–	–	–	–
Magnetic	Various years	i	yes	yes	yes	yes	yes	yes	yes	yes	yes	–
		ii	5.6	6.8	10	–	22.2	21.5	24.45	22.2	21.5	24.45
		iii	0.43	0.43	1.5	–	4.5	4.5	4.5	4.5	4.5	–
		iv	–	–	–	–	–	–	–	–	–	–
Gravity	Various years	i	yes	yes	no	yes	yes	yes	yes	yes	yes	
		ii	5.6	6.8	10		22.2	21.5	24.45	22.2	21.5	24.45
		iii	0.43	0.43	1.5		4.5	4.5	4.5	4.5	4.5	
		iv	–	–	–		–	–	–	–	–	–
Seismic reflection	2009-2010	i	no	no	no	no	yes	yes	yes	yes	yes	yes
		ii	–	–	–	–	22.5	21.5	25.45	22.2	21.5	24.45
		iii	–	–	–	–	4.5	4.5	4.5	4.5	4.5	4.5
		iv	–	–	–	–	–	–	–	–	–	–
Magneto-telluric	2010	i	no	no	no	no	yes	no	yes	yes	yes	yes
		ii	–	–	–	–	22.5	21.5	25.45	22.2	21.5	24.45
		iii	–	–	–	–	4.5	4.5	4.5	4.5	4.5	4.5
		iv	–	–	–	–	–	–	–	–	–	–

Explanations: (i) Was this profile measured using/coincide with the following geophysical method?, (ii) L= measured length [km], (iii) D = approximate interpretation depth in this study [km], (iv) A= area [km^2].

Depending on each survey's aim and the desired investigation depth, different geophysical data were integrated for studying the shallow and deep-scale subsurface geology (Table 1). Although the majority of these geophysical data were previously interpreted and their joint or individual interpretation outlined several results which improved basic geological knowledge of the Skellefte district (Weihed 2015), a joint and multi-scale geophysical study which integrates all these previously acquired geophysical data was lacking. Therefore, we propose new models which are compatible with seismic, MT and potential field data. Moreover, in this study the magnetic and gravity model of Profile E-1 has been proposed to constrain previous interpretations of the resistivity/IP data.

The shallow-scale and deep-scale geophysical studies are classified based on the resolution and investigation depth of each method. Shallow investigations provide a high resolution image of the subsurface down to ~ 1.5 km depth and include 2D and 3D resistivity/IP data (Tavakoli *et al.* 2012a, 2016) constrained with potential field data. Deep investigations include seismic reflection (Dehghannejad *et al.* 2012a), MT (García Juanatey 2012) and potential field data (Tavakoli *et al.* 2012b) which provide a deep image of the subsurface down to ~ 4.5 km depth.

The potential field data have been modeled and interpreted together with resistivity and IP data in a local-scale along 2D profiles and in a 3D area. Tavakoli *et al.* (2012a) studied the effectiveness of the resistivity/IP method to detect the sulphide mineralization, or its alteration halo which was also constrained by magnetic and gravity modeling. The result proved promising, as several high chargeability zones were indicated, among which one with a known mineralization which was already proven by drilling. This was the motivation for deeper geoelectrical investigation thereby the subsurface resistivity down to ca. 1.5 km depth as well as inside a 3D area was imaged (Tavakoli *et al.* 2016). The result from new data strengthened some earlier interpretations, but also provided new insights about anomalies at greater depths (*i.e.*, 0.5-1 km depth). On the other hand, the 3D data, although cover a very small area compared to all other geophysical results, provided valuable result about the distribution of the sulphide ore around the Maurliden deposits (Tavakoli *et al.* 2016). Potential field data was also used to confirm, and in some cases improve the interpretation of the seismic reflection data (Tavakoli *et al.* 2012b). Among others, the TIB mafic intrusive rocks were imaged similarly on seismic reflection and potential field results, imaging the basal detachment zone at 3.5-4.5 km depth. The potential field data constrained the large-scale models in the areas where the seismic reflection sections were blank, or unable to provide information due to, *e.g.*, the verticality of the contacts. The MT data could particularly provide information about the deeper parts of the study area, *e.g.*, the ore related hydrothermally altered metavolcanic rocks within the Skellefte Group are depicted as conductors and have been found at depths between 1 and 6 km in the CSD. Shallower occurrences are not detected or are masked by conductive shales from the Vargfors Group. Also, the Jörn intrusion reaches 6.2 km depth and is characterized by an inhomogeneous distribution of resistivity values (García Juanatey 2012).

Similar studies with diverse range of geophysical data have focused on applying a joint inversion in multi dimensions using different approaches (direct parameter coupling and cross-gradient coupling) in terms of physical parameters and their coupling, for which the results show improvement over single inversions (see, *e.g.*, Moorkamp *et al.* 2011). Represas *et al.* (2015) examines the joint inversion of the gravity and resistivity data and compares it with the individual inversion of each dataset, which indicates that the joint inversion, although more complicated and time consuming, provides more realistic and geologically meaningful models than the ones calculated by inversion of each data set individually. However, this study does not aim to apply the joint inversion due to dimensionality and confidentiality of the data; instead, a joint interpretation is motivated.

Although these geophysical studies revealed interesting results from different scales (varying from ~ 0.5 km down to 4.5 km), they have never been interpreted jointly, together with other collected data from same scale. In this study we aim to integrate these models in the two scales of local and regional and confirm, or otherwise propose new interpretations accordingly. The joint interpretation of the data can then: (i) improve the shallow and deep-scale models of the subsurface by suggesting new interpretations based on the integrated geophysical studies, and (ii) identify new areas to prospect the hosting structures for the sulphide mineralization as a guide for future exploration activities and drilling plans.

2. GEOLOGICAL BACKGROUND

The Skellefte district rocks (Fig. 1) comprise metamorphosed Palaeoproterozoic supracrustal and intrusive rocks that were deformed and metamorphosed (Kathol and Weihed 2005). The 1.9-1.89 Ga Skellefte Group rocks are dominated by subaqueous lava domes, porphyritic cryptodomes, lavas and volcaniclastic rocks with largely rhyolitic composition (Allen *et al.* 1996) and minor occurrences of basalts, andesites and dacites (Billström and Weihed 1996) which form the lowermost stratigraphic unit in the district. Sedimentary rocks of the Vargfors Group overlie the Skellefte Group and form a distinct sub-basin, the Vargfors basin, in the northern part of the central Skellefte district. The lowermost parts of the Vargfors Group stratigraphy comprise turbiditic mudstones and sandstones overlain by monomict conglomerates, whereas the upper parts of the Vargfors stratigraphy are limited to the Vargfors basin and are characterized by polymict conglomerates formed in alluvial fans and braided river systems (Bauer *et al.* 2011, 2013). Metasedimentary rocks to the south of the central Skellefte district are regarded as rocks of the Vargfors Group due to their similar character and lithology. Their transition to Bothnian Supergroup metasedimentary rocks to the south of the district has been arbitrary (Kathol end Weihed 2005). The upper parts of Vargfors stratigraphy are dated at 1875 ± 4 Ma from an intercalated ignimbrite (Billström and Weihed 1996). The oldest intrusive rocks in the Skellefte district are represented by 1.89-1.87 Ga Jörn-type early orogenic intrusive rocks (Fig. 1). The central Skellefte district is characterized by a complex fault pattern of NNW-SSE-striking normal faults and associated NE-SW-striking transfer faults formed during crustal extension (Fig. 1; Bauer *et al.* 2011). The earliest tectonic deformation at 1.89-1.87 Ga is constrained to deeper crustal levels and formed synchronously with upper crustal extension and Skellefte Group volcanism (Skyttä *et al.* 2012). Subsequent SSW-NNE crustal shortening at 1.87 Ga resulted in reactivation of the NNW-SSE syn-extensional faults with S-side-up shearing (Bauer *et al.* 2011) and upper crustal coaxial deformation with steep to sub-vertical mineral lineations (Skyttä *et al.* 2012). The latest major deformation event at 1.82-1.80 Ga (Weihed *et al.* 2002) results from E-W crustal shortening causing reactivation and accompanied reverse shearing along steeply-dipping N-S-striking high-strain zones (Bergman Weihed 2001).

Allen *et al.* (1996) argues that VMS deposits in the Skellefte district could have been formed as sub-seafloor replacement within volcaniclastic sediments in the uppermost parts of the Skellefte Group stratigraphy. Previous studies suggested that the ore-forming hydrothermal fluids utilised the syn-extensional faults as fluid conduits and the ores precipitated in the vicinity of these faults (Allen *et al.* 1996, Bauer *et al.* 2013).

3. GEOPHYSICAL DATA

Seismic reflection, MT and potential field data have been used for deep-scale imaging of the subsurface (~ 4.5 km depth), and resistivity/IP data contributed to model a high resolution image of the geological structures within the shallower parts *(i.e.*, down to ~ 1.5 km). Most of interpretations were constrained by physical properties of the rocks, including magnetic susceptibility (μ), density, resistivity, and chargeability from both outcrop samples, and drill-cores (Tavakoli *et al.* 2012a). A brief description of the geophysical data including field surveys and interpretations are explained in this chapter. Information regarding each specific geophysical method, profile length and investigation depths is summarized in Table 1.

3.1 Resistivity and IP data

During two field campaigns, resistivity and chargeability of the subsurface was measured along three profiles (I, II, and E-1; Tavakoli *et al.* 2012a, 2016) and inside an area in 3D (Tavakoli *et al.* 2016). The field work was carried out using pole-dipole electrode configuration; however, each survey was designed using a specific electrode array and hence different penetration depths were acquired. The pole-dipole electrode array was preferred since the array is sensitive to horizontal variation of the resistivity at depth and according to Nyquist and Roth (2005) has higher signal strength compared to dipole-dipole array. In addition, its lower EM coupling compared to Wenner array and higher penetration depth compared to dipole-dipole makes this array favorite for this study.

2009-resistivity/IP Profiles I and II

In 2009, a geoelectrical campaign was initiated to delineate and model the geometry of the shallow subsurface, down to ~ 430 m (profiles I and II; Tavakoli *et al.* 2012a). The field work was carried out using pole-dipole electrode array; it consisted of five potential electrodes located 200, and 400 m apart. Profile I was extended over 6.8 km and positioned sub-parallel to the 5.6 km long profile II, covering a total length of 12.4 km around the Vargfors basin (Fig. 2). Further description of the field work and data processing is given by Tavakoli *et al.* (2012a).

2010-resistivity/IP Profile E-1

Profile E-1 was measured in 2010 to provide a deeper resistivity/IP image of the subsurface compared to the 2009 profiles (Fig. 2). In contrary to the 2009 profiles for which the potential dipoles (px-py) possessed a constant spacing of 200 m from the current electrode C1, in 2010 C1 was moved between a group of fixed dipoles which were connected differently, depending on the distance of the dipole to C1 (Tavakoli *et al.* 2016). Profile E-1 was 10 km long, covered the SW continuation of Profile II, and imaged the top 1.5 km of the upper crust. Since the field survey was conducted in a forward and reverse manner, the asymmetrical pattern of the data which is often produced when using pole-dipole array is reduced. Further details regarding the field survey and data processing are explained by Tavakoli *et al.* (2016).

2010-3D-resistivity/IP data

The resistivity/IP data were also measured in 3D (Fig. 2), covering a ~2.16 km^2 area of interpretation in the CSD (Tavakoli *et al.* 2016). The main purpose of the 3D resistivity/IP measurement was to better understand the 3D distribution of the sulphide mineralizations in the vicinity of Maurliden-North, East and Central mineralization and to investigate the possibility of

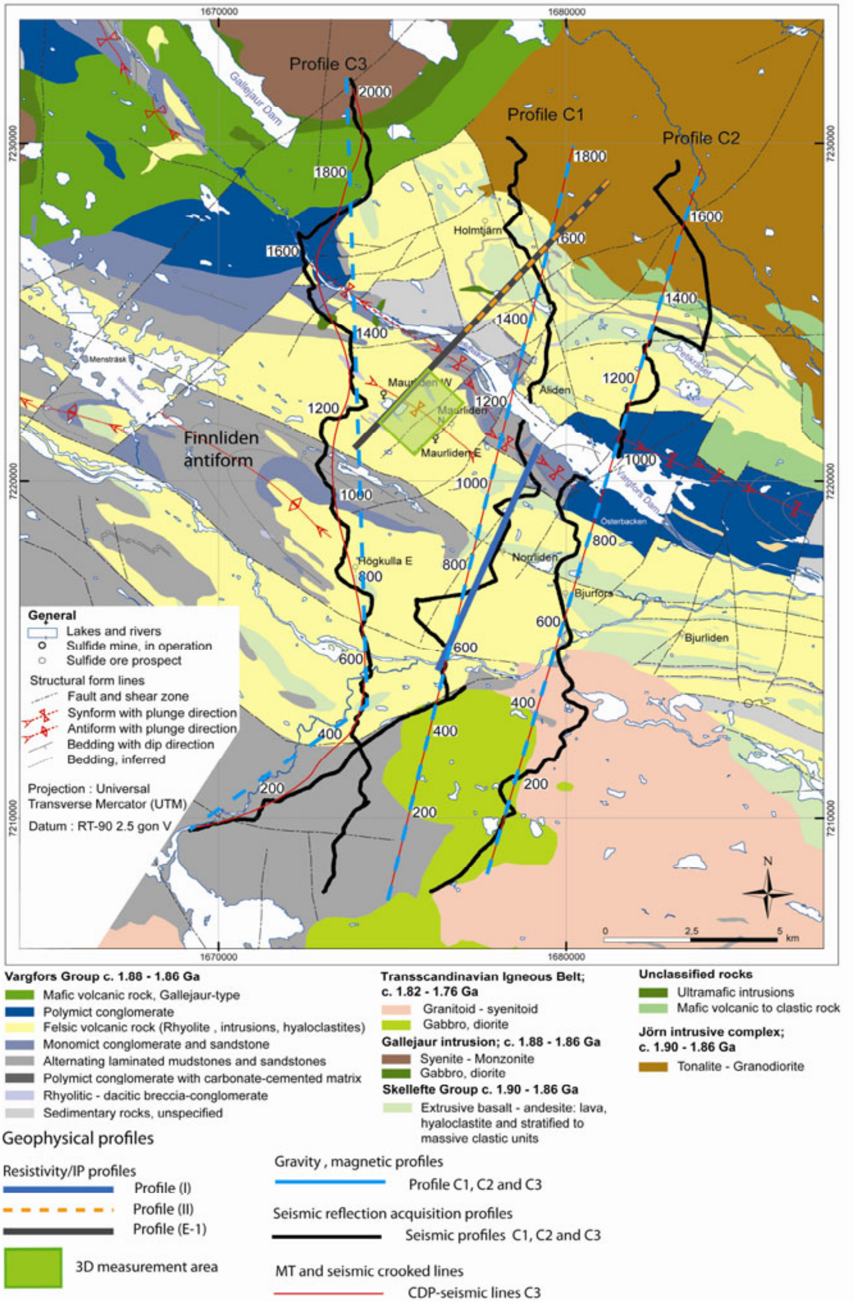


Fig. 2. Geophysical profiles; profiles I, II, E-1 and 3D-area for shallow-scale investigations and profiles C1, C2 and C3 for deep-investigations (modified after Bauer 2010).

detecting new geological features related to the sulphide mineralization. The result imaged the resistivity/chargeability distribution of the top ~450 m of the crust. Further details about the 3D-survey are explained by Tavakoli *et al.* (2016).

3.2 Magnetic and gravity data

Magnetic and gravity data were compiled from a database of SGU and Boliden mineral AB. The magnetic data consisted of a network with 40 m station spacing, whereas the spacing between gravity data varied between 200 and 800 m (Tavakoli *et al.* 2012a). In the sparsely distributed gravity areas along resistivity/IP profiles, additional gravity data were measured by Boliden Mineral AB in 2010, to increase the data density in the vicinity of the profiles. The IGRF correction was applied to the magnetic data with the total field intensity of 52612, inclination 76.7, and declination 6.7 (Model Vision Pro™, Encom Technology). However, the data in the modeled sections are presented in the original form. In addition, all necessary corrections, *i.e.*, latitude, drift, tidal, free air, instrument height, and Bouguer plate have been applied on the gravity data.

The magnetic and gravity modeling was performed on all geophysical profiles in this study (profiles I, II, E-1, C1, C2, and C3) in different scales: (i) shallow 2009-resistivity/IP profiles I and II down to 430 m depth, (ii) 2010-resistivity/IP Profile (E-1) down to 1.5 km, and finally (iii) reflection seismic and MT profiles C1, C2, and C3 down to 4.5 km depth. Potential field modeling of the geophysical profiles was carried out using Model Vision Pro™ (Encom Technology) in 2.5D, where the geological bodies are assumed to have a variable length in their strike direction. Further details about the potential field modeling and their processing procedure are explained by Tavakoli *et al.* (2012a, b).

3.3 Seismic reflection data

During 2009-2010, three sub-parallel and ~N-S trending seismic reflection profiles were acquired (profiles C1, C2, and C3; Fig. 2) to constrain a 3D geological model of the study area (see, Dehghannejad *et al.* 2012a). Each profile was about 30 km long and located approximately 3-7 km apart from the neighboring profile. The profiles were placed perpendicular to the main structural grain of the central Skellefte district. A SERCEL 408UL recording system from the Department of Earth Sciences, Uppsala University, was used for the data acquisition. A hydraulic hammer, VIBSIST, was used to generate the seismic signal and enabled imaging the subsurface structures down to ~4.5 km depth. The vehicle-mounted hammer hit the ground repeatedly at an increasing rate for a predetermined time span (Juhlin *et al.* 2010, Dehghannejad *et al.* 2010, Malehmir *et al.* 2011). The geophones had a natural frequency of 28 Hz. In total, about 3000 shot points were generated along the three profiles with a nominal number of 2-4 sweeps per shot point, depending on the ground condition. Shot spacing was 25 m except in a few places where regular shot spacing was not possible due to a lack of road accessibility. Receiver spacing was 25 m with a large gap where Profiles C1 and C2 cross the Skellefte River. Processing work mainly followed conventional approaches in crystalline environment and experiences obtained from crooked-line data processing in the western part of the Skellefte district (Dehghannejad *et al.* 2010, 2012b). Further details about the seismic data acquisition, processing and results are explained by Dehghannejad *et al.* (2012a).

3.4 MT data

In fall 2010, 34 broadband MT sites were installed in the central Skellefte District, nearby the Maurliden mine. The sites had 1-2 km spacing along two profiles in the NNE-SSW direction with ~23 km length and 3 to 4 km in between. The five MT channels, four for the horizontal electric and magnetic fields and one for the vertical magnetic field, were recorded. The data processing and estimation of the MT transfer functions, was carried out with the algorithm MTU2000 of Smirnov (2003). The obtained transfer functions were in the range of 700 Hz to 200 s, showing a decrease in apparent resistivities, with increasing periods, from 10 000 to 200 Ωm, while phases increase from 40° to 80°. Further details about the acquisition, processing, inversion and interpretation of the data are explained by García Juanatey (2012).

4. RESULTS

In this part, a summary of the geophysical investigations in different scales is presented. The integrated interpretation of the shallow-scale and deep-scale models were then presented and discussed in the discussion parts which includes new interpretations as well as supports to some of the earlier model results.

4.1 Shallow investigations (2D, 3D resistivity/IP and potential field data)

Shallow investigations include interpretation for profiles I, II, E-1 and the 3D-measurement area. In addition to profiles I and II which were modeled with magnetic and gravity data (Tavakoli *et al.* 2012a), the gravity and magnetic response of Profile E-1 is also modeled and presented. The 3D-area partly coincides with Profile E-1 and has a too small surface coverage to be modeled with magnetic and gravity data; hence they have not been modeled with potential field data. The inversion of the resistivity profiles was carried out using Res2Dinv (Loke 2012) with least-square method. The resistivity and chargeability models did not indicate any irregular variation in the values in lower sections. Therefore, we used the default depth weighting factor of 1.05 to compensate for the resolution loss at greater depths. We also allowed the program to determine the depth weighting factor automatically which the resulting section was similar to when 1.05 was used. The effects of the side blocks were also slightly diminished to decrease the effect of artefacts in the inversion result.

Profiles I and II

Basic information about the subsurface geometry of Profile I was acquired from resistivity/IP investigations (Fig. 3a). A study of the resistivity and

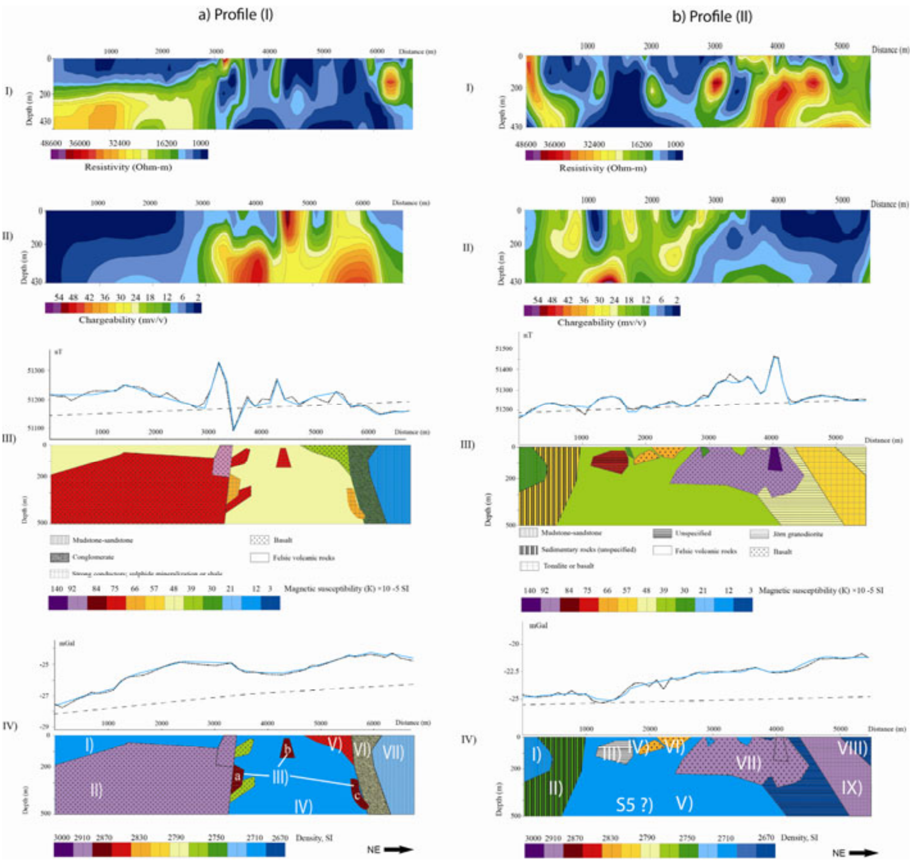


Fig. 3. RIMD models of profiles I and II; (a) Profiles I: (i) resistivity, (ii) IP, (iii) magnetic, (iv) gravity; (b) Profile II: (i) resistivity, (ii) IP, (iii) magnetic, (iv) gravity.

chargeability of the massive sulphide samples from Norrliden-N deposit indicated a zone with high chargeability and high conductivity (Tavakoli *et al.* 2012a). The resistivity/IP result from Profile I imaged three zones (a, b, and c; Fig. 3a-IV). Tavakoli *et al.* (2012a) explain these anomalies as new potential prospects for sulphide mineralization or graphitic schist. The resulted resistivity/IP data was further modeled using potential field data (Figs. 3a-III and 3a-IV). The RIMD (resistivity, IP, magnetic, density) model delineated the geometry of Profile I down to 430 m depth (Fig. 3a). Profile (II) was measured sub-parallel to Profile I and cut through the contact between the Vargfors-Skellefte Group and Skellefte-Jörn units (Tavakoli *et al.* 2012a). The synclinal structure of the Vargfors basin was indicated on both resistivity and IP data, where the inter-sedimentary contact of the Vargfors basin was imaged as a result of a resistivity contrast between the sandstone and unspecified sedimentary rocks (contact I-II; Fig. 3b). In addition, the NE-dipping contact between the Skellefte Group rocks and the Jörn intrusion is indicated on final (RIMD) (contact V-VIII; Fig. 3b). An outstanding high IP anomaly at $\sim x = 1300$ m (body S5; Fig. 3b-IV) was identified within the deeper parts of the Skellefte Group felsic volcanic rocks and therefore interpreted as a likely structure related to the alteration zones which might envelope the sulphide mineralization or graphitic schist. However, since the high IP anomaly was located at the bottommost part of the depth section, additional investigations are required for any further interpretations.

Profile E-1

Profile E-1 is a horizontal extension of Profile II to the SW (Fig. 2), and in comparison to Profile II, images the subsurface down to a greater depth (~ 1.5 km). Several IP anomalies were identified in the inversion result of Profile (E-1) which are mainly associated with low resistivities (S1-S5 in Fig. 4a and b; Tavakoli *et al.* 2016). The high chargeability anomaly labeled S5 in Profile II is associated with S4 and S5 in Profile E-1 which implies that the high IP signature in Profile II was not an artifact, but a possible indication of a conductive structure related to the sulphide mineralization or a graphitic schist (Tavakoli *et al.* 2016). Another IP anomaly interpreted as a graphitic schist (V4; Fig. 4b) which dips to the SW, together with S5 appear in the vicinity of the Skellefte-Vargfors Group contact in the north, and thus are the most likely features related to the alteration zones around sulphide mineralization in Profile E-1 which are enveloped by V4 (Fig. 4a and b; Tavakoli *et al.* 2016).

3D-measurement area

Result from interpretation of the 3D resistivity/IP data includes eight depth sections, starting from 50 m in depth down to ~ 450 m (Fig. 5). A part of the

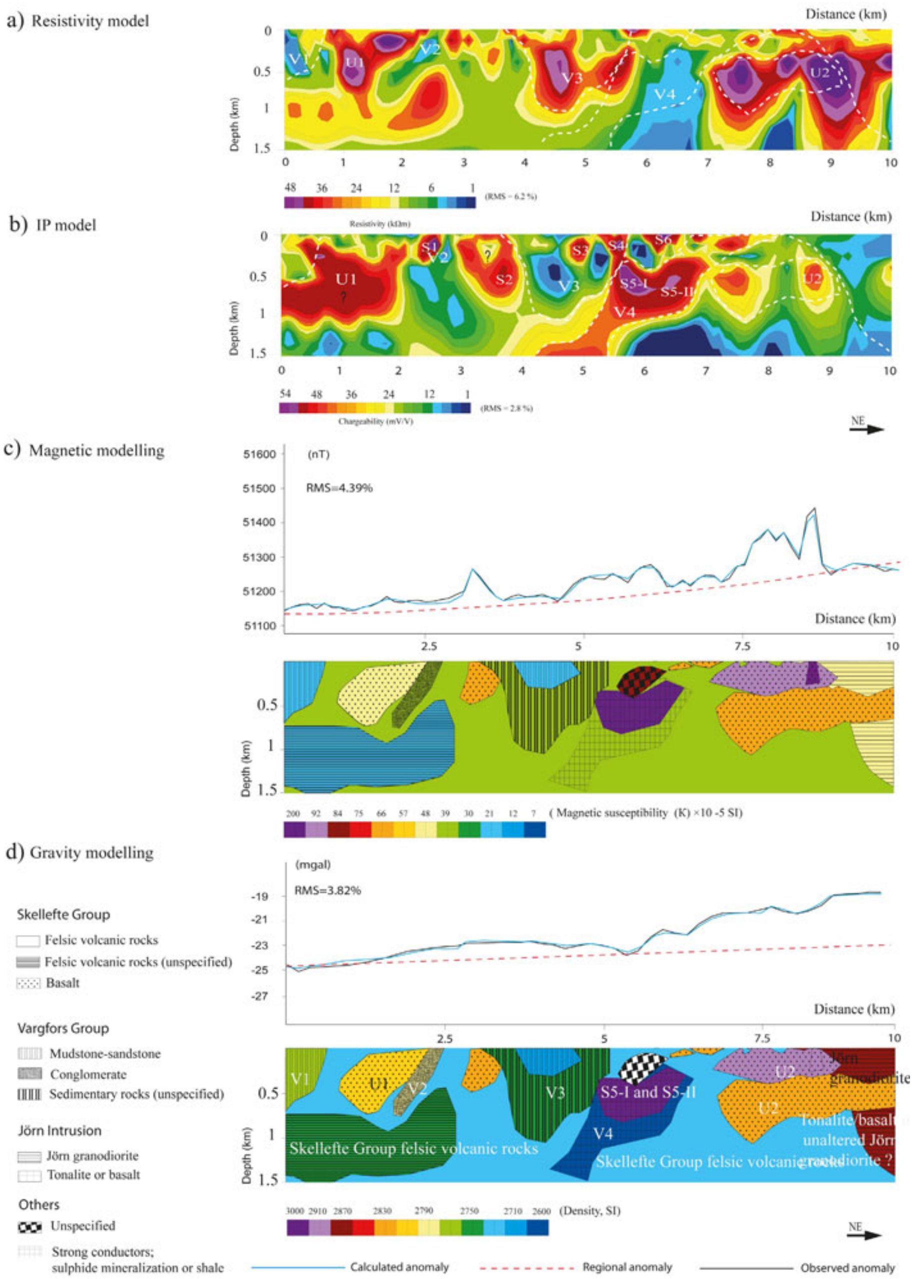


Fig. 4. Integrated geophysical modeling along Profile E-1: (a) resistivity model, (b) IP model, (c) magnetic model, (d) gravity model.

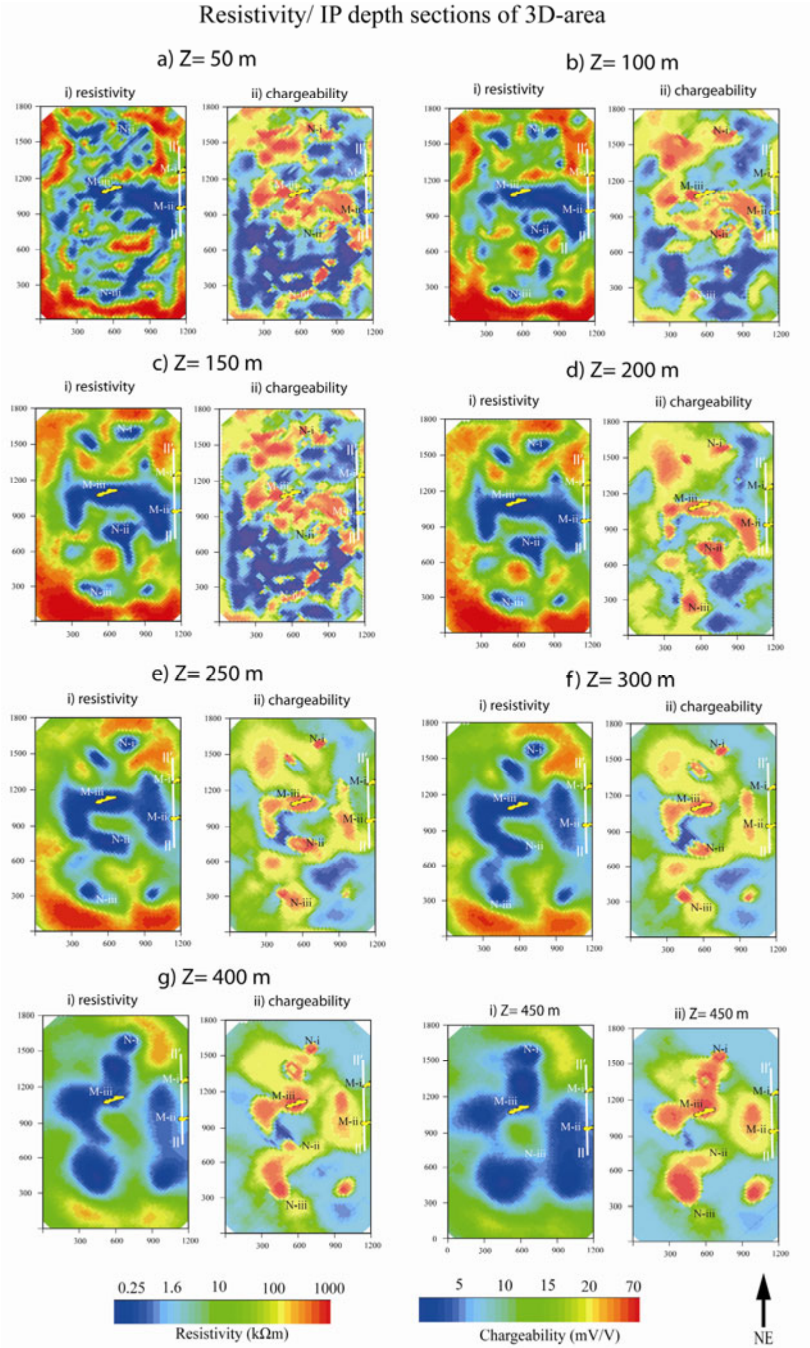


Fig. 5. Resistivity and IP depth sections of the 3D-area for shallow interpretations down to 450 m depth.

3D-area coincides with the profile II-II' of Montelius *et al.* (2007) which studied the distribution of the sulphide mineral related to the Maurliden (North, East and Central) domain.

A Central Conductive Zone (CCZ) which coincides with the Maurliden-East and Central mineralization was identified on both resistivity and IP data from 3D-area. The conductivities are within the range of ~ 1 kΩm and have chargeability ranged between 20 and 70 mV/V, which is within the range of measured chargeability for the VMS deposits (Tavakoli *et al.* 2012a). The unusually high resistivity of the Maurliden-North deposit (~ 50 kΩm) is related to the unaltered structure of the hosting felsic volcanic rocks or low concentration, or disseminated pattern of the VMS deposits (Tavakoli *et al.* 2016). Anomaly M-iii shows the greatest depth extent among the three studied Maurliden mineralizations (Fig. 5). In addition to the Maurliden mineralizations, three new conductive zones were detected inside the 3D-area which may be related to the structures hosting sulphide mineralization (Fig. 5; Tavakoli *et al.* 2016).

4.2 Deep investigations (seismic reflection, potential field, and MT data)

Preliminary interpretation of the seismic profiles C1, C2, and C3 (Fig. 2) was conducted by Dehghannejad *et al.* (2012a). Several reflectors were explained as major faults and shear-zones and were related to lithological contacts (Dehghannejad *et al.* 2012a). The magnetic, and gravity response of the three seismic profiles were calculated down to ~ 5 km in depth to verify or otherwise improve the seismic interpretations (Tavakoli *et al.* 2012b). In addition, the resistivity of the subsurface structures in profiles C1, C2, and C3 has been estimated through inversion of the 3D MT data (García Juanatey 2012).

Seismic reflection data

Seismic reflection data along Profile C1 reveals that the southernmost part of this profile is more reflective than the central and the northern part (Fig. 7a). A series of south-dipping reflections, each with slightly different dip angle, extend to the surface and were interpreted to depict a series of major shear zones (Dehghannejad *et al.* 2012a). The northern parts of Profile C1 are dominated by shorter reflections which are related to the southern contact of the Jörn intrusive complex.

In contrast to the results observed along Profile C1, the central part of C2 is more reflective than its southern part. A synformal structure extending to a ~1.7 km depth corresponded to the Vargfors basin (R8 and R15 in Fig. 6b; see also Bauer 2010). In addition, several weak reflections observed in the northern part of the profile C2 were interpreted to originate from the southern contact of the Jörn intrusive suite (Dehghannejad *et al.* 2012a). A series

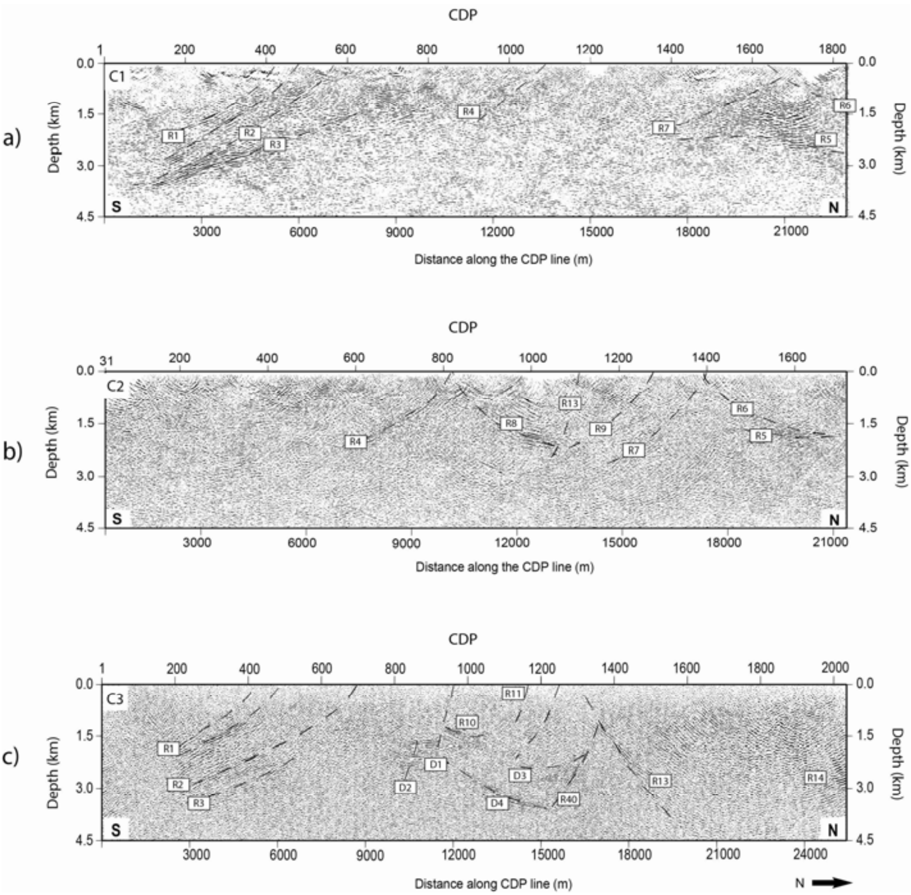


Fig. 6. Migrated seismic section of the seismic reflection profiles along the CDP lines in the central Skellefte district: (a) Profile C1, (b) Profile C2, and (c) Profile C3. Vertical: horizontal (scale) = 90% for a, b, and c. Modified after Dehghannejad *et al.* (2012a).

of south-dipping reflections are observed in the southern part of the Profile C3, similar to the south-dipping reflections identified on Profile C1 (Fig. 7c). These reflections show similar character in both profiles, suggesting identical geological structures (Dehghannejad *et al.* 2012a). The diffractions, as well as some reflections, are segmented by a set of south-dipping shear zones, sub-parallel to other reflections (Dehghannejad *et al.* 2012a).

Potential field modeling

The magnetic and gravity modeling along Profile C1 suggest that among south-dipping reflectors (R1, R2, and R3) in profiles C1 and C3, only R2 co-

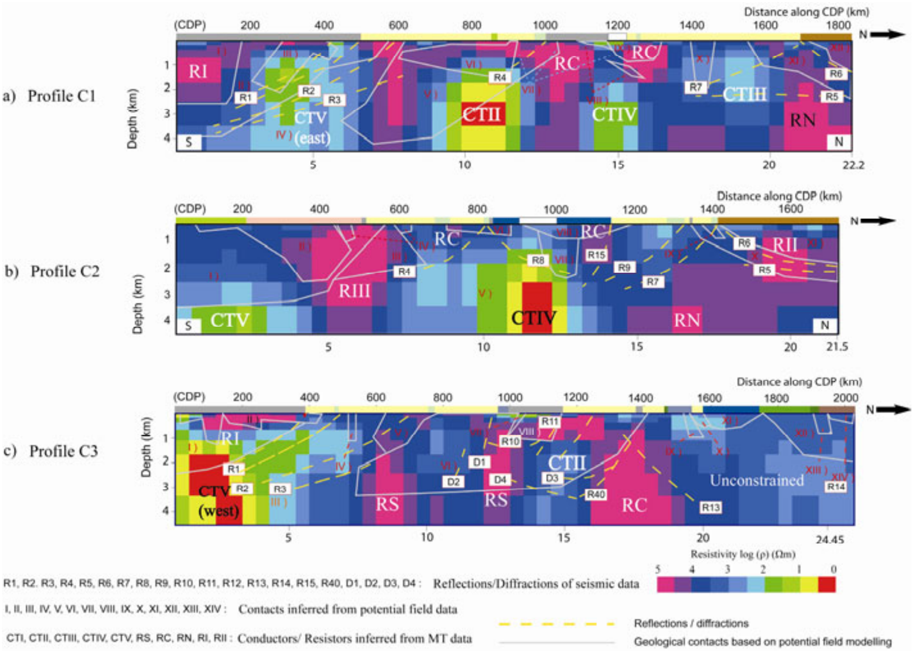


Fig. 7. Integrated deep-investigations along profiles C1, C2, and C3 using seismic reflection (Dehghannejad *et al.* 2012a), potential field data (Tavakoli *et al.* 2012b), and MT data (García Juanatey 2012).

incides with the model inferred from interpretation of the potential field data (Fig. 7). In addition, in the north, reflectors R6 and R7 support the result from magnetic and gravity modeling which delineates the NE-dipping contact between the Jörn intrusion and its underlying rocks (contact XI-XII; Fig. 7a) as well as the contact between the deep basalt and Skellefte Group felsic volcanic rocks (contact IV-X in Fig. 7a; Tavakoli *et al.* 2012b).

Modeling of the magnetic and gravity data greatly contributed to the interpretation of the southern part of Profile C2, since no reflector was observed within the 7 km beginning of this profile (Fig. 7b). Thus, the spatial relationship between lithologies in this part (TIB Gabbro, Skellefte Group, and Vargfors Group rocks) is only delineated and explained by the result of magnetic and gravity data. Further to the north, reflectors R4, R6, and R8 coincide with the models acquired from interpretation of the magnetic and gravity data (Tavakoli *et al.* 2012b).

Geometry of the key geological structures along Profile C3 was well delineated in areas where the seismic reflections/diffraction coincided with the result from magnetic and gravity modeling (R1, R3, R11, R4, and D3; Fig. 6c). Potential field data, however, did not provide an insight on subdivision of the blocks indicated by diffractions D1, D2, and reflection R10, since their surrounding lithology indicates similar petrophysical signatures; this is also due to the large station spacing for the gravity data (Tavakoli *et al.* 2012a).

MT models

The sliced depth sections of the 3D MT data in this study were integrated for interpretation along CDP lines of the profiles C1, C2, and C3. However, in order to compare MT interpretation with the result from seismic and potential field data, deep investigation in this study is limited down to 4.5 km depth.

Prominent and deep sub-vertical conductors at the fringe of the model (*e.g.*, CTV; Fig. 7b) are possibly related to the crustal-scale shear zones that surround the study area. Additionally, several conductors at intermediate depths (between 1-4 km depth) are associated to hydrothermally altered rocks within the ore bearing Skellefte Group (*e.g.*, CTII, CTIII, and CTIV; Fig. 6a and b). These conductors occur at the intersections of near-vertical transfer faults inferred from field observations (Bauer *et al.* 2011), normal listric faults interpreted from seismics (Dehghannejad *et al.* 2012a), and possible vertical faults imaged by the 3D MT model. In addition, several high resistivity structures were identified within upper parts of the models, supporting interpretation of the potential field data (Tavakoli *et al.* 2012b), which relates these anomalies to shallow basalts (*e.g.*, II, IV, V; Fig. 7c).

5. DISCUSSION

5.1 Shallow interpretations down to 1.5 km depth

This section summarizes new conclusions from joint interpretation of profiles I, II, E-1 and a 3D-area based on the resistivity/IP and potential field data. The top 1.5 km of the MT data were also taken into account for possible correlation. The key interpretations inferred from modeling of profiles I, (II), and (E-1) and 3D data are described below:

(i) Resistivity and chargeability data for body II (Fig. 3a-IV) which was interpreted as Skellefte Group basalt agrees with MT interpretation indicating similar high resistivity response on MT and resistivity sections (at CDP 600-900; Fig. 8a).

(ii) The NE-dipping geometry of the Vargors-Skellefte contact in its southern part is consistent in profiles E-1 and I (IV-VI in Profile I and AF-V3 in Profile E-1; Fig. 7). However, understanding the depth to the deeper parts of the Vargfors basin along Profile (I) was not possible due to the limited investigation depth. Yet, Profile (E-1) shows that this part of basin is ex-

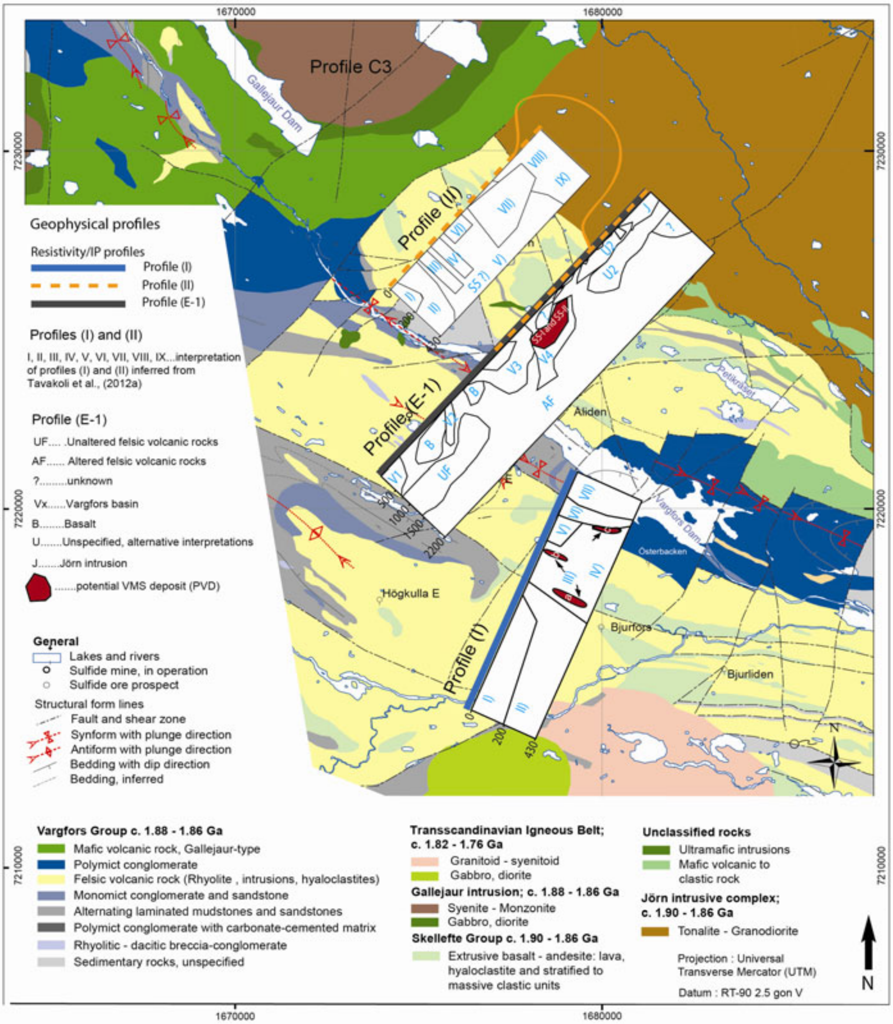


Fig. 8. 2D-shallow profiles and their interpretations inferred from resistivity, IP and potential field data; depth sections have been exaggerated vertically (modified after Bauer 2010).

tended at depth down to a maximum of ~ 1 km, but the maximum depth of the basin is probably more than this.

(iii) A high chargeability body (S5 in Profile II; Fig. 7) which was observed in Profile II in the study by Tavakoli *et al.* (2012a) is probably an indication of structures related to the sulphide mineralization (S5-I and S5-II in Profile E-1; Fig. 7) which implies that S5 is very likely to denote sulphide mineralization and is enveloped by a conductive material such as graphitic schist. This interpretation agrees well with a result from potential field modeling along Profile E-1 where the sink in the gravity anomaly coincides with low density graphitic schist (body V4; Fig. 8) and underlies the Vargfors basin close to the contact with the felsic volcanic rocks (Tavakoli *et al.* 2016). The conductive zone labeled S5 and its enveloping graphitic schist is also identified in the MT data along Profile C1 between 0.5 and 2.5 km depth (CTIII at CDP 1400-1600 in profile C1; Fig. 7a).

(iv) The geometry of the Vargfors basin in profile (E-1) is consistent with the resistivity, IP and magnetic interpretations (Fig. 4a-c). Also, its synformal structure, suggested by Bauer (2010), agrees well with the joint interpretation along the studied profiles I, II, and E-1. However, the unusual gravity high for the basin along 2D-profiles (Figs. 3a-IV, 3b-IV, and 4d) probably resulted from large scale and deep structures at greater depths, or low resolution of the gravity data due to large station spacing.

(v) New potential field modeling along Profile (E-1) reveals that the Jörn intrusion is underlain by unaltered granodiorites, with minor occurrences of basalts, which compensate the gravity high within the northernmost end of Profile E-1. This interpretation is in agreement with pervious resistivity/IP, potential field, seismic and MT studies (reflector R6, resistor RN; Fig. 8a and b) (Dehghannejad *et al.* 2012a, García Juanatey 2012, Tavakoli *et al.* 2012a, b, 2016).

(vi) New result from potential field modeling along Profile E-1 indicates that almost all anomalies labeled S, which were explained as potential features related to the sulphide mineralizations, are associated with magnetic highs. In addition, the SE part of Profile E-1 is associated with magnetic and gravity lows, which is in agreement with the dominant lithology inferred from interpretation of the geophysical data (Fig. 4).

The conductive zones inside the 3D-area imaged the known Maurliden (North, East, and Central) mineralizations (Mi, M-ii, and M-iii; Fig. 5) as well as new zones (N-i, N-ii, and N-iii; Fig. 5) which provide new possibilities for detecting structures related to mineralization. A comparison between the resistivity/chargeability result from 2D and 3D data reveals that S1 can depict the Maurliden mineralization (Figs. 4b and 5). Due to a rather limited depth extent of the 3D-area (~ 450 m), and since the deep profiles C1, C2 or C3 do not coincide with, or pass nearby the 3D-area, seismic reflection or MT, or potential field data are unable to contribute to constrain this interpretation. The top-left corner of the 3D area indicates an interesting anomaly on the chargeability model sections which continues from near surface down to 400 m at depth. S1 could therefore be a projection of this anomaly on Profile E-1, which makes this anomaly an interesting feature related to the VMS mineralization.

5.2 Deep interpretations down to 4.5 km depth

The joint interpretation of the potential field, seismic reflection and MT data along three sub-parallel profiles C1, C2, and C3 was conducted to delineate a deep image of the CSD and new models were proposed. As a result, the earlier magnetic and gravity models have undergone minor modifications to fit the MT, seismic and potential field data simultaneously. The modified models (Fig. 7) are indicated with brown lines in the depth sections (Fig. 9). Seismic reflection data have been used for VMS exploration in other studies and have proven successful. Exceptions are, *e.g.*, short and shallow reflectors that cross-cut through the main lithological contacts which may be difficult to provide explanation for all such reflections (Bellefleur *et al.* 2015). Hedin *et al.* (2013) in an integrated geophysical study in central Sweden conducted 3D interpretation of the subsurface structures in the vicinity of boreholes based on seismic reflection and bedrock geology data, which result was constrained by 3D inversion of the gravity data. The integrative interpretation resulted in better control on the structures of interest for the scientific deep

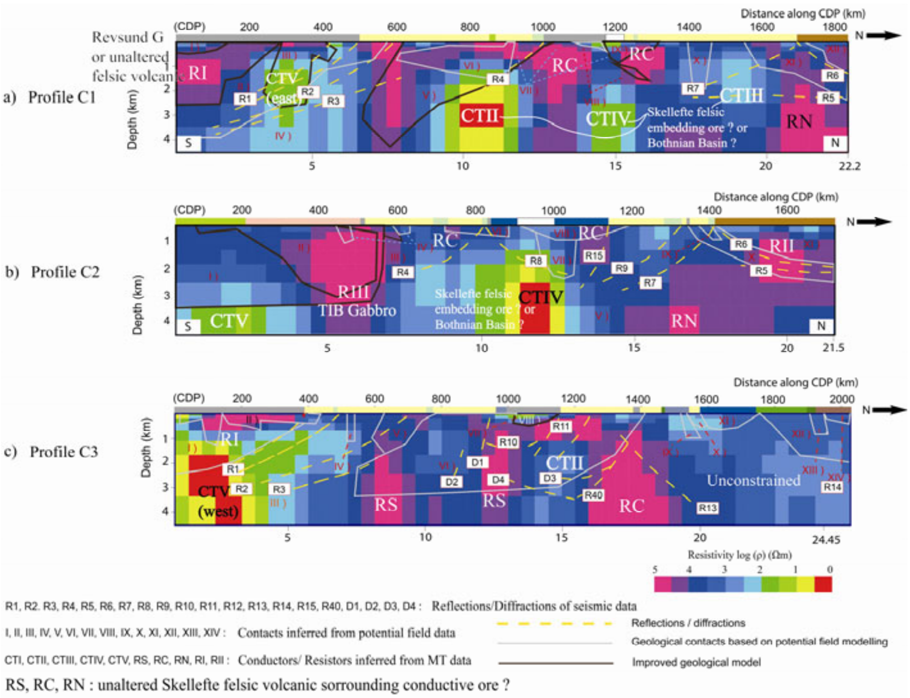


Fig. 9. Improved deep-scale model acquired inferred from seismic reflection, potential field data and MT data. Geometry of the several geological bodies has been modified based on MT interpretations.

drilling project. Hübert *et al.* (2013) used magnetotelluric data to construct 3D resistivity model of the upper crust along the previously collected seismic profiles in Kristineberg area, northern Sweden. The study utilized integration of seismic reflection data into the interpretation that demonstrated good correlation between the reflectors and main electrical features along the MT profiles and provided support to some of the earlier interpretations based on the seismic data, but also proposed new interpretations.

Profile C1

The laminated sandstone-mudstone (body I; Fig. 7a) in the SW end of profiles C1 and C3, demonstrates an unusual high resistivity (100 kΩm) on MT data along Profile C1, which is well above the expected resistivity (~ 4200 Ωm) of the Vargfors sandstone-mudstone based on the study by Tavakoli *et al.* (2012a). Although the seismic data could not image the contact between body I and the surrounding rocks, MT data nicely corresponds to the geometry of the sandstone-mudstone inferred from potential field interpretation (body I; Fig. 7a).

The deep conductor, CTV (east) coincides with bodies II, III, and IV in profile C1. Reflector R2 was coincided and interpreted as a contact between TIB Gabbro and felsic volcanic rocks based on potential field-seismic interpretation (Fig. 7a). However, García Juanatey (2012) relates the deep conductor CTV (east) to either (i) an alteration zone which implies that the Skellefte volcanic rocks are dipping and extending further to the south, or (ii) a near surface signature of the crustal detachment explained by Skyttä *et al.* (2012). However, according to the potential field interpretations (Tavakoli *et al.* 2012b), and given the fact that body II indicates an inconsistent internal structure (density and magnetic susceptibility variations), we suggest that the western part of body II (at x = 3-4 km; Fig. 9a) is attributed to the TIB Gabbro, with high resistivity, density and magnetic susceptibility. Conductor CTIV, with ~ 10-100 Ωm resistivity is explained as an alteration zone embedding the sulphide mineralization (García Juanatey 2012), and therefore supports the interpretation inferred from potential field data, which explains bodies V and IV as two different felsic volcanic rocks (Fig. 9a). Therefore, reflector R4 represents the contact within the felsic volcanic rocks of the Skellefte Group (R4 and contact V-IV; Fig. 9a).

Although bodies VIII, IX, and X do not coincide completely with the MT data (Fig. 7a), they reveal new interpretation which improves the result of potential field modeling. Interestingly, the vicinity of the northern resistor RC to the surface in Profile C1 can signify that body IX, representing the Vargfors basin, has probably a shallower depth extent (~0.8 km) compared to the one (~1 km) inferred from potential field data (Tavakoli *et al.* 2012b) and has its maximum depth in its SW (compare IX in Figs. 7a and 9a). Conductor CTIII, which was interpreted as alteration zone within the Skellefte Group felsic volcanic rocks, is in agreement with the model based on potential field data, as body V4 was interpreted to be a similar feature in the study by Tavakoli *et al.* (2012b) (see CTIII in Fig. 9a and V4 in Fig. 8).

García Juanatey (2012) relates the resistor RN to the basaltic break-back faults which coincide with reflector R5. The faulted contact along the NE parts of Profile C1 (contact XI-XII; Fig. 9a) fits well to the potential field, seismic reflection and MT data. This is indicated by reflector R6 which images the faulted contact between bodies XI and XII (Fig. 9a).

Profile C2

The southern contact of body I fits nicely with resistivity variation near conductor CTV (10-1000 Ωm; Fig. 9b), which García Juanatey (2012) relates to either an alteration zone, or conductive faults associated with graphite/ sulphide. García Juanatey (2012) relates resistor RIII to the TIB type granitoid-syenitoid which partly fits to the model suggested based on interpretation of the potential field data and explains body II as TIB granitoid-syenitoid (Fig. 7b). The geometry of the bodies I and II was therefore slightly modified to make a better fit with both MT and potential field models (Fig. 9b). Due to the lack of seismic reflectors, the seismic data cannot verify or reject this interpretation.

García Juanatey (2012) relates the deep conductor CTIV to either altered felsic volcanic rocks which embed the ore or graphite within the fault. Conductor CTIV in Profile C2 coincides with a part of body V (at CDP 750-1000), and its upper parts are bounded by resistor RC (interpreted as unaltered felsic volcanic rocks) and reflectors R15 and R8, which are interpreted as fault controlled reflectors (Dehghannejad *et al.* 2012a). However, the considerably low resistivity of CTIV between 2.5 and 5 km depth (Fig. 9b) is somewhat too low to explain even strongly altered felsic volcanic rocks, and CTIV is too large to be related to the hosting rocks of the sulphide mineralization. Hence we suggest that CTIV probably images the sedimentary rocks of the Bothnian Basin with high content of graphite, which is regarded as a basement for hydrothermally altered felsic volcanic rocks of the Skellefte Group (Weihed *et al.* 2002, Skyttä *et al.* 2012).

Geometry of the Vargfors basin (body VIII; Fig. 7b) could not be constrained or modeled either with MT or with the seismic reflection data. However, the resistivities ranging from 1 000 Ωm (CDP 950; Fig. 9b) to 10 000 Ωm (CDP 110; Fig. 9b) support the model suggested by Tavakoli *et al.* (2012b). Resistor RII in the NE, fits nicely with the model indicating contact X-XI which depicts the boundary between the Skellefte group basalt and the Jörn granodiorite inferred from potential field interpretations (Tavakoli *et al.* 2012b) and seismic reflection data (reflector R6; Dehghannejad *et al.*

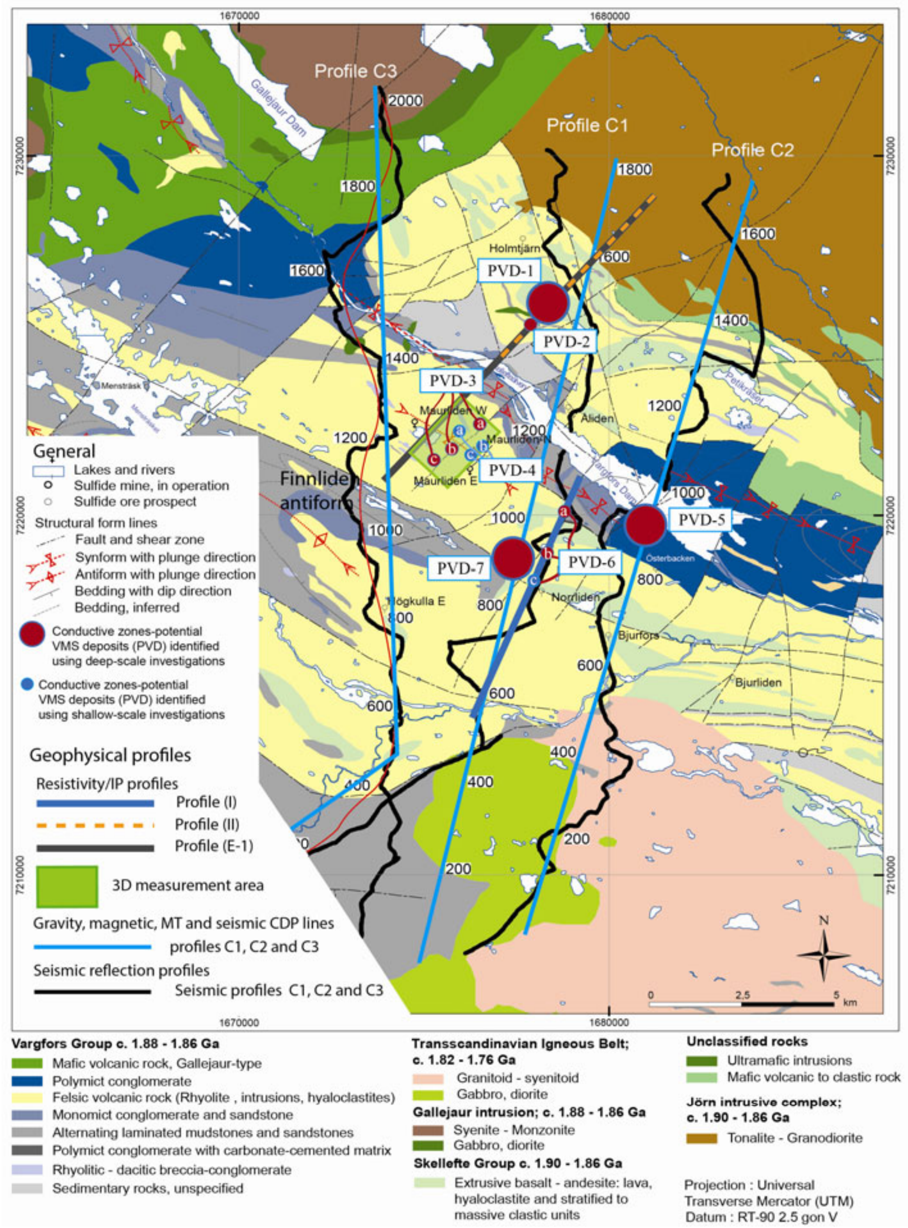


Fig. 10. Distribution of the conductive zones/PVD zones inferred from shallow and deep-scale investigations of the geophysical data in the CSD (modified after Bauer 2010).

2012a). This interpretation is also consistent with the result of deep resistivity/IP investigation for the neighboring Profile E-1, where a resistivity similar to that of RII is observed for the Jörn intrusion and its underlying basalt/unaltered granodiorite (see X-XI in Figs. 9b and 4).

Profile C3

Reflector R1 which Dehghannejad *et al.* (2012a) explained as a product of inverted normal fault, and which isolates the sandstone-mudstone and their underlying rocks (Tavakoli *et al.* 2012b), terminates before reaching the conductive zone CTV west (Fig. 7c). Hence CTV west either images the faults within the laminated sandstone or the shear zone (DNSZ; Fig. 1). Frequent intercalation of the highly magnetized basalts along the laminated sandstone-mudstone was previously explained by Tavakoli *et al.* (2012b) which agrees well with the high resistivity of these basalts inferred from MT interpretation (*e.g.*, body II; Fig. 9c). In addition to the shallow basalts, bodies IV and V, interpreted as deep and sub-vertical basaltic structures, coincide with the high resistivity zone (10 000 Ωm) observed on MT data (at CDP 500-800; Fig. 9c). Unspecified felsic volcanic rocks along Profile C3 (VI; Fig. 9c) which indicate a lower degree of alteration than rhyolites, are bounded with reflectors R3 and R40 and diffraction D3 (Tavakoli *et al.* 2012b). Reflector R11 fits well with the contact VIII-VI which bounds the eastern side of Vargfors basin.

The NE parts of Profile C3 (CDPs 1750-2000) is associated with the TIB-type Gabbro (XIII) and syenite-monzonite (XIV) belonging to the Gallejaur intrusion, which overly the deeply extended Skellefte Group basalts (XII). While this part of Profile C3 is seismically transparent, these intrusive rocks were modeled with the potential field data (Tavakoli *et al.* 2012b). Although XIII and XIV are probably too small to be modeled with MT interpretation, MT result imaged the XII-III boundary (Fig. 9c) which coincides with a group of gently north-dipping reflectors (R14; Dehghannejad *et al.* 2012a).

Several sites were identified as features related to the potential VMS deposits (PVD) based on integrated geophysical and geological interpretations in near-surface and deep-scales. Although these zones were identified based on the modeling of the different datasets, and hence have different certainty, they have great potential for prospecting sulphide mineralization (Fig. 10).

6. CONCLUSION

Multiple geophysical data including magnetic, gravity, resistivity, IP, MT and seismic reflection data were integrated in the central Skellefte district to confirm or improve previous interpretations. The main aim of this study was thus to understand the key lithological features in the two scales of shallow, *i.e.*, down to 1.5 km and deep, *i.e.*, down to 4.5 km in depth.

The result from near-surface investigations confirms that the sulphide mineralization often occurs within the lowermost contact of the Vargfors basin and uppermost part of the Skellefte Group felsic volcanic rocks. The joint interpretation of the 2D-profiles I, II, and E-1 have benefited from potential field modeling along Profile E-1, which adds new insights about the geometry of the area around the Vargfors basin. Several key lithological contacts which were identified in previous studies, *e.g.*, the Jörn-Skellefte contact and the contact between the Skellefte Group-Vargfors Group, were validated in this study. In addition, the basaltic dykes which were inferred from resistivity/IP models fit nicely with the new magnetic model along Profile E-1. Almost all S IP anomalies which were explained as features related to the sulphide mineralization were associated with the magnetic highs which support earlier interpretation for Profile E-1. The deeper resistivity/IP image of Profile E-1 revealed that the anomaly observed on IP model of the Profile II is not an artifact and can depict sulphide mineralization.

Results from deep-investigations revealed new information from greater depths. Joint interpretation of the seismic reflection, potential field and MT data particularly helped to explain several enhanced conductive areas as well as high resistivity zones, which were not detected by seismic and potential field data. Majority of the shallow basalts coincided nicely with the high resistivities observed on MT data; however, some interpretations proved contradictory, which lead to modifying the earlier models and proposed a better model fit consistent with all three datasets. The enhanced conductivities inferred from MT data and confirmed by seismic and/or potential field data suggested new horizons for targeting the sulphide mineralization. The failure in detecting small-scale known mineralization with the MT data, which were detected using high resolution resistivity/IP studies, implies the significance of the multi-scale interpretations which can be determined depending on the aim and scope of studies. The near-surface and deep-scale models of the central Skellefte district are therefore of great importance for further exploration activities in the area. Seven zones with high probability of features related to the potential VMS deposits (PVD zones) were identified within the top 4 km of the crust. These zones, although initially identified based on resistivity, IP or MT data, correlate at least to another geophysical method, *i.e.*, seismic reflection and potential field data in deep-scale and potential field data in shallow-scale which makes these anomalies interesting zones for future focus of the exploration activities in the area.

Acknowledgement. This study is conducted as a part of the VINNOVA 4D modeling project and is financed by VINNOVA and Boliden

Mineral AB. All 4D-modeling members are thanked for their contribution, provision of data and their comments on this manuscript. We also thank Geovista AB for their contributions to the field work and data processing. The authors would like to thank editors of the *Acta Geophysica* journal and two anonymous reviewers for their constructive comments.

References

Allen, R.L., P. Weihed, and S.-Å. Svenson (1996), Setting of Zn–Cu–Au–Ag massive 1256 sulfide deposits in the evolution and facies architecture of a 1.9 Ga marine volcanic 1257 arc, Skellefte District, Sweden, *Econ. Geol.* **91**, 6, 1022-1053, DOI: 10.2113/gsecongeo.91.6.1022.

Bauer, T. (2010), Structural and sedimentological reconstruction of the inverted Vargfors Basin – A base for 4D-modelling, Licentiate Thesis, Luleå University of Technology, Sweden, 44 pp.

Bauer, T.E., P. Skyttä, R.L. Allen, and P. Weihed (2011), Syn-extensional faulting controlling structural inversion – Insights from the Palaeoproterozoic Vargfors syncline, Skellefte mining district, Sweden, *Precambrian Res.* **191**, 3-4, 166-183, DOI: 10.1016/j.precamres.2011.09.014.

Bauer, T., P. Skyttä, R. Allen, and P. Weihed (2013), Fault-controlled sedimentation in a progressively opening extensional basin: the Palaeoproterozoic Vargfors basin, Skellefte mining district, Sweden, *Int. J. Earth Sci.* **102**, 2, 385-400, DOI: 10.1007/s00531-012-0808-x.

Bellefleur, G., E. Schetselaar, D. White, K. Miah, and P. Dueck (2015), 3D seismic imaging of the Lalor volcanogenic massive sulphide deposit, Manitoba, Canada, *Geophys. Prospect.* **63**, 4, 813-832, DOI: 10.1111/1365-2478.12236.

Bergman Weihed, J. (2001), Palaeoproterozoic deformation zones in the Skellefte and Arvidsjaur areas, northern Sweden. **In:** P. Weihed (ed.), *Economic Geology Research*, Vol. 1, Geological Survey of Sweden, C 833, 46-68.

Billström, K., and P. Weihed (1996), Age and provenance of host rocks and ores in the Paleoproterozoic Skellefte district, northern Sweden, *Econ. Geol.* **91**, 6, 1054-1072, DOI: 10.2113/gsecongeo.91.6.1054.

Dehghannejad, M., C. Juhlin, A. Malehmir, P. Skyttä, and P. Weihed (2010), Reflection seismic imaging of the upper crust in the Kristineberg mining area, northern Sweden, *J. Appl. Geophys.* **71**, 4, 125-136, DOI: 10.1016/j.jappgeo.2010.06.002.

Dehghannejad, M., T.E. Bauer, A. Malehmir, C. Juhlin, and P. Weihed (2012a), Crustal geometry of the central Skellefte district, northern Sweden – constraints from reflection seismic investigations, *Tectonophysics* **20**, 87-99, DOI: 10.1016/j.tecto.2011.12.021.

Dehghannejad, M., A. Malehmir, C. Juhlin, and P. Skyttä (2012b), 3D constraints and finite difference modeling of massive sulfide deposits: The Kristineberg seismic lines revisited, northern Sweden, *Geophysics* **77**, 5, 69-79, DOI: 10.1190/geo2011-0466.1.

García Juanatey, M.A. (2012), Seismics, 2D and 3D inversion of magnetotellurics : Jigsaw pieces in understanding the Skellefte Ore District, Ph.D. Thesis, Uppsala University, Uppsala, Sweden.

Hedin, P., A. Malehmir, D.G. Gee, C. Juhlin, and D. Dyrelius (2013). 3D interpretation by integrating seismic and potential field data in the vicinity of the proposed COSC-1 drill site, central Swedish Caledonides, *Geol. Soc. London Spec. Pub*. **390**, 301-319, DOI: 10.1144/SP390.15.

Hübert, J., M. García Juanatey, A. Malehmir, A. Tryggavson, and L. Pedersen (2013), The upper crustal 3-D resistivity structure of the Kristineberg area, Skellefte district, northern Sweden revealed by magnetotelluric data, *Geophys. J. Int.* **192**, 2, 500-513, DOI: 10.1093/gji/ggs044.

Juhlin, C., M. Dehghannejad, B. Lund, A. Malehmir, and G. Pratt (2010), Reflection seismic imaging of the end-glacial Pärvie Fault system, northern Sweden, *J. Appl. Geophys*. **70**, 4, 307-316, DOI: 10.1016/j.jappgeo.2009.06.004.

Kathol, B., and P. Weihed (eds.) (2005), *Description to Regional Geological and Geophysical Maps of the Skellefte District and Surrounding Areas*, Sveriges Geologiska Undersökning, Ser. Ba, 197 pp.

Loke, M.H. (2012), Tutorial: 2-D and 3-D electrical imaging surveys, Geotomo Software, Malaysia.

Malehmir, A., P. Dahlin, E. Lundberg, C. Juhlin, H. Sjöström, and K. Högdahl (2011), Reflection seismic investigations in the Dannemora area, central Sweden: insights into the geometry of poly-phase deformation zones and magnetite-skarn deposits, *J. Geophys. Res*. **116**, B11, B11307, DOI: 10.1029/2011JB008643.

Montelius, C., R.L. Allen, S.-Å. Svenson, and P. Weihed (2007), Facies architecture of the Palaeoproterozoic VMS-bearing Maurliden volcanic centre, Skellefte district, Sweden, *GFF* **129**, 3, 177-196.

Moorkamp, M., B. Heincke, M. Jegen, A.W. Roberts, and R.W. Hobbs (2011), A framework for 3-D joint inversion of MT, gravity and seismic refraction data, *Geophys. J. Int.* **184**, 1, 477-493, DOI: 10.1111/j.1365-246X.2010.04856.x.

Nyquist, J.E., and M.J.S. Roth (2005), Improved 3D pole-dipole resistivity surveys using radial measurement pairs, *Geophys. Res. Lett*. **32**, 21, L21416, DOI: 0.1029/2005GL024153.

Represas, P., F.A. Monteiro Santos, J.A. Ribeiro, A. Andrade Afonso, J. Ribeiro, M. Moreira, and L.A. Mendes-Victor (2015), On the applicability of joint inversion of gravity and resistivity data to the study of a tectonic sedimentary basin in Northern Portugal, *Pure Appl. Geophys*. **172**, 10, 2681-2699, DOI: 10.1007/s00024-014-0920-x.

Skyttä, P. (2012), Crustal evolution of an ore district illustrated - 4D-animation from the Skellefte district, Sweden, *Comput. Geosci*. **48**, 157-161, DOI: 10.1016/j.cageo.2012.05.029.

Skyttä, P., T.E. Bauer, S. Tavakoli, T. Hermansson, J. Andersson, and P. Weihed (2012), Pre-1.87 Ga development of crustal domains overprinted by 1.87 Ga transpression in the Palaeproterozoic Skellefte District, Sweden, *Precambrian Res*. **206-207**, 109-136, DOI: 10.1016/j.precamres.2012.02.022.

Smirnov, M. (2003), Magnetotelluric data processing with a robust statistical procedure having a high breakdown point, *Geophys. J. Int*. **152**, 1, 1-7, DOI: 10.1046/j.1365-246X.2003.01733.x.

Tavakoli, S., S-Å. Elming, and H. Thunehed (2012a), Geophysical modelling of the central Skellefte district, Northern Sweden; an integrated model based on the electrical, potential field and petrophysical data, *Appl. Geophys*. **82**, 84-100, DOI: 10.1016/j.jappgeo.2012.02.006.

Tavakoli, S., T.E. Bauer, S.-Å. Elming, H. Thunehed, and P. Weihed (2012b), Regional-scale geometry of the central Skellefte district, northern Sweden-results from 2.5D potential field modeling along three previously acquired seismic profiles, *Appl. Geophys*. **85**, 43-58, DOI: 10.1016/j.jappgeo.2012.06.012.

Tavakoli, S., T.E. Bauer, T.M. Rasmussen, P. Weihed, and S.-A. Elming (2016), Deep massive sulphide exploration using 2D and 3D geoelectrical and induced polarization data in Skellefte mining district, northern Sweden, *Geophys. Prospect*., DOI: 10.1111/1365-2478.12363.

Weihed, P. (2010), Palaeoproterozoic mineralized volcanic arc systems and tectonic evolution of the Fennoscandian shield: Skellefte district Sweden, *GFF* **132**, 1, 83-91.

Weihed, P. (2015), *3D, 4D and Predictive Modelling of Major Mineral Belts in Europe*, Mineral Resource Reviews, Springer, 331 pp.

Weihed, P., K. Billstrom, P. Persson, and J. Weihed (2002), Relationship between 1.90-1.85 Ga accretionary processes and 1.82-1.80 Ga oblique subduction at the Karelian craton margin, Fennoscandian Shield, *J. Geol. Soc. Sweden* **124**, 3, 163-180.

Received 21 August 2015
Received in revised form 23 November 2015
Accepted 27 January 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 4, Dec. 2016, pp. 2200-2213
DOI: 10.1515/acgeo-2016-0100

Application of Electrical Resistivity Imaging for Engineering Site Investigation. A Case Study on Prospective Hospital Site, Varamin, Iran

Amin AMINI and Hamidreza RAMAZI

Department of Mining and Metallurgical Engineering,
Amirkabir University of Technology, Tehran, Iran;
e-mail: ramazi@aut.ac.ir

Abstract

The article addresses the application of electrical resistivity imaging for engineering site investigation in Pishva Hospital, Varamin, Iran. Some aqueduct shafts exist in the study area backfilled by loose materials. The goals of this study are to detect probable aqueduct tunnels and their depth, investigate filling quality in the shafts as well as connection(s) between them. Therefore, three profiles were surveyed by dipole-dipole electrode array. Also, to investigate the potentially anomalous areas more accurately, five additional resistivity profiles were measured by a Combined Resistivity Sounding-Profiling array (CRSP). According to the results of 2-D inversion modelling, a main aqueduct tunnel was detected beneath the central part of the site. Finally, the resistivity pattern of the detected aqueduct system passing the investigated area was provided using the obtained results.

Key words: electrical resistivity imaging, cavity detection, CRSP array, Qanat.

Ownership: Institute of Geophysics, Polish Academy of Sciences;

1. INTRODUCTION

Civil engineering projects are usually dealing with geological and geotechnical problems. Conducting these projects in urban areas has several critical and special problems (Dindarloo and Siami-Irdemoosa 2015, Fang *et al.* 2016). Lack of scientific attention to them can cause noticeable damages or even disasters (*e.g.*, BBC News 2015). Channels, underground cavities, wells, aqueducts and other obsolete man-made underground structures which may be buried now or in the near future are among the most crucial geotechnical and geological problems in urban areas (Fig. 1). The underground void spaces of this kind (empty or filled by loose materials) usually cause weak points in the ground and, therefore, there is an urgent need to identify, map and compile them. Various methods have been used to study the abovementioned phenomena. Geophysical methods are considered as the fastest and most affordable of them (Benson 1995, Burger 1992, Cosenza *et al.* 2006, Gautam *et al.* 2000).

Electrical resistivity method has been widely used due to low operation costs and its flexibility for many geological conditions (Van Schoor 2002, Zhou *et al.* 2002). The purpose of the resistivity method is to detect underground inhomogeneities and interpret them as modifications in underground materials or structures. Resistivity method is widely used in fields such as mineral exploration, environmental investigations, geotechnical studies, hydrology and hydrogeology (*e.g.*, Asfahani and Abou Zakhem 2013, Bayrak

Fig. 1. Surface settlement of Navab Str., Tehran, Iran, due to an obsolete Qanat tunnel collapse (Donya-e eqtesad 2009).

and Senel 2012, Candansayar and Başokur 2001, Fehdi *et al.* 2011, Giang *et al.* 2013, Hee *et al.* 2010, Loke and Barker 1996, Narayan *et al.* 1994, Papadopoulos *et al.* 2007, Wilkinson *et al.* 2010, Winters *et al.* 2015).

In Iran, systems of underground aqueducts called Qanats were constructed in past millenniums, including a series of well-like vertical shafts, connected by gently sloping tunnels. This technique taps into subterranean water in a manner that delivers water to the surface without the need for pumping (Fig. 2). Most of the old Qanats are obsolete nowadays, but their structures still exist (Semsar Yazdi and Askarzadeh 2007, ICQHS 2015).

In this paper a case study is discussed to demonstrate the application of the electrical resistivity method to investigate foundation of a prospective hospital site to detect probable weak zones and/ or cavities due to an abandoned Qanat system.

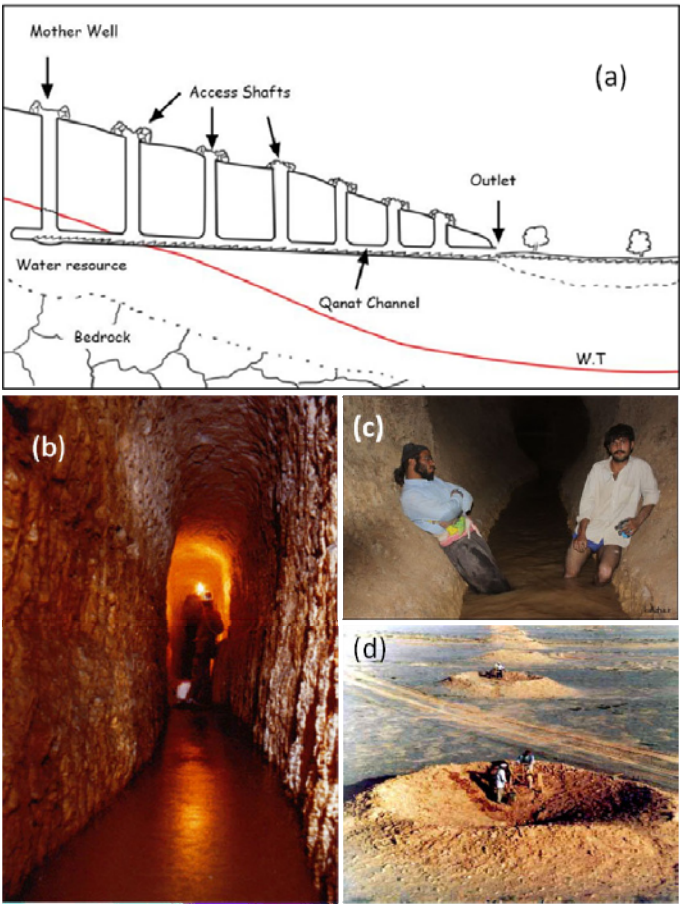


Fig. 2: (a) Schematic section showing the Qanat system; (b, c, and d) typical pictures showing Qanats features.

2. STUDY AREA

The Pishva Hospital site is located in Pishva town, 70 km southeast of Tehran, the capital city of Iran (Fig. 3). Pishva town is located on a large plain which is dominantly covered by agricultural soils with thicknesses ranging from 1 m to more than 5 m laid on young fine grain alluviums. Previous geotechnical investigation at the area reported that the thickness of the alluvium layer changes from 8 m in the north to over 18 m in the southern part of the town. Although the second layer is partially formed of a few sub-layers with different thicknesses, they can be considered together as a single geoelectrical unit. The alluvium layer is laid on an impermeable silty layer. Pishva town has been constructed on old farmlands, so interestingly but not surprisingly lots of Qanat shafts still exist in the town, as expected. The main field information is the following: Previous investigations on Qanats show that ancient people dug the Qanats tunnels into the silt layer in this area, and several abandoned Qanat shafts traces could be observed at the Pishva Hospital site, suggesting that a Qanat system exists beneath the area. Most of the shafts were backfilled by people because of safety problems. Part of the study area has been levelled by municipality contractors (see Fig. 3). Based on local habitants' sayings there should be seven aqueduct shafts in the area.

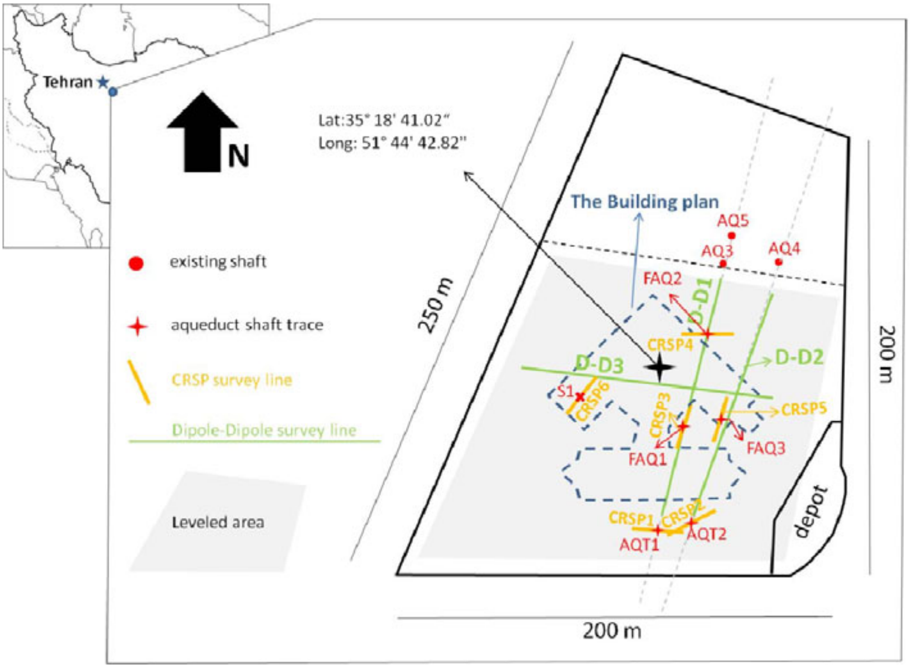


Fig. 3. Plan view of Pishva Hospital site. Locations of the survey lines are identified in the figure.

The existing shafts are located along two more-or-less parallel lines with a north-south trend, so we supposed that two parallel Qanat tunnels connect N-S trending contiguous shafts. For example, we assumed that shafts AQ5, AQ3, and AQT1 are connected together as well as shafts AQ4 and AQT2 through parallel tunnels.

3. SURVEY DESIGNING AND DATA ACQUISITION

In this study, we used a CRSP array (Ramazi 2005) to conduct detailed electrical resistivity imaging (ERI) surveys. The data obtained from this array could be applied not only for ERI but also for vertical electrical soundings (VES) interpretation. The main purpose of the application of this array is to compile ERI sections. By CRSP array, three VES are constructed simultaneously by a set of measurements. In this array the distance of measuring stations is equal to the spacing of the potential electrodes (Fig. 4). As shown in the figure, CRSP could be similar to the Schlumberger and Wenner-Schlumberger arrays in central measurements; however, potential electrode spacing could be decreased for shorter current electrode distances. The CRSP provides an enhanced lateral and in depth coverage, due to the acquisition of more data in a section (Fig. 4). The CRSP has been successfully applied to different mineral exploration and engineering site investigations (see Amini and Ramazi 2016, Ramazi and Jalali 2014, and Ramazi and Mostafaie 2013, for further details).

Designing of the survey lines was done in four steps as follows:

- Two CRSP profiles, with a maximum current length of 150 m and electrode spacing of 3 m, were surveyed in the locations of AQT1 and AQT2, mainly to study the backfilled shafts and their filling quality.

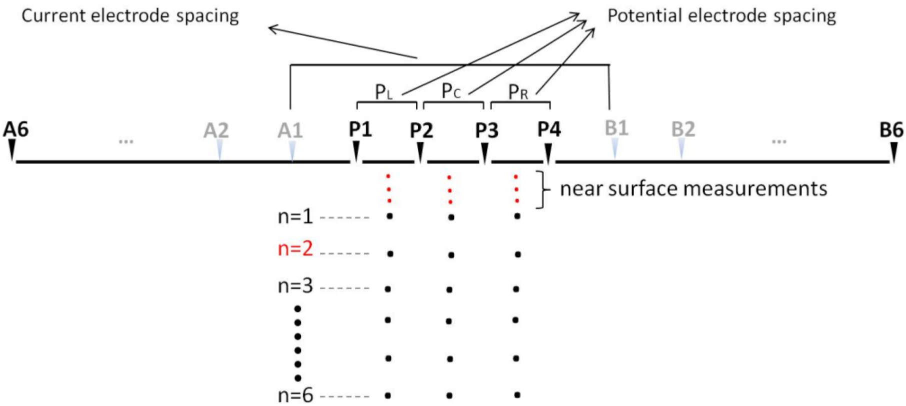


Fig. 4. Schematic section along CRSP electrode array.

- Three dipole-dipole profiles were conducted to have data coverage of the study area. Two parallel conventional inline surface dipole-dipole profiles of 110 m in length, dipole spacing of 5 m and a maximum dipole length of 80 m ($n = 16$) were conducted in a direction of north-south. Also, to detect other probable existing aqueduct tunnels, a dipole-dipole profile with a length of 110 m and dipole spacing of 5 m was designed in the west-east direction.
- Three CRSP profiles with a maximum current length of 150 m and electrode spacing of 3 m were conducted for obtaining more detailed data about the anomalies found in the previous step. The CRSP profiles were carried out on the anomalies locations (CRSP3, 4, and 5).
- A CRSP profile, with a maximum current length of 150 m and electrode spacing of 3 m, was carried out in location of S1, to obtain a sequence of geological layers. Based on the observations, this location was expected to be fresh and geological layers should be normal.

The data acquisition was carried out by a WDJD-3 RS-IP measurement system using single channel mode. All measurements were repeated for three to five times to ensure data quality. The acquired apparent resistivity data were inverted by the RES2DINV (commercially available) software Ver. 3.54 and the results were drawn into 2D resistivity sections using the SURFER software Ver. 10. All profiles were surveyed in the flattened area, so no topographic correction was needed to be applied on the acquired data (Fig. 3). Inversion parameters were selected so that the best image of the subsurface would be achieved. Finite differences method was applied with finest grid option to carry out forward models; the method of inversion was set as combined inversion method (*i.e.*, the combination of Marquardt and Occam inversion) and the starting model was set as the average of the measured apparent resistivity data).

4. RESULTS AND DISCUSSION

Figures 5 to 8 show the resistivity inverted images along surveyed profiles. From the inverted resistivity images, three resistivity layers were overall distinguished (in the all dipole-dipole and CRSP images and also in the VES curve obtained from CRSP6 measurements), which are more or less coincided with geological layers in the area: a surficial resistivity layer of 45-70 ohm-m consisting top soils and loose materials and having a thicknesses of 2-5 m; an intermediate resistivity layer of 15-40 ohm-m is inferred to be fine grained alluviums with an average thickness of 12-13 m; and a deeper third resistivity layer of 5-15 ohm-m that seems to relate to the basal silty layer (as the bedrock). Along the dipole-dipole profiles, three high resistivity anomalies were detected. The detected anomalies could be related to the

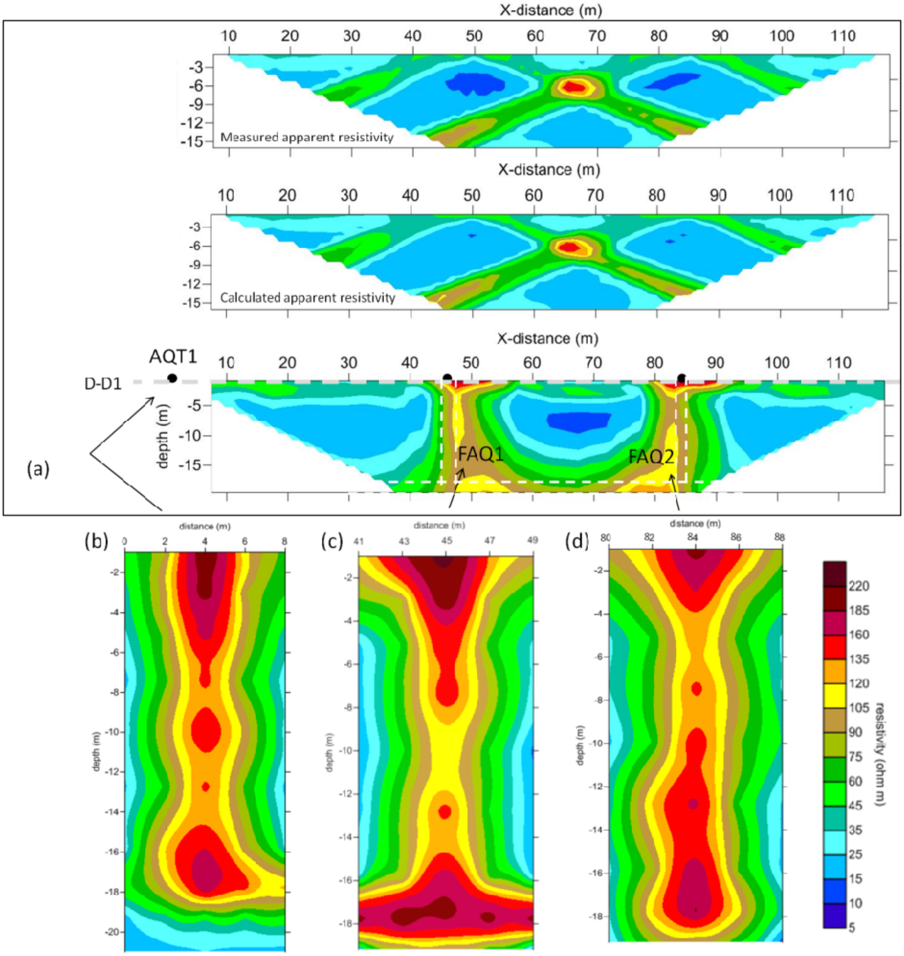


Fig. 5: (a) Inverted resistivity image along survey line D-D1. Pseudo-sections of measured and calculated apparent resistivity data (iteration 5) are also included in the figure. The RMS error was calculated as 0.07; (b) Inverted resistivity image along survey line CRSP1. This image shows the filling quality of the shaft AQ1; (c) Inverted resistivity image along survey line CRSP3; (d) Inverted resistivity section along the survey line CRSP4. The direction of the profile line is perpendicular to D-D1.

Qanat shafts. Figure 5a shows the inverted resistivity image along the survey line D-D1. Two high resistivity anomalies (FAQ1 and FAQ2) are observed in the image which seems to be due to hidden backfilled Qanat shafts. As seen in Fig. 5a, a tunnel passes along the survey line orientation. The depth of the tunnel is estimated at almost 17 to 18 m. It could be inferred that at least Qanat shafts AQT1, AQ3, AQ5 and two other shafts found are connected through the detected tunnel. Figure 5b shows the inverted resistivity image along survey line CRSP1. Because of constrains with southern fence of the site, we decided to conduct CRSP1 in a direction of west-east so that the probable tunnel connecting AQT1 to AQT2 could be detected. The AQ1 filling quality could be followed through CRSP1 inverted resistivity image. The length of the shaft is estimated to be a little more than 18 m. Figure 5c illustrates the inverted resistivity image along the survey line CRSP3 carried out on the location of FAQ1. The detected tunnel in profile D-D1 is also observed. A local partial tunnel collapse has made a considerable cavity at the intersection of the shaft and the tunnel. Figure 5d shows the inverted resistivity image along the survey line CRSP4 which was conducted to obtain a more detailed image on found anomaly FAQ2. Through this section, shaft FAQ2 filling quality can be followed. The length of the shaft is estimated to be about 18 m. The rounded section of the bottom of the shaft confirms that the connected tunnel to the FAQ2 is perpendicular to the survey line direction.

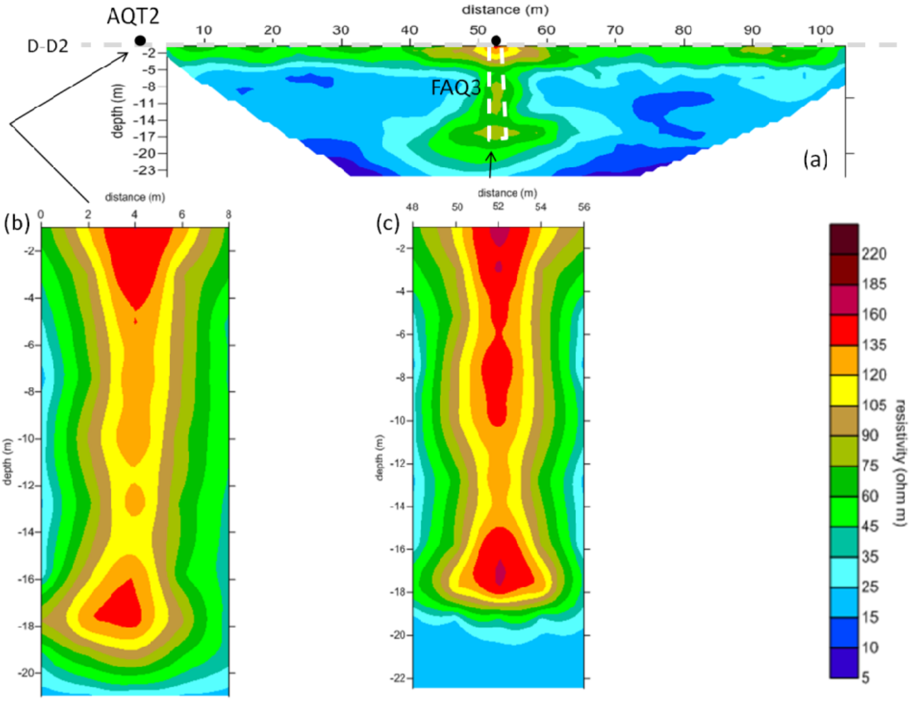


Fig. 6: (a) Inverted resistivity image along survey line D-D2. FAQ3 is distinguishable in the image; (b) Inverted resistivity image along survey line CRSP2. Filling quality of shaft AQT2 is seen in this section; (c) Inverted resistivity section along the survey line CRSP5. The direction of the profile line is parallel to D-D2.

Figure 6a shows the inverted resistivity image along survey line D-D2. A relatively high resistivity anomaly (FAQ3) is visible in the section, suggesting a backfilled Qanat shaft's trace but no tunnel could be detectable along the survey line orientation. According to the shape of the anomaly it could be inferred that survey line D-D2 is not exactly over FAQ3 location (Fig. 3); therefore, the intensity of this anomaly is lower than the anomalies found in the survey line D-D1. The inverted resistivity image along survey line CRSP2 is shown by Fig. 6b. The direction of CRSP2 is such that the tunnel connecting the AQT2 to the AQT1 could be detected. The length of shaft AQ2 is estimated at about 18 m. We conducted a CRSP survey to obtain a more detailed image on the detected anomaly by D-D2 profile (Fig. 6c). Through this section, the FAQ3 filling quality can be followed. The length of the shaft FAQ3 is estimated to be 18 m. The shape of the shaft's bottom confirms the connected tunnel to shaft FAQ3 is perpendicular to the survey line CRSP5; considering the observation and detected shaft FAQ1 along survey line D-D1, FAQ3 seems to be connected to FAQ1.

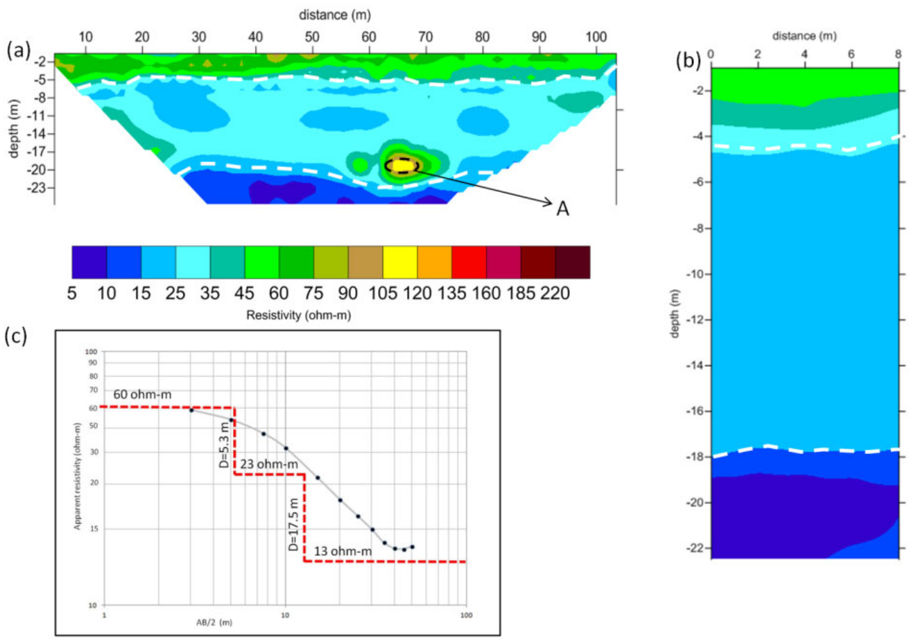


Fig. 7: (a) Inverted resistivity image along survey line D-D3. Three main resistivity layers are distinguishable in the image. The anomaly remarked with "A" belongs to a potential connection tunnel; (b) Inverted resistivity image along the survey line CRSP6. In conjunction with the inverted resistivity image along the survey line D-D2, the main resistivity layers could be also distinguished; (c) 1-D resistivity inversion along CRSP5-c sounding curve.

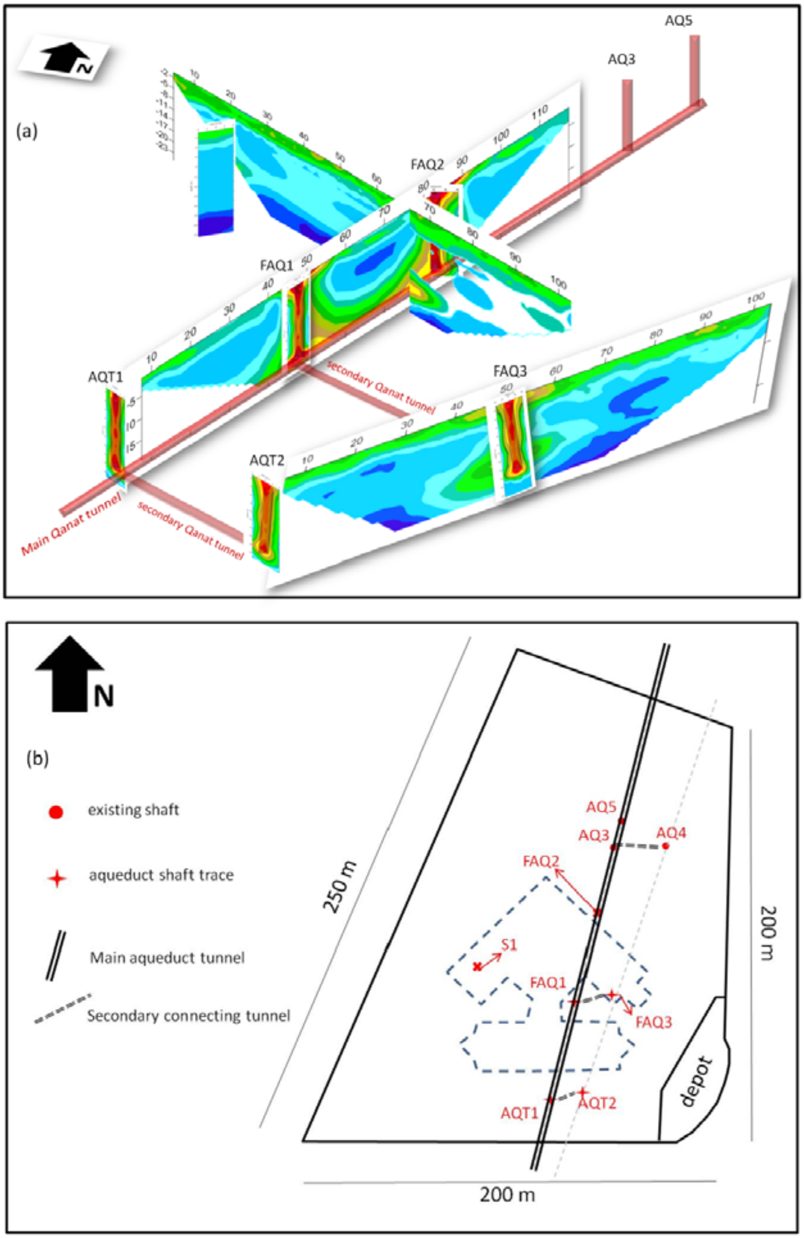


Fig. 8: (a) Quasi-3D schema of the aqueduct system derived from inverted resistivity images; (b) resulted plan of the aqueduct system paths. The main tunnel passes the shafts AQT1, FAQ1, FAQ2, AQ3, and AQ5. Shafts AQT2, FAQ3, and AQ4 are connected to shafts AQT1, FAQ1, and AQ3, respectively, through secondary tunnels.

Figure 7a presents the inverted resistivity image along survey line D-D3. A high resistivity anomaly of 100-120 ohm-m is marked in the figure with letter “A”, which seems to be the effect of a passing tunnel across the survey line. Considering the location of the detected anomaly, it should be the trace of detected tunnel in survey line D-D1. Survey line CRSP6 was conducted to explore a sequence of geological layers of the area. As expected, no important anomalies could be detected in this section. We also carried out 1-D inversion on sounding data along this profile (Fig. 7c). The results confirm three main resistivity layers in this section.

Finally, with regard to the information gained by resistivity imaging along the survey lines, the plan of the abandoned aqueduct paths was drawn. The main tunnel of the Qanat system is located in the central part of the hospital site and its orientation is almost north-south. The shafts located in the eastern side of the area are connected to the main tunnel through secondary tunnels with almost west-east orientation (Fig. 8).

5. CONCLUSION

The main goal of this article was investigating probable weak zones and/or cavities, at the of Pishva Hospital site, by using 2D electrical resistivity imaging. Our results indicate that the conducted method could detect the objectives with an acceptable precision. Three main resistivity layers were distinguished in the resistivity images D-D3 (as well as in inversion image CRSP6), which are in agreement with the local stratigraphy. The results arisen from the dipole-dipole profiles provided a satisfying coverage of the investigation area so that the main Qanat tunnel was detected as a resistivity anomaly in the resistivity section D-D1, as well as the aqueduct shafts traces FAQ1 and FAQ2. The effect of shaft FAQ3 was also detected through D-D2 profile. In conjunction with the dipole-dipole inverted resistivity images, the CRSP could provide detailed information of the aqueduct shafts so that cavities and/or weak points along the shafts fillings could be observed in the resulting inverted resistivity images, due to their anomalous resistivity signature.

References

Amini, A., and H.R. Ramazi (2016), Anomaly enhancement in 2-D Electrical resistivity imaging method using residual resistivity technique, *J. Southern Afr. Inst. Min. Metall.* **116**, 2, 161-168, DOI: 10.17159/2411-9717/2016/v116n2a7.

Asfahani, J., and B. Abou Zakhem (2013), Geoelectrical and hydrochemical investigations for characterizing the salt water intrusion in the Khanasser valley, northern Syria, *Acta Geophys*. **61**, 2, 422-444, DOI: 10.2478/s11600-012-0071-3.

Bayrak, M., and L. Senel (2012), Two-dimensional resistivity imaging in the Kestelek boron area by VLF and DC resistivity methods, *J. Appl. Geophys.* **82**, 1-10, DOI: 10.1016/0926-9851(95)90040-3.

BBC News (2015), Grays sinkhole caused by quarry tunnel collapse, say experts, BBC News, available from: http://www.bbc.com/news/uk-england-essex-32921411(accessed: 11 November 2015).

Benson, A.K. (1995), Applications of ground penetrating radar in assessing some geological hazards: examples of groundwater contamination, faults, cavities, *J. Appl. Geophys.* **33**, 1-3, 177-193, DOI: 10.1016/0926-9851(95) 90040-3.

Burger, H.R. (1992), *Exploration Geophysics of the Shallow Subsurface*, Prentice-Hall, Englewood Cliffs, 489 pp.

Candansayar, M.E., and A.T. Basokur (2001), Detecting small scale targets by the 2D inversion of two-sided three- electrode data: application to an archaeological survey, *Geophys. Prospect*. **49**, 1, 13-25, DOI: 10.1046/j.1365-2478.2001.00233.x.

Cosenza, P., E. Marmet, F. Rejiba, Y.J. Cui, A. Tabbagh, and Y.Charlery (2006), Correlations between geotechnical and electrical data: A case study at Garchy in France, *J. Appl. Geophys*. **60**, 3-4, 165-178, DOI: 10.1016/j.jappgeo. 2006.02.003.

Dindarloo, S., and E. Siami-Irdemoosa (2015), Maximum surface settlement based classification of shallow tunnels in soft ground, *Tunn. Undergr. Sp. Tech.* **49**, 320-327, 320-327, DOI: 10.1016/j.tust.2015.04.021.

Donya-e eqtesad (2009), Obsolete Qanats caused Tohid tunnel collapse, Donya-e eqtesad newspaper, available from: http://www.donya-e-eqtesad.com/news/ 568474 (accessed: 10 November 2015).

Fang, Q., Q. Tai, D. Zhang, and L. Wong (2016), Ground surface settlements due to construction of closely-spaced twin tunnels with different geometric arrangements, *Tunn. Undergr. Sp. Tech*. **51**, 144-151, DOI: 10.1016/j.tust. 2015.10.031, 144-151.

Fehdi, C., F. Baali, D. Boubaya, and A. Rouabhia (2011), Detection of sinkholes using 2D electrical resistivity imaging in the Cheria Basin (north-east of Algeria), *Arab. J. Geosci*. **14**, 1-2, 181-187, DOI: 10.1007/s12517-009-0117-2.

Gautam, P., R.P. Surendra, and A. Hisao (2000), Mapping of subsurface karst structure with gamma ray and electrical resistivity profiles: a case study from Pokhara valley, central Nepal, *J. Appl. Geophys*. **45**, 2, 97-110, DOI: 10.1016/S0926-9851(00)00022-7.

Giang, N.V., N.B. Duan, L. Thanh, and N. Hida (2013), Geophysical techniques to aquifer locating and monitoring for industrial zones in North Hanoi, Vietnam, *Acta Geophys*. **61**, 6, 1573-1597, DOI: 10.2478/s11600-013-0147-8.

Hee, S.H., S.K. Dae, and J.P. Inn (2010), Application of electrical resistivity techniques to detect weakened fracture zones during underground construction, *Environ. Earth Sci*. **60**, 4, 723-731, DOI: 10.1007/s12665-009-0210-6.

ICQHS (2015), ICQHS official website available from: http://icqhs.org/ SC.php?type=staticandid=25 (accessed: 11 November 2015).

Loke, M., and R. Barker (1996), Rapid least squares inversion of apparent resistivity pseudosections using a quasi Newton's method, *Geophys. Prospect*. **44**, 1, 131-152, DOI: 10.1111/j.1365-2478.1996.tb00142.x.

Narayan, S., M.B. Dusseault, and D.C. Nobes (1994), Inversion techniques applied to resistivity inverse problems, *Inverse Probl*. **10**, 3, 669-686, DOI: 10.1088/0266-5611/10/3/011.

Papadopoulos, N.G., P. Tsourlos, G.N. Tsokas, and A. Sarris (2007), Efficient ERT measuring and inversion strategies for 3D imaging of buried antiquities, *Near Surf. Geophys*. **5**, 6, 349-361, DOI: 10.3997/1873-0604.2007017.

Ramazi, H.R. (2005), Combined resistivity sounding and profiling and its application in mineral exploration and site investigation, Technical Report, Tehran, 21 pp. (in Persian).

Ramazi, H.R., and M. Jalali (2014), Contribution of geophysical inversion theory and geostatistical simulation to determine geoelectrical anomalies, *Stud. Geophys. Geod*. **59**, 1, 97-112, DOI: 10.1007/s11200-013-0772-3.

Ramazi, H.R., and K. Mostafaie (2013), Application of integrated geoelectrical methods in Marand (Iran) manganese deposit exploration, *Arab. J. Geosci*. **6**, 8, 2961-2970, DOI: 10.1007/s12517-012-0537-2.

Semsar Yazdi, A., and S. Askarzadeh (2007), A historical review on the Qanats and historic hydraulic structures of Iran since the first millennium B.C., International History Seminar on Irrigation and Drainage, Tehran, Iran.

Van Schoor, M. (2002), Detection of sinkholes using 2D electrical resistivity imaging, *J. Appl. Geophys*. **50**, 4, 393-399, DOI: 10.1016/S0926-9851(02)00166-0.

Wilkinson, P.B., P.I. Meldrumm, O. Kuras, J.E. Chambers, S.J. Holyoake, and R.D. Ogilvy (2010), High-resolution electrical resistivity tomography monitoring of a tracer test in a confined aquifer, *J. Appl. Geophys*. **70**, 4, 268-276, DOI: 10.1016/j.jappgeo.2009.08.001.

Winters, G., I. Ryvkin, T. Rudkov, Z. Moreno, and A. Furman (2015), Mapping underground layers in the super arid Gidron Wadi using electrical resistivity tomography (ERT), *J. Arid Environ*. **121**, 79-83, DOI: 10.1016/j.jaridenv.2015.05.008.

Zhou, W., B.F. Beck, and A.L. Adams (2002), Effective electrode array in mapping karst hazards in electrical resistivity tomography, *Environ. Geol*. **42**, 8, 922-928, DOI: 10.1007/s00254-002-0594-zv.

Received 7 September 2015
Received in revised form 26 January 2016
Accepted 3 February 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2214-2231
DOI: 10.1515/acgeo-2016-0082

Enhancement of Seismic Data Processing and Interpretation of Fracture Zones on the Upper Part of Granitic Basement in Cuu Long Basin, Vietnam

Mai Thanh TAN[1], Mai Thanh HA[2], Kurt J. MARFURT[3], Nguyen Trung HIEU[4], and Nguyen Thi My HANH[2]

[1]Hanoi University of Mining and Geology, Dong Ngac, Tu Liem, Hanoi, Vietnam
e-mail: mttan44@gmail.com

[2]Petrovietnam Exploration and Production Corporation, Tran Duy Hung, Hanoi, Vietnam

[3]The University of Oklahoma, Norman, Oklahoma, USA

[4]Vietnam Petroleum Institute, Trung Kinh, Tu Liem, Hanoi, Vietnam

Abstract

The fractured granite basement is the primary oil and gas reservoir in the Cuu Long Basin, Vietnam. Due to the complexity of this non-layered unconventional target, combined with complicated fault and fracture systems, the seismic data quality near and within the basement section is very low. For this reason, it is important to apply improved seismic data processing workflows, filtering and migration techniques, as wells as attribute processing methods to enhance the imaging quality.

Our studies show that applying different types of filters, including the f-k, Radon transform and Tau-P, improves signal to noise ratio, removing multiples, revealing basement's top and its related fractured and fault zones. In addition, the application of multi-arrival-solution migration algorithms, such as Kirchhoff Migration and Controlled Beam Migration, provides improved imaging for identifying basement top and faults and fractures within the basement. Furthermore, the application of seismic attributes such as curvature, apparent dip, or energy gradient, is important in locating faults and fractures, whereas mapping of intensity and orientation of such structures assists the delineation of “sweet spots” and assists the planning of exploration.

Key words: Cuu Long Basin, fractured basement, faults, filters, migration, imaging, attributes.

1. INTRODUCTION

The Cuu Long Basin is one of major Cenozoic marginal sedimentary basins of the continental shelf of Vietnam. This basin comprises fragmented Mesozoic basement complexes and overlying Cenozoic strata of differing age, origin and spatial distribution. The basement complexes, which compose mostly felsic intrusive and subordinate sedimentary rocks, are commonly strongly fractured, dismembered and partly hydrothermally altered as a consequence of numerous tectonic events. The basement rocks were exposed to weathering before subsidence and covered by Cenozoic sedimentary units during the crustal rifting and basin formation processes. These processes created local significant structural variability such as horst and grabens, a network of fractures with sufficient pore space and permeability within the basement complexes. All of these features have contributed to the formation of a unique type of hydrocarbon reservoirs within the basement granitic rocks of the Cuu Long Basin, which became primary targets for oil and gas exploration in Vietnam.

In the Cuu Long Basin, petroleum exploration activities depend largely on the results of seismic data acquisition and analysis, especially for the basement complexes. However, due to the non-layered nature and the structural complexity of the fractured basement, the seismic data contain a lot of noise and multiples. Although seismic processing improves data quality for Cenozoic sediments, the seismic data quality for fractured zones in the basement is still very limited.

The resolution of seismic data in fractured basement is often very low and therefore it is difficult to identify sufficient structural information and correlation interpreted data with well logs. Therefore, it is important to improve the seismic data quality, remove noise and multiples, enhance faults and fracture related signals for assisting the geological interpretation and identifying exploration targets as well as optimizing well locations. This can be done by using different types of filters, including f-k, Tau-P, Radon transform to reprocess the seismic data. In addition, applying multiple-arrival migration approach such as Controlled Beam Migration (CBM) also assists revealing complicated, steeply dipping faults. Furthermore, application of geometric attributes, such as apparent dips and energy gradients, using shaded-relief maps from orthogonal attribute scan help highlighting fault systems. All these processes help improving the interpretation of fault and fracture networks in the basement complexes and providing better results for the identification of potential reservoirs and exploration targets. This paper describes the results of those processing sequences of a seismic dataset from Cuu Long Basin to reveal the sub-unconformity, structural complex basement architectures.

2. GENERAL GEOLOGY OF CUU LONG BASIN

The Cuu Long Basin is located on the southeast continental shelf of Vietnam (Fig. 1). This basin is a typical intracratonic rift basin formed on the Pre-Cenozoic basement, filled with Miocene and Oligocene terrigenous sedimentary sequences (Tapponier *et al.* 1986, Le Pichon *et al.* 1992, Gwang *et al.* 2001, Dong 2012). The sedimentary sequences are variable greatly in spatial distribution and thickness. In some portion of the basin, the sediment's thickness is up to about 7000-8000 m. The Pre-Cenozoic basement rocks comprise dominantly the Jurassic to Paleocene intrusive complexes and subordinate extrusive rocks and sedimentary units. The structure of the Pre-Cenozoic basement in the Cuu Long basin is very complex, which was caused by multiple deformational events, hydrothermal alteration, uplift, and weathering processes (Dong *et al.* 1999, Tan 1990).

Fig. 1. Cuu Long Basin in the Southeast Continental Shell of Vietnam (image courtesy of PetroVietnam).

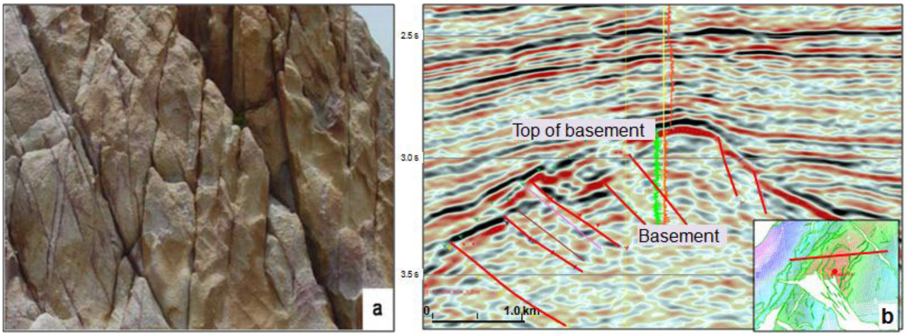


Fig. 2. Outcrops (a) and seismic cross-section (b) showing faults and fractures inside granite basement.

The multiple deformational events created numerous tectonic and/or solution-enhanced fracture systems with different directions and sizes in the basement. The deformation processes include shearing, extension and compression, which produced a network of cross-cutting faults and fractures within the basement rocks as well as commonly created horst and graben structures. The fracture and/or fault systems in the basement are variable in both trend and dip but are commonly steeply dipping (55-70°) (Cuong and Warren 2009, Duc 2014). Highly fractured zones with good reservoir properties are usually found in the basement highs that are controlled by regional-scale faults. In addition, the influence of surface processes on the basement rocks prior to the subsidence to form sedimentary basin, particularly weathering, also created thick zones of high pore space within the upper part of the basement complexes, and, as such, were improving basement reservoir quality (Areshev *et al.* 1992, Tan 1995, San *et al.* 1997). The hydrothermal processes, however, play negative role in reservoir quality by the infiltration of secondary materials into the pore space and lead to the reduction of porosity and permeability of the basemen rocks (Tan and Bo 1997).

Thus, the combined effects of tectonic deformation and subsequent subsurface processes in the basement magmatic rocks produced a mixture of fractures, vugs, and empty spaces that are favorable for hydrocarbon to migrate from the Oligocene–Miocene source rocks and accumulate into the highs near the top of the basement. Figure 2 shows faults and fractures of granite basement on (a) outcrops and (b) seismic section.

3. METHODS FOR IMPROVING FAULTS AND FRACTURES SIGNATURE WITHIN BASEMENT

The unique characteristics of basement reservoir require a careful evaluation and unique processing and interpretation methods. Identification of distribution of the fracture zones is important for extrapolation of favorable target for exploration (Tan and Thap 2001). However, in the absence of stratified, coherent reflectors, illumination of basement faults is more problematic than illumination of faults within the sedimentary cross-sections. With the complex interference pattern, non-layered and steeply dipping nature of faults and fractures, it is important to enhance seismic data processing, reducing noise and multiple reflections (Tan and Ha 2013). Seismic filters such as f-k, Tau-P, Radon transform were used to remove multiple reflections, revealing the weaker seismic signals. Seismic migration methods such as Kirchhoff Migration and Controlled Beam Migration (CBM) were used to image steeply dipping fault and fracture plane events. And selective seismic attributes methods were applied to further delineate and map out fractures.

3.1 Multiple attenuation by using filters

Analysis of seismic profiles from Cuu Long Basin has shown that seismic signals reflected from the fractured granitic basement is of very low quality, containing a lot of noise and multiples. Multiples are related with strong reflection surfaces of the Miocene, Oligocene sediments, and top-basement. Multiple suppression methods can be used based on either prediction criteria or normal move-out differential to improve the weak primary reflections. Using filters such as f-k, Tau-P, and Radon transform enhances the signal-to-noise ratio and improves seismic signals significantly.

***F-k filtering*:** Multiple reflections are filtered from seismograms by transforming them into an f-k array representing amplitude as a function of frequency (f) and wave number (k). Noise and seismic signals may overlap in time domain, but are distinguishable in f-k plots, because of the difference in frequency, and wave number between noise and signal. The inverse of the f-k transform of the multiple reflections is generated. The f-k array of the seismograms is filtered by weighting all samples with the inverse of the f-k transform of the multiple reflections (Duncan and Beresford 1994).

***Tau-P filtering*:** is an invertible transformation of seismic shot records expressed as a function of time and offset into the intercept time (t) and ray parameter (p) domain. The main advantage of this transformation is to present a point source shot record as a series of plane wave experiments. The data in this "dip" domain then can be transformed back into the time domain. An un-stacked seismic record or a common-midpoint gather can be described in terms of slope $dt/dx = p$ and intercept time τ, the arrival time t obtained by prospecting the slope back to $x = 0$ (Sheriff 1999). Hyperbolic reflections are then transformed into ellipses, straight events such as head waves and direct wave into points. Filtering can be done on the Tau-P map and the filtered result transformed back into record. Multiples often have a greater move-out than primaries arriving at the same time. Multiples are periodic in the Tau-P domain, so that it can be suppressed by applying predictive deconvolution in this domain (Benoliel *et al.* 1987, van der Baan and Kendall 2002). After normal move-out correction, errors are approximately parabolic and tend to map to points in the parabolic Radon domain.

***Radon transform*:** Radon transforms map data before and after move-out correction into points, and multiples can be recognized in the Radon domain. The identified multiple signals can then be subtracted from the data to improve the interpretation process. Radon transforms can effectively model such events, which can then be subtracted from the original data. The Radon transform attempts to more equally suppress multiples for all traces, including near- and far-offset traces (Chapman 1981, Foster and Mosher 1992).

Thus, using filters such as f-k, Tau-P, and Radon transform for multiple attenuation is not only successful in improvement of seismic signals for sedimentary units but also works well in fractured, non-layered and structural complex basement in the Cuu Long Basin.

3.2 Improving image by using Seismic Migration

Migration of seismic data corrects reflection from dipping surface to their true positions, collapses diffractions, increases spatial resolution and resolves areas of complex geology. A major difference in migration algorithms arises from the way the velocity field is utilized (Gardner 1985, Wang and Pann 1996). Kirchhoff Migration is one of seismic migration methods that are extensively used in both layered environment and non-layered fractured basement by summing the energy along diffraction curves (Sun *et al.* 2000, Yilmaz 2001). The advantage of this type of migration is that it can enhance signals of steeply dipping objects and moderate lateral velocity variations. In a single-arrival migration algorithm, only one arrival is imaged, depending on certain predefined criteria. This is the weakness of Kirchhoff Migration method that requires ray-tracing to handle strong lateral velocity variations.

In order to reduce the disadvantage, Controlled Beam Migration (CBM) can handle multi-arrival ray paths, and preserve steeply dipping reflections, resulting in a cleaner image (Raz 1987, Hill 2001, Gray *et al.* 2009). A good review of CBM with applications to this study can be found in Sun *et al.* (2007), Bone *et al.* (2008), and Elkady *et al.* (2008).

3.3 Improving image by using Geometrical Attributes

For further identification of fault signature, the attribute illumination approach has been applied to well-established vector attributes including structural dip and azimuth and amplitude energy gradients to provide greater interpreter interaction (Chopra and Marfurt 2007, Ha *et al.* 2014). Based on

Barnes (2003), we could mathematically generate simple axis rotations and project the two orthogonal dip or energy gradient components along the surface against the direction of illumination. A planar surface, such as dipping horizon or faults, can be presented by its true dip azimuth θ and strike ψ. The true dip θ can be presented by apparent dips θ_x and θ_y along the x and y axes. For time-migrated seismic data, it is more convenient to measure apparent seismic time dip (p_x, p_y) components along inline and cross-line directions. For depth-migrated seismic data, such as our Cuu Long survey, we compute θ_x and θ_y and display them either as components or as dip magnitude, θ, and dip azimuth, ψ, or alternatively as dimensionless (p_x, p_y) measured. The relationship between apparent seismic time/depth dips and apparent angle dips are:

$$p_x = 2 * \tan\theta_x / v \quad \text{and} \quad p_y = 2 * \tan\theta_y / v \ ,$$

where v is an average time to depth conversion velocity. We can compute apparent dip at any angle ψ from the north through a simple trigonometric rotation (Marfurt 2006):

$$p_\psi = p_x \cos(\psi - \varphi) + p_y \sin(\psi - \varphi) \ ,$$

where ϕ is the angle of the inline seismic axis from the north.

Amplitude gradient vector attribute that has inline and cross-line components (g_x, g_y) is also described. We can therefore compute an amplitude gradient at any angle, ψ, from the north:

$$g_\psi = g_x \cos(\psi - \varphi) + g_y \sin(\psi - \varphi) \ .$$

Using these equations, we are able to animate, through a suite of apparent dip, amplitude gradient images at increments of 30° to see which perspective best illuminates structural features of interest.

4. RESULTS

4.1 Improvement of seismic signal from fractured basement by using filters

In order to apply seismic filtering and migration methods for improving seismic signal from fractured basement, we create seismic models representing fractured basement in Cuu Long Basin, with the velocity varying vertically and laterally (Fig. 3).

The models show that there are sedimentary boundaries with strong reflection coefficient that could create strong multiples. The multiples from these boundaries can be imbedded into the lower sections, especially the basement section. With the results from the models, we can improve our

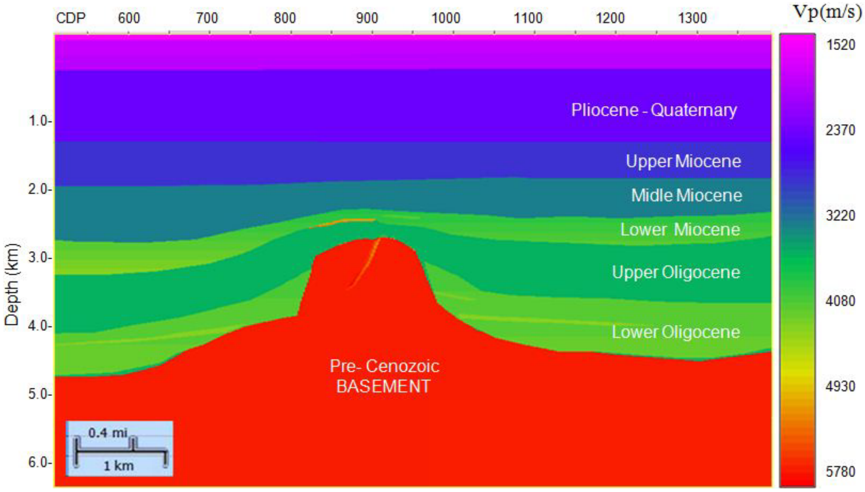


Fig. 3. Seismic velocity model representing fractured basement in Cuu Long Basin.

processing workflow, predict the parameters for the filters for better removing the multiples, and enhance the true signals from the basement. The granite matrix velocity is estimated to be about 5.4-6.0 km/s; however, some imaging studies show that the velocity for the granite section in the model of about 4.6 km/s is more applicable for the migration and stacking response. The velocity contrast between the basement and sedimentary layers varies in different regions. Some parts of the basement covered by lower Oligocene with lower acoustic impedance contrast gives weak seismic response. Characteristics of fractured zone are different from fresh basement; the ratio between *P*- and *S*-wave velocity in fractured zone increases from 1.7-1.9 to 2.0, and acoustic impedance (AI) decreases by 10% (Tan 1990).

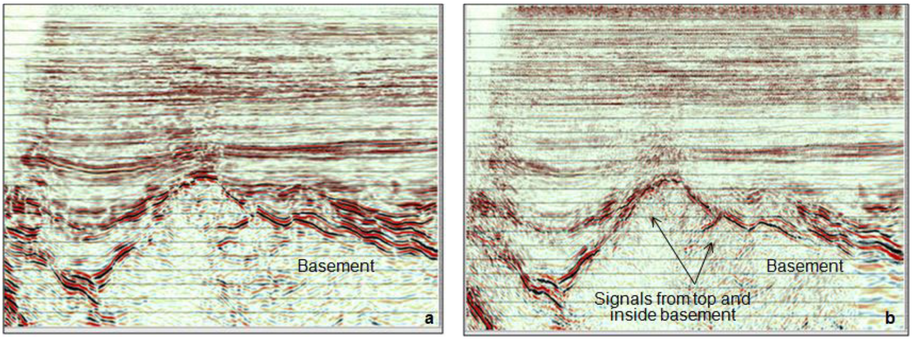


Fig. 4. f-k filtering could improve seismic signal from top and inside the basement: (a) middle and far offset stacked section, (b) middle and far offset stacked section with f-k filtering.

The f-k filtering was applied for full, middle and far offset. Figure 4 shows that by applying f-k filter on middle and far offset, the upper part of basement is significantly improved and the signature of fractured zones within the basement is revealed better.

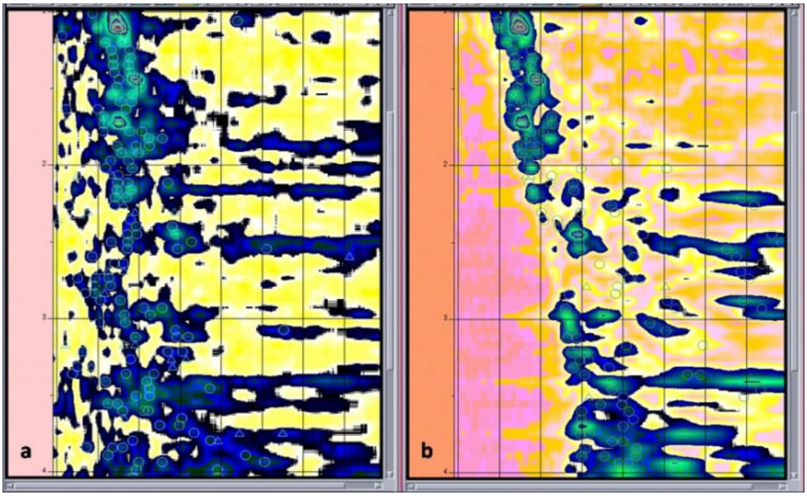


Fig. 5. Enhanced resolution and velocity analysis with Radon transform: (a) velocity spectrum without Radon transform, (b) with Radon transform.

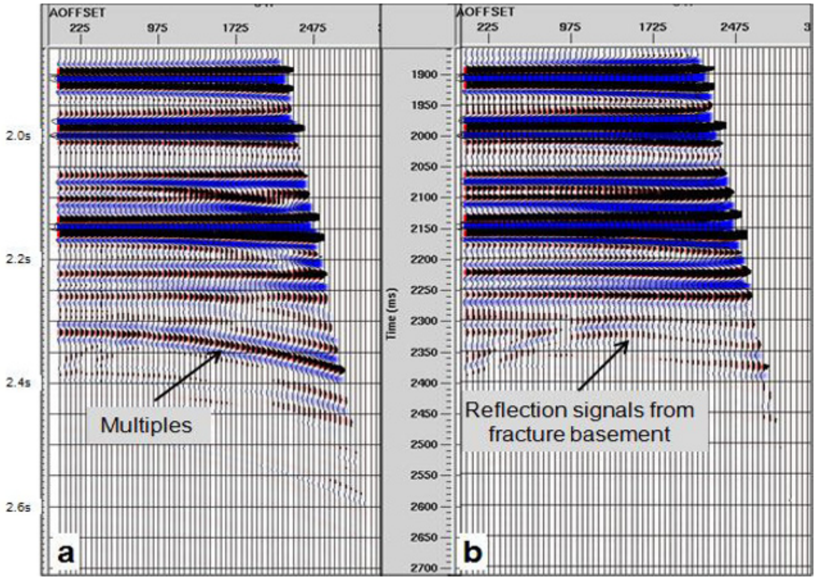


Fig. 6. Enhanced reflection signals from fractured basement with Radon transform: (a) common midpoint gather without Radon transform, (b) with Radon transform.

When applying Radon transform, the velocity spectrum and resolution is enhanced (Fig. 5), so that the velocity analysis can be done more accurately. With improved velocity analysis using Radon transform, multiples are suppressed, and signals are enhanced. This is especially important for the basement because of its originally weak S/N ratio. The signal from fractured system is greatly enhanced, as shown in seismic gather (Fig. 6) and seismic section (Fig. 7).

Figure 8 shows the results of applying combination of Radon transform and Tau-P filters. After effectively eliminating multiples by the Radon transform and Tau-P filtering, reflective signals from the fractured zone within the basement are better visible (Fig. 8b). The faults and fractures can be better interpreted on the filtered data. These results are then verified and conform with sonic, gamma ray characteristics on well data (Fig. 9).

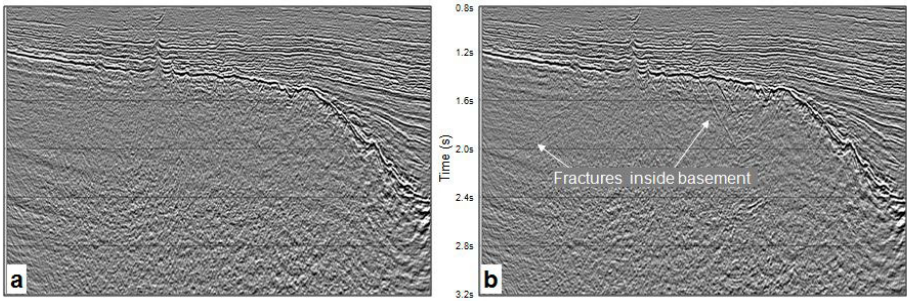


Fig. 7. Compare seismic section before (a) and after velocity accuracy correction (b). In the seismic section (b) the image of fractures on the roof of basement distinguished by the velocity analysis was enhanced.

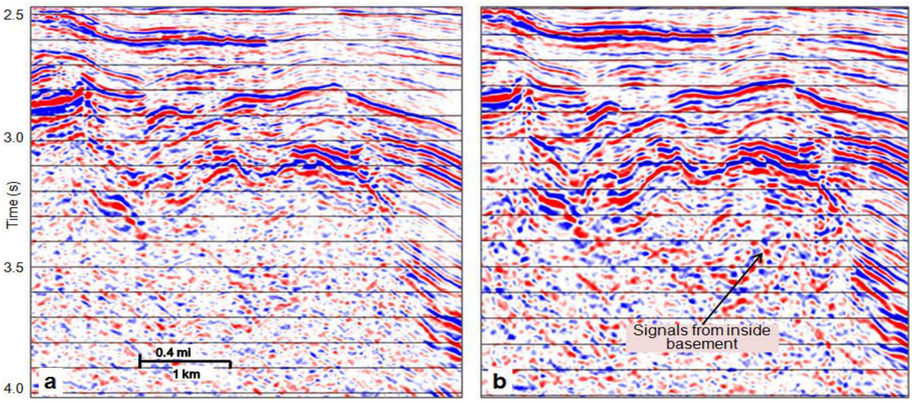


Fig. 8. Seismic section showing the effect of combining Radon transform and Tau-P filter: (a) seismic section, (b) seismic section with Radon transform and Tau-P filter.

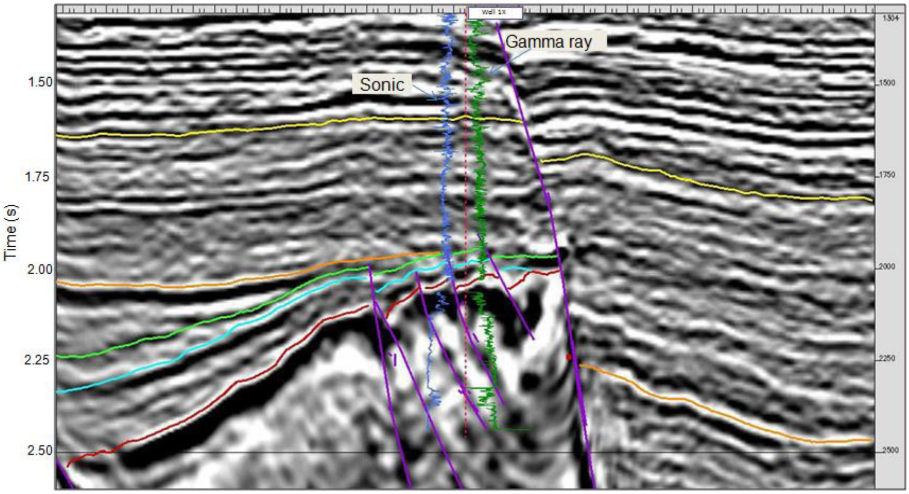


Fig. 9. Reflections from fracture zones have been confirmed by well data.

4.2 Improving image in fractured basement by Kirchhoff and Control Beam Migration

In order to qualitatively define the nature of fractures within the basement in Cuu Long Basin, besides using seismic filters, both Kirchhoff Migration and CBM approaches were used for improving image of fractures within the basement. Kirchhoff Migration provides better imaging for both sediments and the top of basement (Fig. 10). Even though the results of Kirchhoff Migration improves the image of fractured basement, it is still necessary to

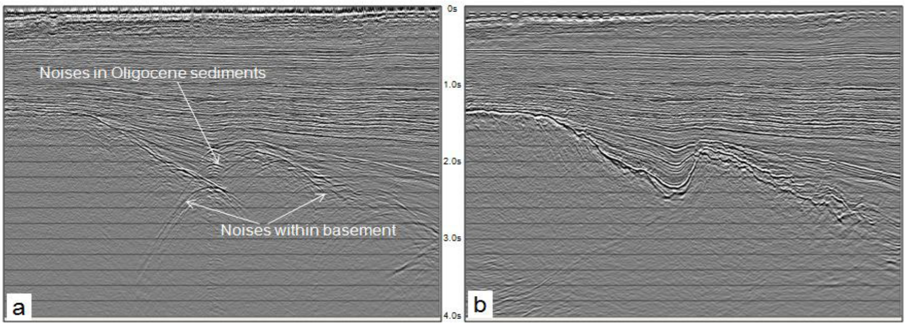


Fig. 10. The quality of seismic section related with top and inside basement is improved by Kirchhoff Migration. (a) Seismic section before using Kirchhoff Migration is of very low quality and contains a lot of noise, (b) Seismic section after using Kirchhoff Migration could provide good imaging for the identification of the cover and within the basement.

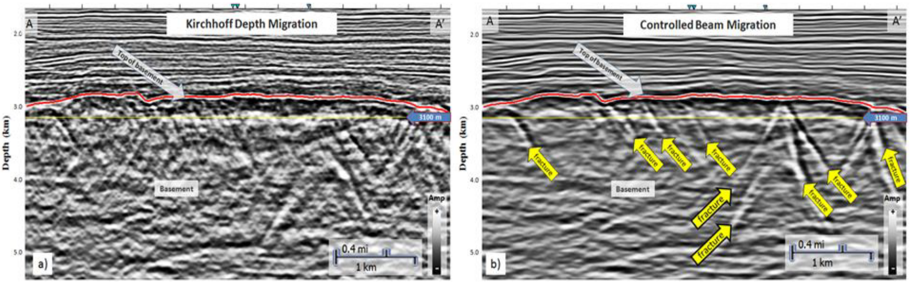


Fig. 11. Comparison of Kirchhoff and Controlled Beam Migration in vertical seismic section (the fault indicates location of depth slice displayed in Fig. 12). (a) Kirchhoff depth migration, (b) Controlled Beam Migration.

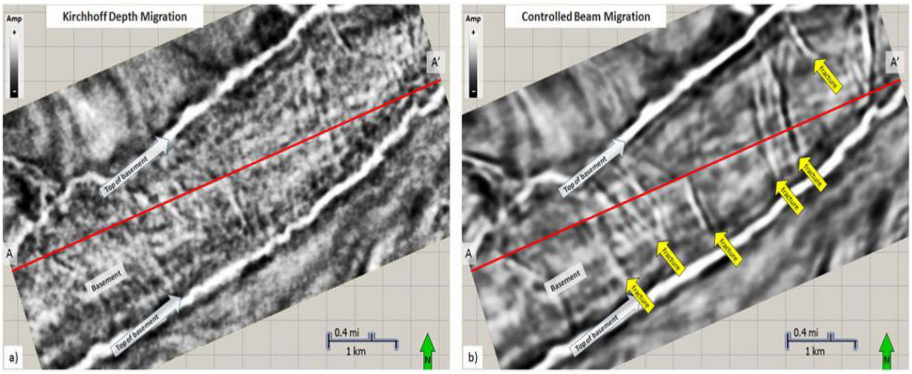


Fig. 12. Comparison of Kirchhoff and Controlled Beam Migration on depth slice at $z = 3100$ m (vertical seismic section in line AA′ is shown in Fig. 13). (a) Kirchhoff depth migration, (b) Controlled Beam Migration. While the lateral resolution is slightly lower, the Controlled Beam Migration much better indicates the fractures.

apply CBM to solve the imaging problem of the basement, which is naturally complex. Figure 11 shows a comparison of the results applying Kirchhoff Migration and CBM. The faults lineaments are much easier identified in the CBM section, with less noise interfered. A comparison of Kirchhoff Migration and CBM on depth slices $z = 3100$ m is shown in Fig. 12. While the lateral resolution is slightly lower, the CBM can indicate fractures much better, which helps improving interpretation.

4.3 Seismic Attribute illumination of basement faults

Geometrical Attributes such as structural dip and azimuth and amplitude energy gradients are used to improve locating faults and fractures inside the basement. The results obtained are shown in Figs. 14-17. Figure 14 shows

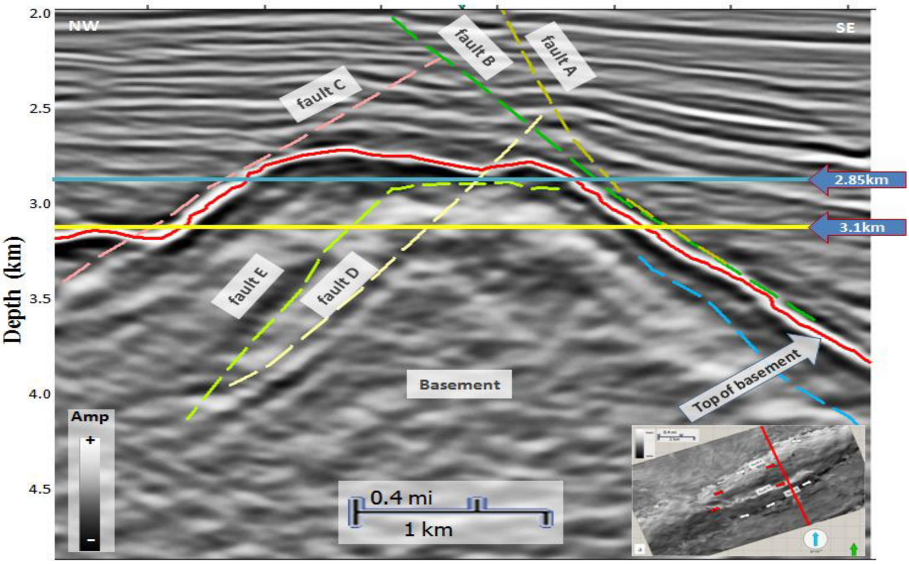


Fig. 13. Vertical seismic section showing top of basement and interpreted faults.

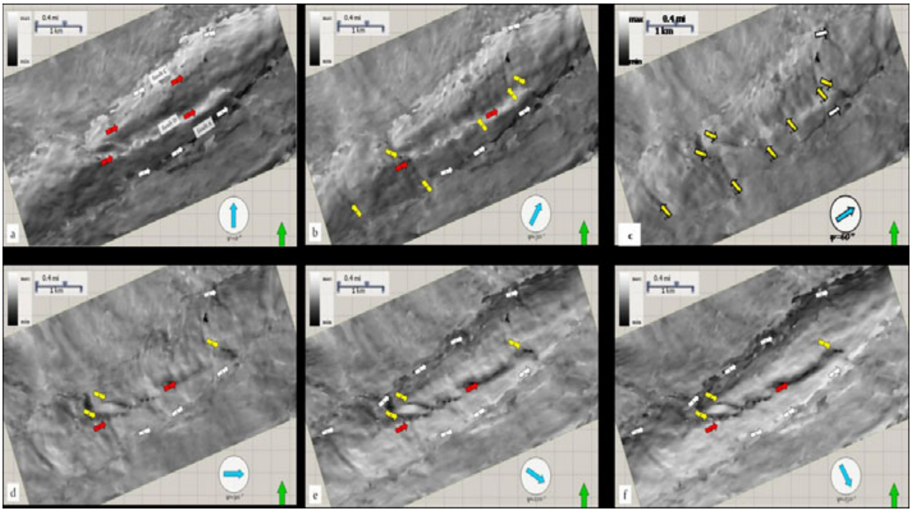


Fig. 14. Depth slices at $z = 2850$ m through apparent dip, p_y, computed at apparent direction $\psi = 0°$ (a), 30° (b), 60° (c), 90° (d), 120° (e), and 150° (f) from the north. Arrows (1) indicate lineaments which were interpreted as main NE-SW faults running along basement top. Arrows (2) indicate the faults that cut across the basement, in N-S and NW-SE direction, and Arrows (3) indicate subtle faults running NE-SW and cutting into the basement.

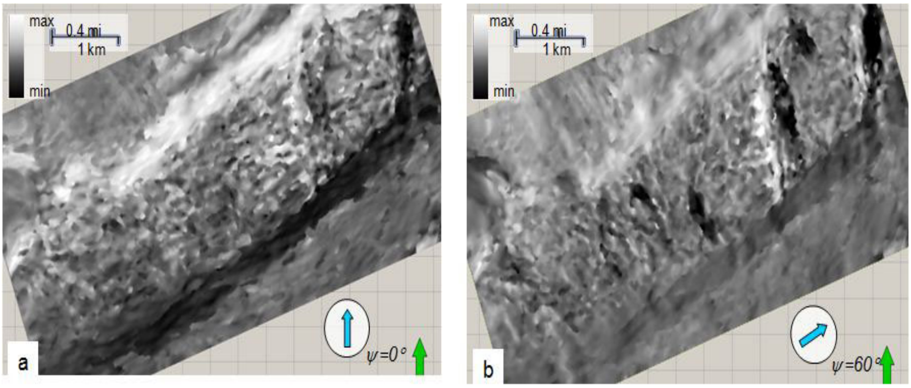


Fig. 15. Depth slices at $z = 3100$ m through apparent dip, p_ψ, computed at apparent direction $\psi = 0°$ and 60° from the north.

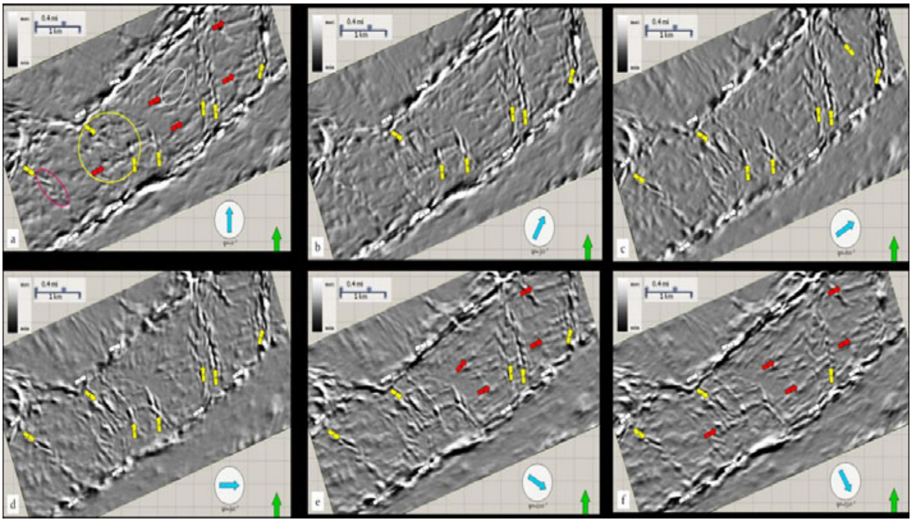


Fig. 16. Depth slices at $z = 3100$ m through apparent amplitude gradients, g_ψ, computed at apparent direction $\psi = 0°$ (a), 30° (b), 60° (c), 90° (d), 120° (e), and 150° (f) from the north. White arrows (1) indicate lineaments which were interpret as main NE-SW faults running along basement top. Yellow arrows (2) indicate the faults that cut across the basement, in N-S and NW-SE direction. Red Arrows (3) indicate subtle faults running NE-SW and cutting into the basement.

depth slices at the top of basement level (2850 m) through the apparent dip volume, p_ψ, as a function of azimuth, using the equation in Subsection 3.3 for $\psi = 0°$, 30°, 60°, 90°, 120°, and 150°. In these depth slices through the

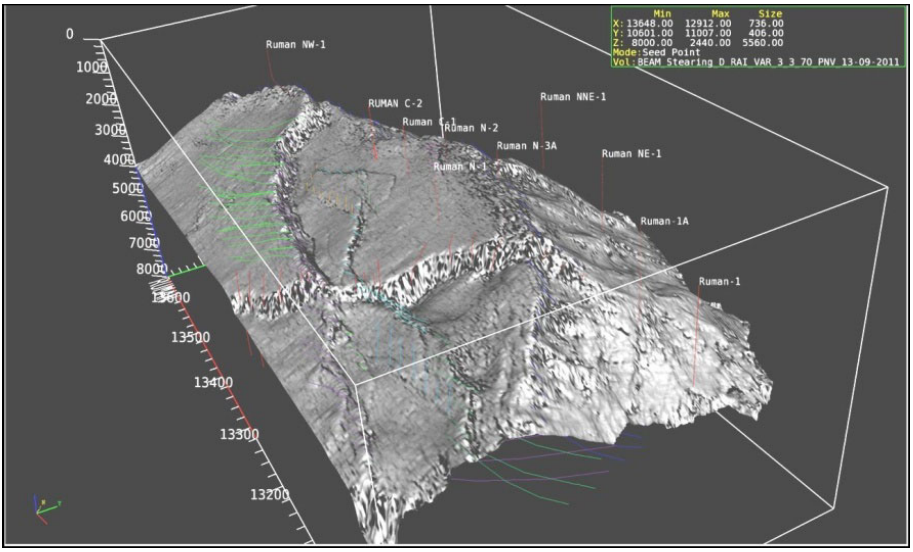


Fig. 17. Curvature, dip and azimuth attributes of top basement in 3D cube.

apparent dip, the NE-SW major faults running along the basement top, the N-S and NW-SE faults cutting across the basement, and the NE-SW subtle faults cutting into the basement have appeared clearer. The reflector dip enhances different lineament features as the direction of illumination is rotated, as we can see by comparing Figs. 15a and 15b. Since the dip attribute measures the dip of a reflector surface, the dip attribute computed on or near the top of basement reveals lineaments well. However, if we look deeper inside the basement, the dip estimates become noisier, making it harder to interpret the results.

In contrast, amplitude gradients are computed along local dip and better delineate high energy cross-cutting fractures. Apparent amplitude energy gradient results were generated at the illumination direction $\psi = 0°$, 30°, 60°, 90°, 120°, and 150° from the north. Figure 16 shows depth slice at 3100 m cutting through the top of the granite basement. In this depth slices through apparent amplitude gradients, the faults like the ones shown in Fig. 14 are distinguished. From the results of study of structural dip and azimuth and amplitude energy gradients, the imaging of faults and fractured on the top basement expressed in 3D cube (Fig. 17).

5. CONCLUSIONS

Highly fractured zones in pre-Cenozoic, structurally complex granitic basement highs in the Cuu Long Basin form potential petroleum reservoirs. The seismic data quality of these sections is poor with the interference of strong noise and multiple reflections.

The application of combining filtering techniques, including f-k, Tau-P, and Radon transform, enhances signal-to-noise ratio significantly and allows to image the fractured basement signal better.

Kirchhoff Migration and Controlled Beam Migration, on the other hand, effectively image the top basement and intra-basement events in which the image quality of CBM is clearly superior to that of Kirchhoff Migration. Geometric attributes such as dip and amplitude energy gradients, are multi-component in nature and are thus amenable to visualization from different user-controlled perspectives.

Thus, the application of a suite of filters, migrations, and attribute analysis enhances the clarity of internal structures within the basement complexes within the Cuu Long Basin such as faults and fractures that might otherwise be miss-interpreted.

Acknowledgements. This work is funded by the Vietnam's National Foundation for Science and Technology Development (NAFOSTED), Project No. 105.04.27.09. The support by PetroVietnam and OU Attribute-Assisted Seismic Processing and Interpretation (AASPI) consortium is acknowledged.

References

Areshev, E.G., T.L. Dong, N.T. San, and O.A. Shnip (1992), Reservoirs in fractured basement on the continental shelf of Southern Vietnam, *J. Petrol. Geol.* **15**, 4, 451-464, DOI: 10.1111/j.1747-5457.1992.tb01045.x.

Barnes, A.E. (2003), Shaded relief seismic attribute, *Geophysics* **68**, 4, 1281-1285, DOI: 10.1190/1.1598120.

Benoliel, S.D., W.A. Schneider, and R.N. Shurtleff (1987), Frequency wavenumber approach of the Tau-P transform: Some applications in seismic data processing, *Geophys. Prospect.* **35**, 5, 517-538, DOI: 10.1111/j.1365-2478.1987.tb00833.x.

Bone, G., N.T. Giang, D.N. Quy, V.N. An, D. Pham, J. Sun, and Q. Tang (2008), Improvements in seismic imaging in fractured basement, Block 15-1, Offshore Vietnam: Fractured basement reservoir, *Petrovietnam Rev.* **4**, 63-69.

Chapman, C.H. (1981), Generalized Radon transforms and slant stacks, *Geophys. J. Roy. Astr. Soc.* **66**, 2, 445-453, DOI: 10.1111/j.1365-246X.1981.tb05966.x.

Chopra, S., and K.J. Marfurt (2007), *Seismic Attributes for Prospect Identification and Reservoir Characterization*, SEG Geophysical Developments Series, No. 11, Society of Exploration Geophysicists, Tulsa, USA, 456 pp.

Cuong, T.X., and J.K. Warren (2009), Bach Ho field, a fractured granitic basement reservoir, Cuu Long basin, offshore SE Vietnam: a "Buried-Hill" play, *J. Petrol Geol*. **32**, 2, 129-156, DOI: 10.1111/j.1747-5457.2009.00440.x.

Dong, H.N. (2012), Geological and tectonic characterization in the Northern part of Cuu Long Basin during Eocene-Oligocene, Ph.D. Thesis, Hanoi University of Mining and Geology, Hanoi, Vietnam.

Dong, T.L., P.H. Long, H.V. Quy, P.D. Hai, and T.S. Hau (1999) Distribution of fractures, faults and their formation in the basement of South Vietnam continental shelf and adjacent areas, *Petrovietnam Rev*. **3**, 4-13.

Duc, N.A. (2014), Characteristics of Pre-Tertiary fractured granitoid basement in Hai Su Den field, Cuu Long basin, *Petrovietnam J*. **5**, 15-22.

Duncan, G., and G. Beresford (1994), Slowness adaptive f-k filtering of prestack seismic data, *Geophysics* **59**, 1, 140-147, DOI: 10.1190/1.1443525.

Elkady, N., Y.C. Tan, R. Baker, Z.J. Zhou, and A. Tan (2008), Imaging improvement of fractured and faulted basement through Controlled Beam Migration. Examples from Diamond field in Vietnam. **In:** *International Petroleum Technology Conference, Kuala Lumpur, Malaysia*.

Foster, D.J., and C.C. Mosher (1992), Suppression of multiple reflections using the Radon Transform, *Geophysics* **57**, 3, 386-395, DOI: 10.1190/1.1443253.

Gardner, G.H.F. (1985), *Migration of Seismic Data*, SEG Monograph Series, 462 pp.

Gray, S.H., Y. Xie, C. Notfors, T. Zhu, D. Wang, and C. Ting (2009), Taking apart Beam Migration, *The Leading Edge* **28**, 9, 1098-1108, DOI: 10.1190/1.3236380.

Gwang, H.L., L. Keumsuk, and J.S. Watkins (2001), Geologic evolution of the Cuu Long and Nam Con Son basins, offshore Southern Vietnam, South China Sea, *AAPG Bull*. **85**, 6, 1055-1082.

Ha, M.T., O. Elebiju, and K.J. Marfurt (2014), Attribute illumination of basement faults, examples from Cuu Long Basin basement, Vietnam and the Midcontinent, USA, *Interpretation* **2**, 1, SA119-SA126, DOI: 10.1190/INT-2013-0091.1.

Hill, N.R. (2001), Prestack Gaussian-Beam depth migration, *Geophysics* **66**, 4, 1240-1250, DOI: 10.1190/1.1487071.

Le Pichon, X., M. Fournier, and L. Jolivet (1992), Kinematic, topography, shortening, and extrusion in the India-Eurasia collision, *Tectonics* **11**, 6, 1085-1089, DOI: 10.1029/92TC01566.

Marfurt, K.J. (2006), Robust estimates of 3D reflector dip and azimuth, *Geophysics* **71**, 4, 29-40, DOI: 10.1190/1.2213049.

Raz, S. (1987), Beam stacking: a generalized preprocessing technique, *Geophysics* **52**, 9, 1194-1210, DOI: 10.1190/1.1442383.

San, N.T., N. Giao, and T.L. Dong (1997), Pre-Tertiary basement – The new objective for oil and gas exploration and production in the Continental Shelf of

South Vietnam. **In:** *Proc. Int. Conf. on Petroleum Systems of SE Asia and Australia*, 461-465.

Sheriff, R.E. (1999), *Encyclopedic Dictionary of Exploration Geophysics*, SEG.

Sun, J., P. Don, J. Sun, Q. Tang, G. Bone, and N.T. Giang (2007), Imaging of fractures and faults inside granite basement using controlled beam migration. **In:** *ASEG 19th International Geophysical Conference & Exhibition.*

Sun, Y., F. Qin, S. Checkles, and J.P. Leveille (2000), 3D Prestack Kirchhoff Beam Migration for depth imaging, *Geophysics* **65**, 5, 1592-1603, DOI: 10.1190/1.1444847.

Tan, M.T. (1990), The enhancement of seismic prospecting effectiveness for oil and gas under the conditions of the sedimentary basins in the continental shelf of Vietnam, *Zesz. Nauk AGH Geofiz. Stos. Bull.* **6**, 1323, 142.

Tan, M.T. (1995), Seismic stratigraphic studies of the continental shelf of Southern Vietnam, *J. Petrol. Geol.* **18**, 3, 345-354, DOI: 10.1111/j.1747-5457.1995.tb00910.x.

Tan, M.T., and M.T. Ha (2013), Enhancement of the seismic technology to improve imaging quality of fractured reservoir inside granite basement. **In:** *Proc. 11th SEGJ Int. Symp., Yokohama, Japan*, 160-163.

Tan, M.T., and N.Q. Thap (2001), The possibility of applying geophysical methods to study fractured basement reservoir in the continental shelf of Vietnam. **In:** *Proc. Offshore Technology Conf.*, 451-456.

Tan, T.K., and N.Q. Bo (1997), Geological modeling and reservoir properties of basement rocks of the South Vietnam continental shelf. **In:** *Proc. Int. Conf. on Petroleum Systems of SE Asia and Australia*, 539-544.

Tapponier, P., M. Peltzer, and R. Armijor (1986), On the mechanisms of the collision between India and Asia, *Geol. Soc. London Spec. Publ.* **19**, 115-157.

van der Baan, M., and J.M. Kendall (2002), Estimating anisotropy parameters and traveltimes in the τ-p domain, *Geophysics* **67**, 4, 1076-1086, DOI: 10.1190/1.1500368.

Wang, B., and K. Pann (1996), Kirchhoff migration of seismic data compressed by matching pursuit decomposition. **In:** *SEG Technical Program Expanded Abstracts*, 1642-1645.

Yilmaz ,O. (2001), *Seismic Data Analysis. Processing, Inversion, and Interpretation of Seismic Data*, Series: Investigations in Geophysics, Vol. 1, Society of Exploration Geophysicists, Tulsa, USA.

Received 7 October 2015
Received in revised form 17 December 2015
Accepted 8 February 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2232-2243
DOI: 10.1515/acgeo-2016-0097

An Easy Method for Interpretation of Gravity Anomalies Due to Vertical Finite Lines

İbrahim KARA and Nihan HOSKAN

Department of Geophysical Engineering, Faculty of Engineering, Istanbul University, Avcılar, Turkey; e-mail: nihan@istanbul.edu.tr

Abstract

A new method is introduced to determine the top and bottom depth of a vertical line using gravity anomalies. For this, gravity at a distance x from the origin and horizontal derivative at that point are utilized. A numerical value is obtained dividing the gravity at point x by horizontal derivative. Then a new equation is obtained dividing the theoretical gravity equation by the derivative equation. In that equation, assigning various values to the depth and length of vertical line, several new numerical values are obtained. Among these values, a curve is obtained for the one that is closest to the first value from attending the depth and length values. The intersection point of these curves obtained by repeating this procedure several times for different points x yield the real depth and length values of the line. The method is tested on two synthetics and field examples. Successful results are obtained in both applications.

Key words: gravity anomaly, vertical finite line, depth and length curves.

1. INTRODUCTION

Interpretation of the potential-based data is usually faced with multi-solution. For example, gravity anomalies of sphere and vertical line cannot be distinguished by naked eye. In most cases, lines are accepted to be lateral in evaluations (Odegard and Berg 1965, Gay 1965, Abdelrahman 1990,

Ownership: Institute of Geophysics, Polish Academy of Sciences;

Abdelrahman *et al.* 1989). For the interpretation of the spherical sources, Mohan *et al.* (1986) used the Mellin transform, Shaw and Agarwall (1990) applied the Walsh transform, Abdelrahman and H.M. El-Araby (1993) utilized correlation factors, and Abdelrahman and T.M. El-Araby (1993) operated the least squares approach. Gupta (1983) used the least squares technique for interpretation of the vertical infinite line and also Kara and Kanli (2005) developed nomograms for the vertical finite line. To analyze the gravity anomalies due to simple geometrical structures, several methods have been developed (Essa 2007a, 2011, 2012, 2014). In a study similar to the method presented in this work, Essa (2007b) calculated shape factor and source depths of a sphere, and horizontal and vertical infinite lines. However, in his study, measurement values should be very accurate, otherwise the method becomes unsuccessful.

In this study, a new method has been developed for determination of depth and length of vertical finite line. Here, along a profile, gravity and horizontal derivative values are utilized. In order to show the validity of the method, it is tested on two numerical examples and then top and bottom depths of the source body are determined assessing the Louga gravity anomaly (USA).

2. FORMULATION OF THE PROBLEM AND SOLUTION

The gravitational acceleration produced by a vertical line extending to a finite depth to the bottom at an observation point, P, displayed in Fig. 1, is expressed with the following relation by Nettleton (1942):

$$g(x) = A\left[\frac{1}{\left(x^2+h^2\right)^{1/2}} - \frac{1}{\left[x^2+\left(h+L\right)^2\right]^{1/2}}\right] , \tag{1}$$

where A is the amplitude coefficient ($A = G\rho_l$, where G is gravity constant, ρ_l is the linear density of anomalous mass in the line), h is the depth of upper surface , L is the length of the body, and x is the distance from observation point to the origin.

The horizontal derivative of Eq. 1 is:

$$g_x(x) = A\left[\frac{-x}{\left(x^2+h^2\right)^{3/2}} - \frac{-x}{\left[x^2+\left(h+L\right)^2\right]^{3/2}}\right] , \tag{2a}$$

but the horizontal first gradient values in the field study are obtained as:

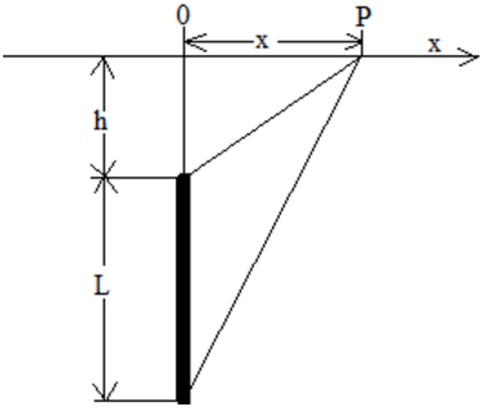


Fig. 1. Vertical line extending to a finite depth.

$$g_x(x) = \left[g\left(x+dx\right) - g\left(x-dx\right)\right]/2dx \ , \tag{2b}$$

where dx is the grid-spacing.

If Eqs. 1 and 2 are divided by each other and left and right sides of the resulting equation are termed F_1 and F_2, respectively, the following equations are obtained:

$$F_1 = \frac{g(x)}{g_x(x)} \ , \tag{3a}$$

$$F_2 = \frac{\dfrac{1}{\left(x^2+h^2\right)^{1/2}} - \dfrac{1}{\left[x^2+\left(h+L\right)^2\right]^{1/2}}}{\dfrac{-x}{\left(x^2+h^2\right)^{3/2}} - \dfrac{-x}{\left[x^2+\left(h+L\right)^2\right]^{3/2}}} \ , \tag{3b}$$

where F_1 is obtained dividing the measurement value at any observation point ($x \neq 0$) by horizontal derivative value at the same point. Then, F_2 is computed for a small h value (for example $h = 1$) and increasing L values (for example $L = 1, 2, 3.....50$). Meanwhile, in each calculation, the difference between F_1 and F_2 is determined. For the least value of this difference, h and L values are saved. Then h is slightly increased (for example $h = 2$) and the same procedure is repeated. The values of h and L (as h is the vertical axis and L is the horizontal axis) obtained from here are plotted in a coordinate system and a curve is obtained. The flow chart of the method is illustrated in Fig. 2. This process is repeated several times for different observation points and the intersection point of curves drawn on the same co-

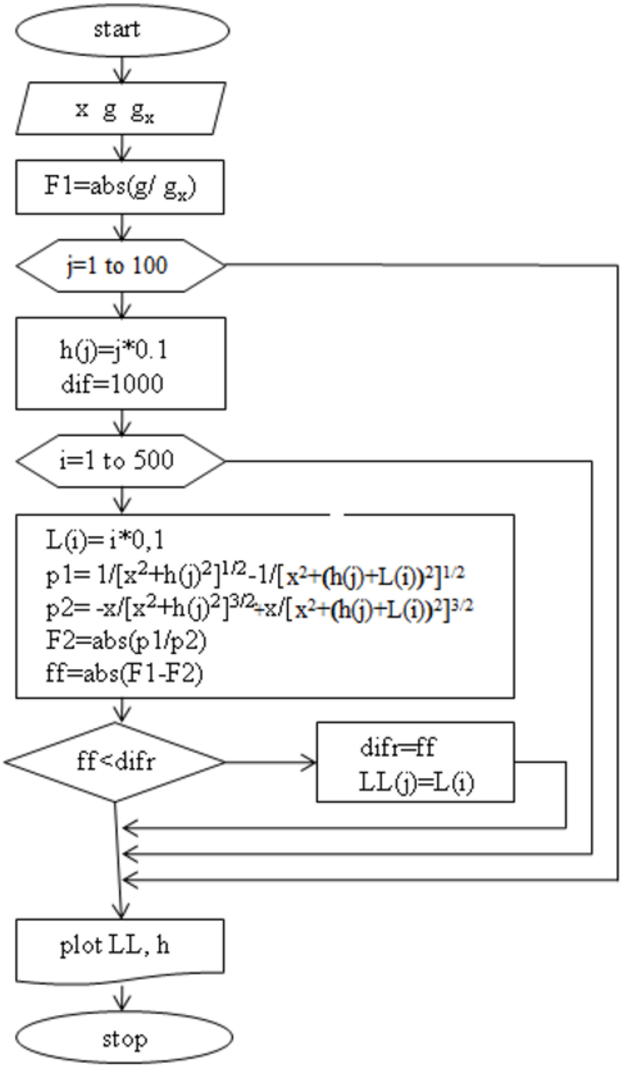


Fig. 2. The flow chart of the method for a single curve.

ordinate system yield the depth and length of the vertical line. If these curves do not intercept, it means that *h* or *L* or both are not sufficiently extended and they are extended until the intersection point of these curves is provided.

3. SYNTHETIC EXAMPLES

For testing the validity of the proposed method, the application is carried out on two theoretical models. In the first application, the gravity anomaly of a vertical finite line with $A = 100$ mGal*m, $h = 5$ m and $L = 30$ m is used (Fig. 3a).

The horizontal derivative anomaly of this anomaly is shown in Fig. 3b.

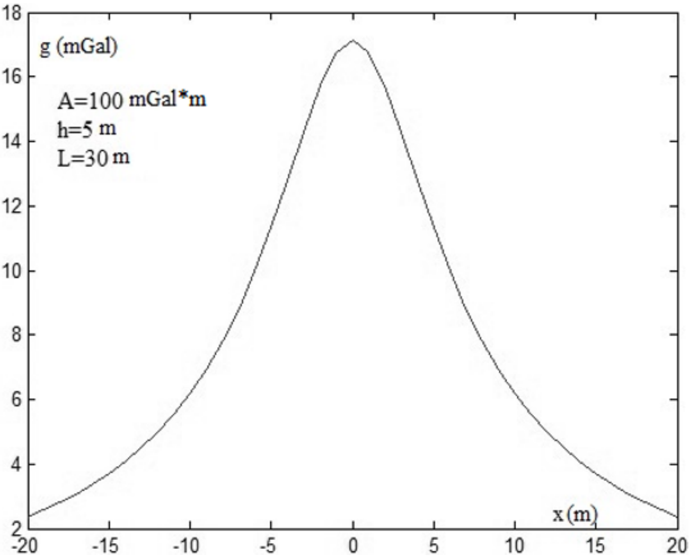


Fig. 3a. Gravity anomaly used in the first application.

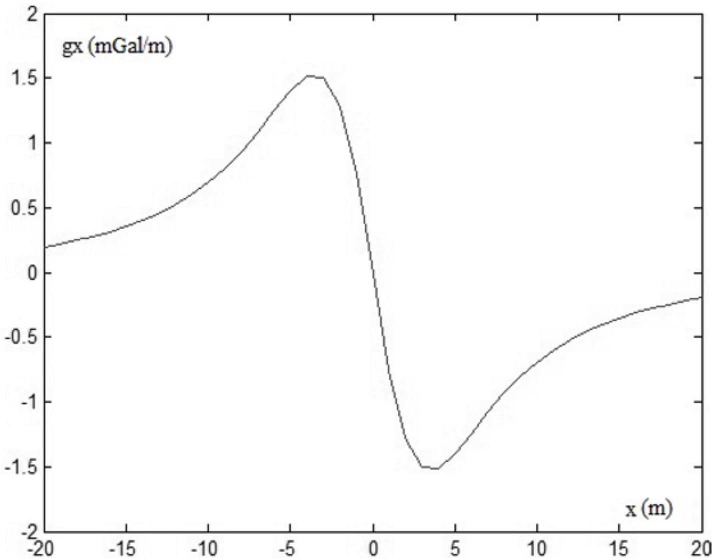


Fig. 3b. Horizontal derivative anomaly of the anomaly shown in Fig. 3a.

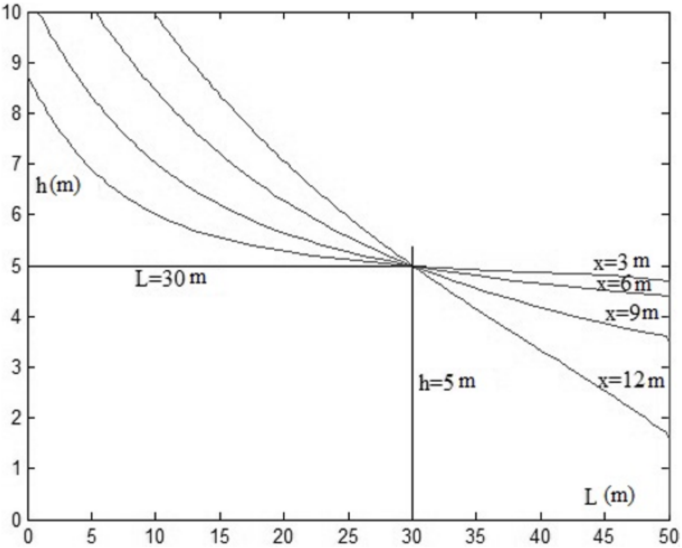


Fig. 3c. The curves obtained when the present method is applied to anomalies in Fig. 3a and b.

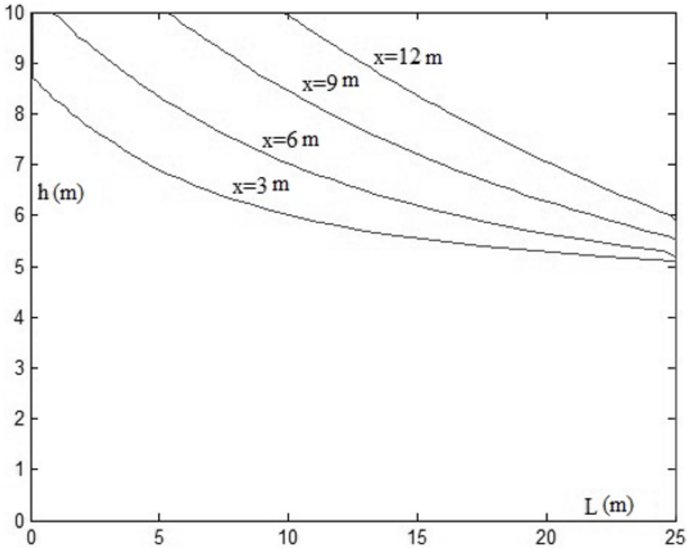


Fig. 3d. For the same application, the curves that are not intersecting due to insufficiently extended *L*.

Then the method is applied using gravity and horizontal derivative values at observation points $x = 3\text{-}6\text{-}9\text{-}12$ m and curves community given in Fig. 3c is obtained. As shown therein, $h = 5$ m and $L = 30$ m are found.

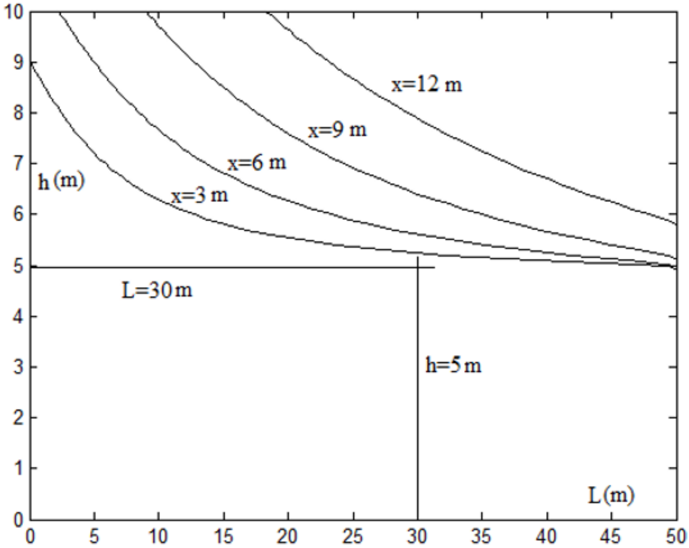


Fig. 3e. The curves obtained by adding 1 m to the reference level and applying the present method to the anomaly displayed in Fig. 3a.

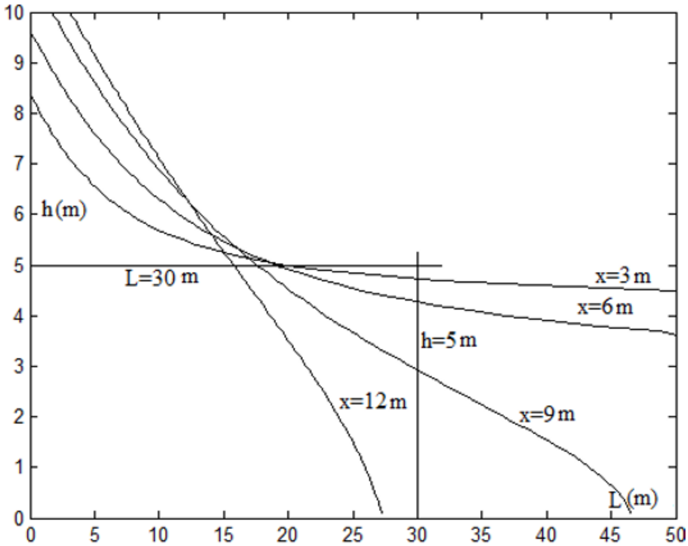


Fig. 3f. The curves obtained by subtracting 1 m from the reference level and applying the present method to the anomaly displayed in Fig. 3a.

These values are in excellent agreement with the initial parameters and this result indicates the validity of the present method. In addition, as mentioned previously, if h and L are chosen smaller than their real values, the curves have no tendency to cross each other. In this synthetic example, $L = 30$ m is taken. If Eq. 3b is calculated to $L = 25$ m and the curves are drawn, these curves do not cross each other (Fig. 3d).

In order for the curves to intersect each other, the value of L should be slightly increased, as illustrated in Fig. 3c. Besides, if the reference level of anomaly is detected incorrectly, the curves mentioned above either do not cross each other or intersect at the point that is far away from where they must intersect. Namely, intercepting the curves each other in one point depends on selecting the reference level correctly. For explaining the importance of this fact, Fig. 3e is obtained by adding 1 m and Fig. 3f is obtained by subtracting 1 m from the reference level of the anomaly displayed in Fig. 3a. Hence, using this method in practice, it must be changed several times with small intervals to the reference level of the anomaly until the intersection point of the curves with minimal dispersion is identified. Even though the curves still do not cross each other, the intersection point is accepted at the mid-point of the closest area of the curves.

For the second synthetic example, $A' = 1000$ mGal*m^2 and $h = 8$ m are chosen, using the following Eq. 4

$$g = A' \frac{h}{\left(x^2 + h^2\right)^{3/2}} , \tag{4}$$

from which the gravity anomaly of a sphere is obtained (here, $A' = 4/3\pi G R^3 \rho$). After horizontal derivative values of this sphere are calculated, using g and g_x values of this sphere at distances of $x = 3$-6-9-12 m, the

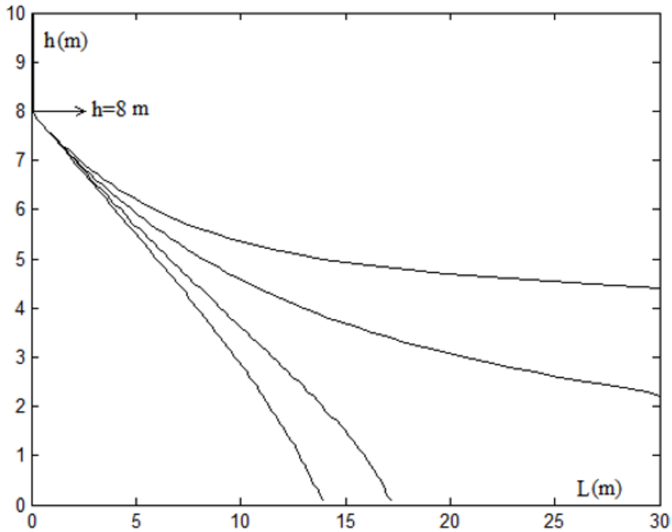


Fig. 4. Curve assemblage obtained for the second synthetic example (sphere).

proposed method is applied and the curves given in Fig. 4 are obtained. As shown in Fig. 4, the curves intersect at $h = 8$ m. In fact, the accepted h is at 8 m. However, careful inspection of Fig. 4 reveals that $L = 0$, indicating that the source body is a sphere.

4. FIELD EXAMPLE

For the field example, Louga gravity anomaly (USA) is used (Fig. 5a). The solid line in Fig. 5a represents observation values sampled in 1 km intervals. Thereafter, horizontal derivative values of the curve are obtained (Fig. 5b). The curves shown in Fig. 5c are obtained from application of the present method to gravity and horizontal derivative values at distance of $x = 2, 4, 6, 8$ km.

From the intersection point of curves, depth (h) and length (L) of the vertical finite line are found to be 5.75 and 16.3 km, respectively. Then in Eq. 1, x is taken as 0, and if h and L values calculated above and max gravity value in the anomaly is replaced in this equation, A is found as 602 mGal*Km. The values of A, h, and L computed above are replaced into Eq. 1 to find a new gravity anomaly (dotted line in Fig. 5a). It is seen that the observed and calculated anomalies are very close to each other.

In addition, Mohan *et al.* (1986) found the depth of sphere to be 9.31 km, applying the Mellin transform and assuming that Louga gravity anomaly belongs to a sphere. If this anomaly belongs to a sphere, the curves similar to those in Fig. 4 would be obtained. Thus, this anomaly is determined to be a vertical finite line.

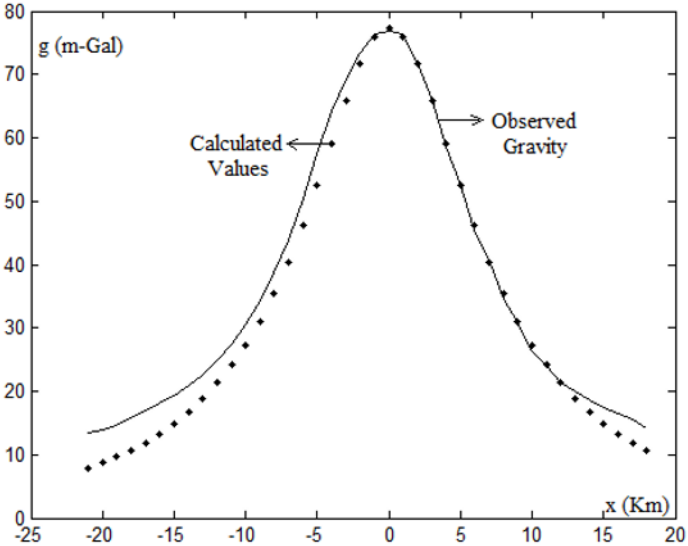


Fig. 5a. North-south Louga gravity anomaly, USA (Nettleton 1976).

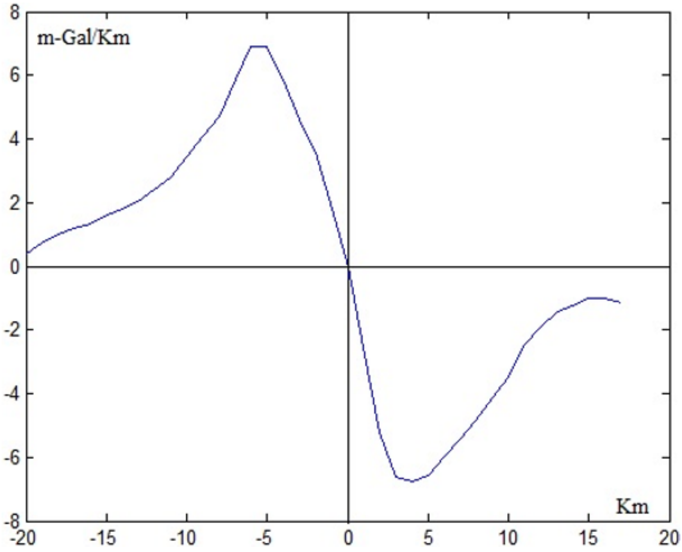


Fig. 5b. Horizontal first derivative anomaly of the Louga gravity anomaly.

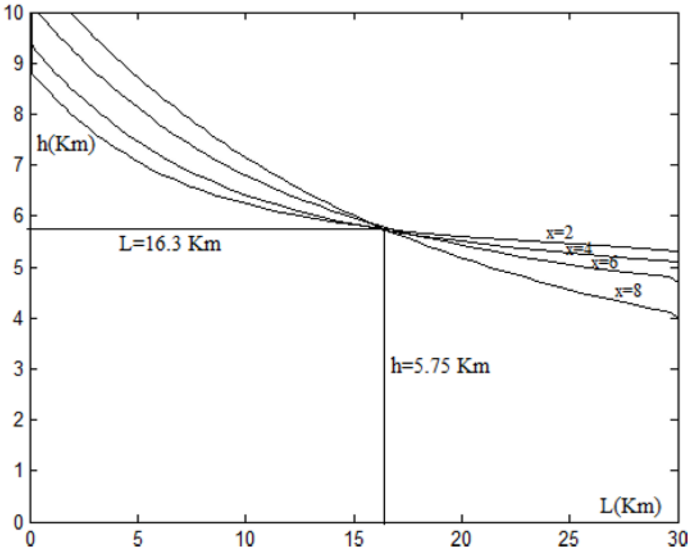


Fig. 5c. Curves obtained by the application of proposed method to the Louga gravity anomaly.

5. CONCLUSION

In this study, depth and length of a vertical finite line are determined by the interpretation of residual gravity anomaly of the source body. In this interpretation, the ratio of measurement values at any observation point to horizontal derivative values at the same point is obtained numerically and then, for various depths and lengths, the ratio of theoretical vertical finite line equation to theoretical horizontal derivative equality is computed. The calculated values are plotted length versus depth and a curve is obtained. This process is repeated for several observation points and several curves are plotted on the same coordinate axis. Intersection point of these curves reflects the real depth and length of vertical finite line.

In order for the method to be completed accurately, regional effects and the noises in the anomaly should be eliminated and the reference level should be determined accurately. Besides, if the anomaly has noises, since it is possible for a mistake to occur, particularly when calculating horizontal derivative values, the anomaly must be smoothed before applying the proposed method. Furthermore, if another body exists near the vertical line, the proposed method may become unsuccessful because the anomaly of this body affects the anomaly of the vertical line.

This method has been tested on two synthetic examples and applied to field anomaly and the results are compared with those of another author who previously assessed this anomaly.

Acknowledgements. We thank Editor and the anonymous reviewers for their detailed and constructive comments which helped to improve this article.

References

Abdelrahman, E.M. (1990), Magnetic interpretation of long horizontal cylinders using correlation factors between successive least-squares residual anomaly profiles, *Pure Appl. Geophys.* **132**, 3, 521-531, DOI: 10.1007/BF00876927.

Abdelrahman, E.M., and T.M. El-Araby (1993), A least-squares minimization approches to depth determination from moving average residual gravity anomalies, *Geophysics* **58**, 12, 1779-1784, DOI: 10.1190/1.1443392.

Abdelrahman, E.M., and H.M. El-Araby (1993), Shape and depth solutions from gravity data using correlation factors between succesive least-squares residual, *Geophysics* **58**, 12, 1785-1791, DOI: 10.1190/1.1443393.

Abdelrahman, E.M., A.I. Bayoumi, Y.E. Abdelhayt, M.M. Gobashy, and H.M. El-Araby (1989), Gravity interpretation using correlation factors between successive least-squares residual anomalies, *Geophysics* **54**, 12, 1614-1621, DOI: 10.1190/1.1442629.

Essa, K.S. (2007a), Gravity data interpretation using the s-curves method, *J. Geophys. Eng.* **4**, 2, 204-213, DOI: 10.1088/1742-2132/4/2/009.

Essa, K.S. (2007b), A simple formula for shape and depth determination from residual gravity anomalies, *Acta Geophys.* **55**, 2, 182-190, DOI: 10.2478/s11600-007-0003-9.

Essa, K.S. (2011), A new algorithm for gravity or self-potential data interpretation, *J. Geophys. Eng.* **8**, 3, 434-446, DOI: 10.1088/1742-2132/8/3/004.

Essa, K.S. (2012), A fast interpretation method for inverse modeling of residual gravity anomalies caused by simple geometry, *J. Geol. Res.* **2012**, 327037, DOI: 10.1155/2012/327037.

Essa, K.S. (2014), New fast least-squares algorithm for estimating the best-fitting parameters of some geometric-structures to measured gravity anomalies, *J. Adv. Res*. **5**,1, 57-65, DOI: 10.1016/j.jare.2012.11.006.

Gay, S.P. Jr. (1965), Standard curves for magnetic anomalies over long horizontal cylinders, *Geophysics* **30**, 5, 818-828, DOI: 10.1190/1.1439656.

Gupta, O.P. (1983), A least-squares approach to depth determination from gravity data, *Geophysics* **48**, 3, 357-360, DOI: 10.1190/1.1441473.

Kara, I., and A.I. Kanli (2005), Nomograms for interpretation of gravity anomalies of vertical cylinders, *J. Balkan Geophys. Soc.* **8**, 1, 1-6.

Mohan, N.L., L. Anandabadu, and R. Seshagari (1986), Gravity interpretation using the Melin transform, *Geophysics* **51**, 1, 114-122, DOI: 10.1190/1.1442024.

Nettleton, L.L. (1942), Gravity and magnetic calculation, *Geophysics* **7**, 3, 293-310, DOI: 10.1190/1.1445015.

Nettleton, L.L. (1976), *Gravity and Magnetic in Oil Prospecting*, McGraw-Hill Book Co.

Odegard, M.E., and J.W. Berg (1965), Gravity interpretation using the Fourier integral, *Geophysics* **30**, 3, 424-438, DOI: 10.1190/1.1439598.

Shaw, R.K., and B.N.P. Agarwall (1990), The application of Walsh transforms to interpret gravity anomalies due to some simple geometrically shaped causative sources: A feasibility study, *Geophysics* **55**, 7, 843-850, DOI: 10.1190/1.1442898.

Received 27 May 2015
Received in revised form 15 January 2016
Accepted 22 February 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2244-2263
DOI: 10.1515/acgeo-2016-0085

Vibration Effect of Near Earthquakes at Different Depths in a Shallow Medieval Mine

Markéta LEDNICKÁ and Zdeněk KALÁB

Institute of Geonics of the CAS, Ostrava-Poruba, Czech Republic
e-mail: lednicka@ugn.cas.cz, kalab@ugn.cas.cz

Abstract

The shallow medieval Jeroným Mine is located at a distance of about 25 km southeast of the Nový Kostel focal zone where the most intensive seismic activity in West Bohemia (Czech Republic) has been documented. Permanent seismological monitoring has been carried out since 2004 in this mine. During the 2011 and 2014 seismic swarms, more than 1000 triggered records comprising almost 1500 earthquakes were recorded at the permanent station in the mine. Three short-term seismological experiments were accomplished during these swarms. Several temporary seismic stations were simultaneously placed in different parts of underground spaces which enabled comparison of vibration effect caused by near earthquakes in different parts of the mine. Although the depth of the lowest parts of mine is only about 60 m, a vibration effect generated by earthquakes from the Nový Kostel focal zone is not the same for the whole underground complex.

Key words: Jeroným mine, vibration, earthquake, seismic swarm, underground space.

1. INTRODUCTION

The Jeroným (Jerome) Mine is located near Čistá in the Sokolov Region, West Bohemia. This medieval mine was declared the National Cultural Heritage Site in 2008 and it represents a part of the European Mining Heritage

Network. The oldest parts of the mine are related to the extraction and processing of tin during the second half of the 16th century. In the underground complex we can find mine workings made by different mining methods such as extracting by a picker and miner's hammer, fire-setting, underhand stoping or overhand stoping, chamber mining, *etc*. Nowadays, small part of the mine is opened to the public as a mining museum. Stability of the mine, the parts of which are more than 400 years old, is the priority both in the light of preserving the unique spaces for next generations and in the light of safety of people visiting the mine. Therefore, a detailed geomechanical monitoring has been performed there including a stability analysis of underground spaces especially of the critical ones. A dynamic loading evaluation of underground spaces is included in the stability analysis and therefore a seismological monitoring is also performed in the mine. Earthquakes from the Nový Kostel focal zone in West Bohemia represent the most significant seismic loading of this mine. According to Fischer *et al.* (2010), limited seismic potential of the focal zone would amount to an earthquake of $M_L \sim 5$.

Permanent seismological monitoring has been carried out since 2004 in the Jeroným Mine. A seismic pillar has been installed in the mine about 30 m below the surface in one of the largest chambers signed as K1. Three modified SM3 seismometers in geographical orientation are anchored on the concrete pillar; the seismic station is signed as JER1. The seismic recording apparatus PCM3-EPC4 (Knejzlík and Kaláb 2002) has a special modification for the high air humidity environment and drip water. Seismic recording is based on a trigger regime. Sampling frequency of digital signal is 250 Hz per channel and the frequency range is 0.5-30 Hz.

Since 2008, more than 2000 earthquakes from West Bohemia were recorded on JER1 station during three intensive seismic swarms in the years 2008, 2011, and 2014. The first analysis of vibration effect in the Jeroným Mine during these seismic swarms was described in the papers by Kaláb and Lednická (2011), Lednická and Kaláb (2013), Lyubushin *et al.* (2014), and Kaláb *et al.* (2015). According to the presented results, the measured vibration velocity values reached up to 0.8 $\mathrm{mms^{-1}}$ at JER1 station for an earthquake with local magnitude M_L 3.6. Based on the measured data and information about local magnitude of recorded earthquakes (according to database of WEBNET network; IG CAS 1991), maximum space component of vibration velocity in the mine of 13 $\mathrm{mms^{-1}}$ was extrapolated for an expected earthquake with local magnitude M_L 5.0. As it is mentioned in the performed studies, maximum vibration velocity expected in the Jeroným Mine should not reach limit values for the first damage underground, as stated in the literature (*e.g.*, Dowding and Rozen 1978). Therefore, it is possible to say that the Jeroným Mine as the whole complex of underground spaces should be stable from the viewpoint of damage caused by vibrations. Nevertheless, determined maximum space component of vibration velocity is valid only for the place of JER1 station. If there will be amplification of vibration effect in other parts of the mine compared to the effect at JER1 station, it will be necessary to perform other detailed studies of vibration effect at these parts and to determine corrected maximum of expected vibration velocity value for stability evaluation.

It is necessary to study vibration effect especially in critical places where movements along discontinuities may occur or places where a deformation of roofs may take place, especially in shallow underground workings. Some movements along discontinuities in the highest levels of the mine were detected within last years. These movements were detected by cracked glass targets installed across the discontinuities. These targets are checked visually, approximately five times per year during irregular inspections of the mine. At present we can only speculate about the time and reason of glass target cracking. Cracked glass targets were detected on the fissured residual pillar in the chamber K4 and on the roof in the highest part of the chamber K1.

During seismic swarms in 2011 and 2014, short-term seismological experiments were made in the mine, enabling to analyse a vibration effect at different parts of the mine caused by near earthquakes (epicentre distances up to 25 km). Five temporary seismic stations (signed ST1-ST5) were placed in three large chambers on different levels of the mine during the experiment in 2011 to cover the different parts of the whole underground system of mine workings. Seismometers had to be placed directly on the rock massif during the measurement to have possibility to compare vibration response of the rock massif (not the response of infilling rock material located in the chambers). Especially in sloping chambers K3 and K1, there is practically no uncovered rock massif at the bottom of the chamber suitable for placing a seismometer. The chambers are filled with the rock blocks from the collapsed roofs and pillars and with the weathered rock material; moreover, the lowest parts of these two chambers are permanently flooded. The lowermost temporary seismic station was situated in the lowest accessible part, which means, on the level of drainage adit. The highest placing of the seismic station was on the same level as the permanent station JER1 but in the different part of the chamber K1 – in a small niche (see Figs. 1 and 2 and Table 1). The difference in height of the lowermost and the highest station was 19 m. Based on the remarks resulting from the measurement in 2011, next experimental measurement was performed during 2014 seismic swarm with temporary seismic station (signed ST6) located in the highest accessible part of K1 chamber near collapsed roof and cracked pillar.

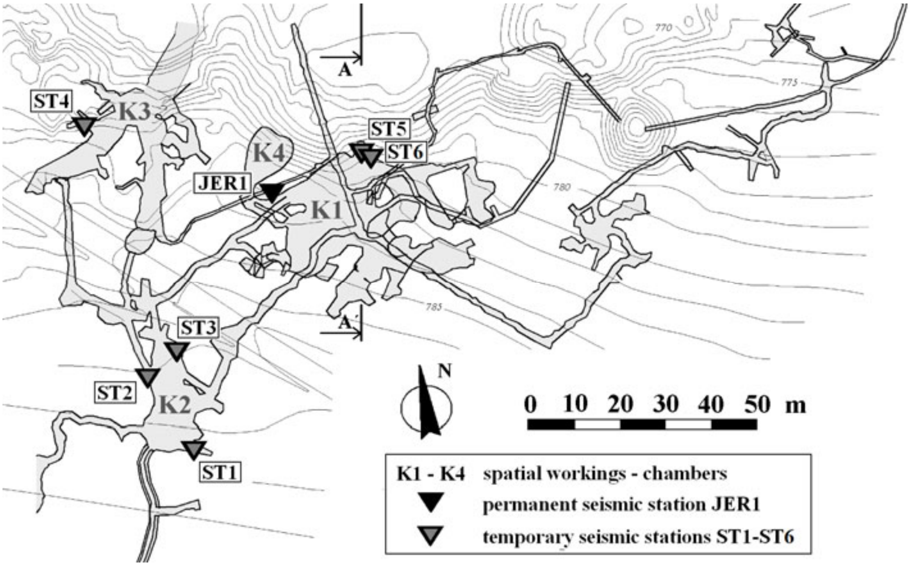


Fig. 1. Sketch of the Jeronýn Mine and locations of seismic stations.

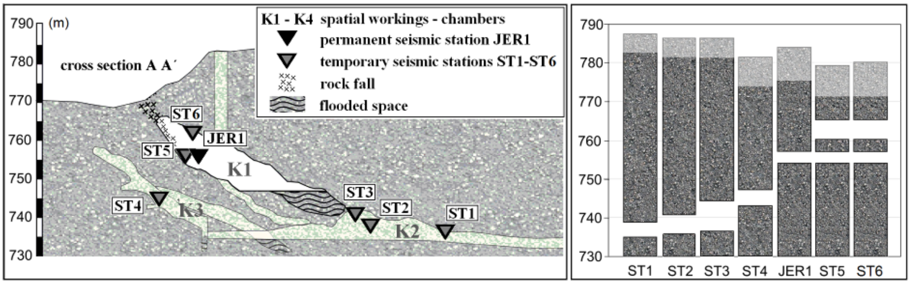


Fig. 2. General cross-section of the Jeroným Mine with position of seismic stations and detail cross-sections in the seismic station places. Light grey colour of near-surface layers on the right figure represents zone of low velocities indicated by geophysical measurement (according to Beneš 2011).

Table 1

Seismic stations

Seismic station	JER1	ST1	ST2	ST3	ST4	ST5	ST6
Altitude above sea level [m]	754	735	736	739	743	754	760
Depth below the surface [m]	30	53	51	48	37	24	20

2. LOCALITY DESCRIPTION

Some detailed information about the Jeroným Mine and mining history in this locality was described in the papers by, *e.g.*, Žůrek and Kořínek (2001/2002), Tomíček (2011). Geological and geomechanical descriptions of this locality can be found in numerous papers (according to Kaláb *et al.* 2006, 2008). From the geological viewpoint, the territory consists of metamorphosed rocks of the Slavkov mantle crystalline complex and of Variscian granites of the Ore Mountains pluton. Rock mass underground had been exposed to devastation and weathering for many years and places with lower stability were formed as a result of those processes (Lednická and Kaláb 2012). Some critical places underground are represented mostly by fissured and weathered supporting pillars or hanging layers on the roof in chambers. On the residual pillars there is probably an influence of recent tectonics induced by anthropogenic factors. These, most likely, tension cracks roughly trace directions of natural tectonics. Some other critical places are represented by ground deformations above the mine where a few collapses occurred due to broken sub-surface rock layers into shallow workings (Kukutsch *et al.* 2011). The sub-surface rock structure is possible to investigate, *e.g.*, through a seismic survey (Arosio *et al.* 2013). The refraction seismic survey was applied also at the locality of the Jeroným Mine and the results of the performed measurements have shown that the rock environment is to a great extent inhomogeneous. The overlying layers above the cavities are at heavily loosened places. This is shown by significant declines of the seismic velocities. The existence of unknown cavities was confirmed by the testing boreholes which were situated on the basis of the geophysical measurement results (Beneš 2011).

The Jeroným Mine consists of an underground system of workings, galleries, shafts, and chambers on at least three horizontal levels. The lowest level is permanently flooded (Fig. 2). Some parts of the underground space have so far been unexplored. All historical maps and documents were destroyed by the fire in the Bureau of Mines in 18th century. Therefore, there is no detailed information about any historical mining activities and a scale of subsurface and underground space in this location. At present, the lowest accessible part of the mine is on the dewatered level of about 55 m below the surface.

3. WEST BOHEMIA SEISMIC SWARMS

A weak seismic activity in the form of seismic swarms is typical for the West Bohemia Region. It means hundreds up to thousands earthquakes occurring over several days or months. An earthquake swarm activity in this area has been monitored by local seismic network WEBNET operated by the

Institute of Geophysics and the Institute of Rock Structure and Mechanics, both coming under the Czech Academy of Sciences, Prague (IG CAS 2000). The oldest observations of earthquakes in the region date back to the Middle Ages. The magnitude of the strongest earthquakes in 1824, 1897-1908, and 1985-2011 did not exceed M_L 4.5 (Fischer *et al.* 2014).

The analysis of the 2011 seismic swarm is presented in the study by Fischer *et al.* (2014). The hypocentre depths in the Nový Kostel focal zone range between 6.5 and 11 km with some clusters down to 13 km. Two sub-clusters, the northern and the southern one, are distinguishable. The 2011 seismic swarm took place in the northern sub-cluster of the Nový Kostel focal zone and the duration of the main swarm period (earthquakes with local magnitude $M_L \geq 2.5$) was only 2 weeks. Almost 1400 earthquakes were recorded at the seismic station JER1 within the phase from 23 August to 11 September. Maximum component vibration velocity recorded at JER1 station reached the 0.35 mms^{-1} during one of the most intensive earthquakes (4 September, M_L 3.3). Epicentres of earthquakes from the Nový Kostel focal zone, which were recorded at the JER1 station during the 2011 seismic swarm, are plotted in Fig. 3.

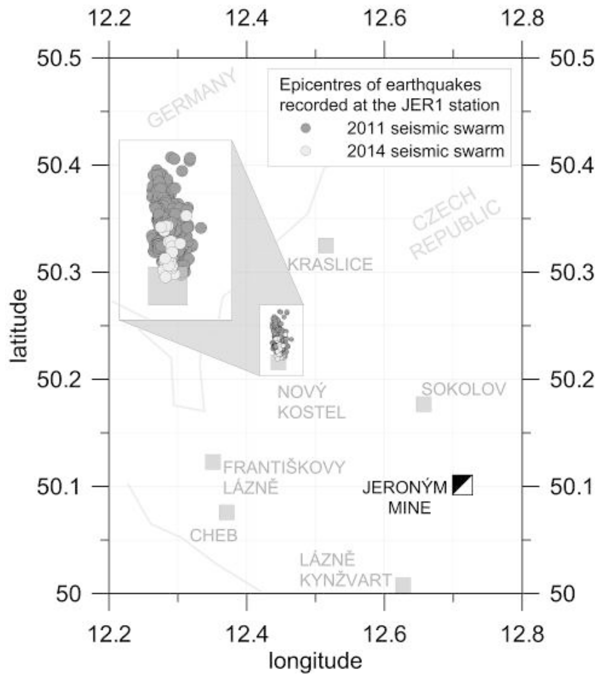


Fig. 3. Epicentres of earthquakes in West Bohemia region occurred during the 2011 and 2014 seismic swarms, which were recorded at the JER1 station; according to the WEBNET seismic network (IG CAS 1991).

The last intensive seismic activity was recorded in 2014 (according to http://www.ig.cas.cz/struktura/observatore/zapadoceska-seismicka-sit-webnet/aktualne-o-seismicke-aktivite-v-zapadnich-c-0). On 24 May 2014, intensive earthquake of local magnitude M_L 3.5 occurred in the Nový Kostel focal zone at a depth of about 9 km. On 31 May 2014, intensive earthquake occurred at 10:37:20 UTC in the same focal zone at a depth of 8.5 km with the local magnitude M_L 4.4. This earthquake counts among the two strongest earthquakes registered in this area over the last 100 years and is comparable to the earthquake in 1985. At night from 3 to 4 August 2014 next intensive earthquake of local magnitude M_L 3.6 occurred in the Nový Kostel focal zone. The depth of the epicentre was approximately 9 km. During this seismic activity, almost 100 earthquakes were recorded at JER1 station. Unfortunately, there was 10 day's gap in the seismic monitoring at JER1 station due to some reconstruction activities in the mine and the most intensive earthquake with M_L 4.4 was not recorded. Maximum component vibration velocity recorded at JER1 station reached the 0.76 mms^{-1} during the M_L 3.6 earthquake (3 August). Epicentres of earthquakes from the Nový Kostel focal zone, which were recorded at the JER1 station during the 2014 seismic swarm, are plotted in Fig. 3.

4. MEASUREMENT

Analysis of short-term measurements (signed as A, B, and C) performed in the mine during the 2011 and 2014 seismic swarms is presented in this paper. Periods of these measurements are illustrated in Fig. 4. The permanent seismic station JER1 was simultaneously in operation and more than 1000 triggered records comprising almost 1500 earthquakes were recorded at this station within these two swarms. The values of maximum component vibration velocity for all recorded earthquakes during the evaluated periods of the seismic swarms in 2011 and 2014 are plotted on the graph in Fig. 4. During one of the strongest earthquakes (25 August 2011, M_L 3.5), the maximum amplitude range of seismic channel ($\pm$ 0.25 mms^{-1}) was exceeded so the maximum component vibration velocity was not possible to be determined for this earthquake.

A short-term measurement A was performed in the Jeroným Mine in order to evaluate a vibration effect in different parts of the mine complex. Five solitary seismic stations (signed as ST1-ST5) with three component sensors were installed in three large chambers of the mine on different levels (Figs. 1 and 2). Some detailed information about the location of individual seismic stations is given in Table 1. This experiment was performed within the period from the afternoon of 30 August to the morning of 1 September 2011. Unfortunately, during this short-term measurement there was a gap in seismic activity and only a few weak earthquakes with local magnitude $M_L \leq 2.5$

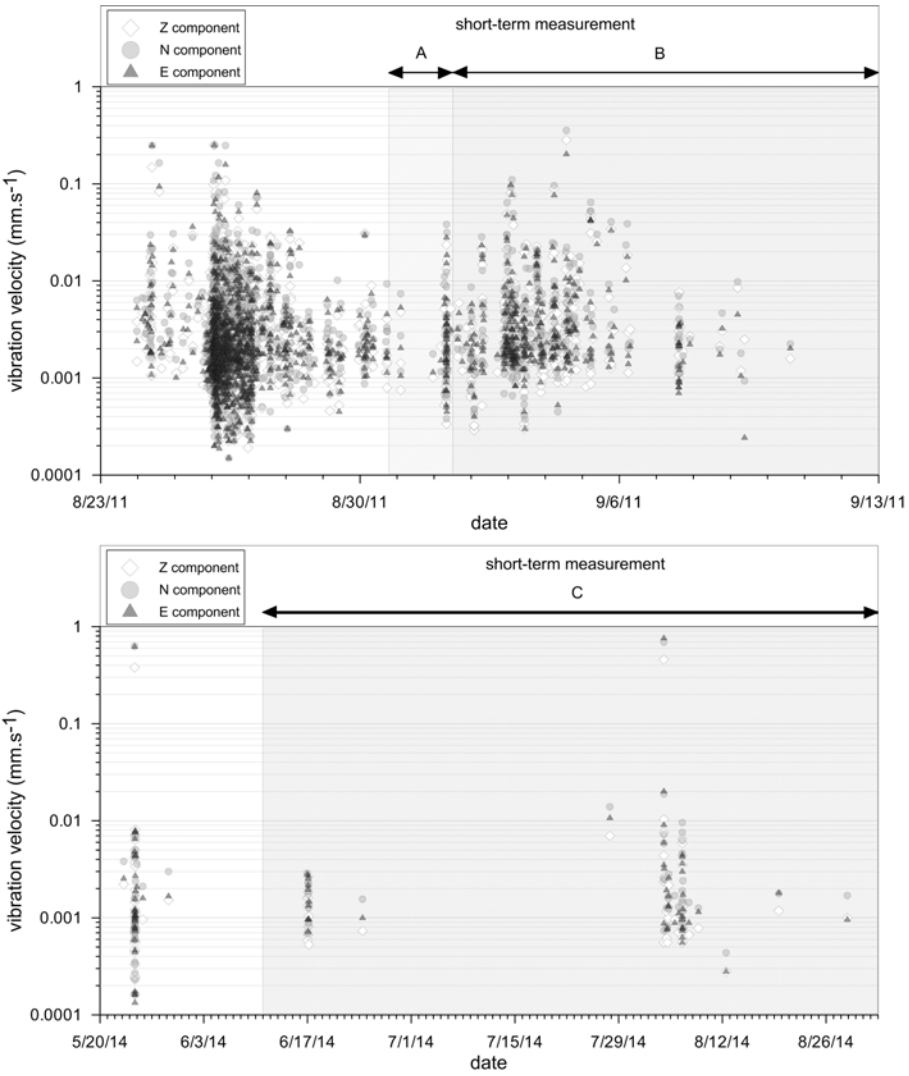


Fig. 4. Chronology of peak ground velocities for all recorded earthquakes at the seismic station JER1; A, B, C – periods of short-term measurements.

were recorded. Only 10 earthquakes (data set A) were possible to be used for elaboration out of this measurement. After this short-term experiment, the station ST1 was kept at the lowest level for subsequent monitoring (measurement B). On 1 September 2011 the next phase of seismic activity started in the Nový Kostel focal zone and within the phase from 1 to 11 September 2011 more than 200 earthquakes were recorded in the mine with local magnitude $M_L \leq 3.3$ (data set B). Third short-term measurement, C, was performed in the mine to extend the knowledge about the vibration effect in the highest parts of the mine. One temporary seismic station (signed as ST6) with three component sensor was placed at the higher level in K1 chamber at a depth of 20 m below the surface; the difference in height of the ST6 and JER1 station was 6 m. This measurement was performed within the period from 11 June 2014 to the beginning of October 2014 and almost 50 earthquakes from the Nový Kostel focal zone were recorded with the local magnitude $M_L \leq 3.6$ (data set C).

5. DATA ANALYSIS

As mentioned above, first analysis of the data recorded at the permanent station JER1 during the seismic swarm in 2011 was presented by Lednická and Kaláb (2013) and the data analysis was focused especially on the evaluation of influence of near earthquakes on the stability of discussed mine. Data from the 2014 seismic swarm are still analysed. In order to summarize main results from the above-mentioned paper we can say that extrapolated values of the component vibration velocity at the station JER1 are in the range from 3.5 to 7.5 mms^{-1} (space component 13 mms^{-1}) for expected local magnitude M_L 5.0. It is necessary to mention that determined equations are derived only for earthquakes from the Nový Kostel focal zone and only for location of the station JER1. Other relations will be necessary to determine if the vibration effect will vary in different parts of the mine, especially for the places where the vibration effect can be higher compared to the effect at the place of the station JER1. Permanent station JER1 was selected as a reference station for all calculations performed in next sections and ratios of peak ground velocities (PGV) from the given station to the reference station were calculated.

5.1 Short-term measurement A

Only 10 weak earthquakes (see Fig. 4) were elaborated with local magnitude $M_L \leq 2.5$ from the Nový Kostel focal zone. The PGV during the phase of *S*-wave (almost all of recorded earthquakes have the maximum value of recorded vibration velocity during the phase of *S*-wave) was determined for each recorded earthquake at each station. PGV ratios from the given station to the reference station were then calculated and elaborated by using box-whisker plots. The results show the lowest vibration effect at the station ST1 on the adit level (Fig. 5), *i.e.*, half the effect at JER1 station. It is possible to detect a slight increase of the vibration effect at the ST2 and subsequently at the ST3 station in the chamber K2. The vibration effect in the central part of the chamber K3 is approximately 1.3 times lower than on the JER1 station and it is slightly different for all three components. At the station ST5 in the

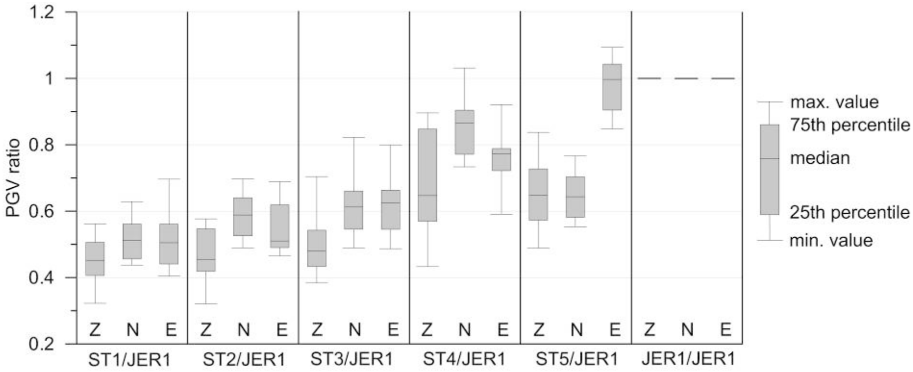


Fig. 5. The PGV ratios of the given stations to the reference station calculated for each component (data set A).

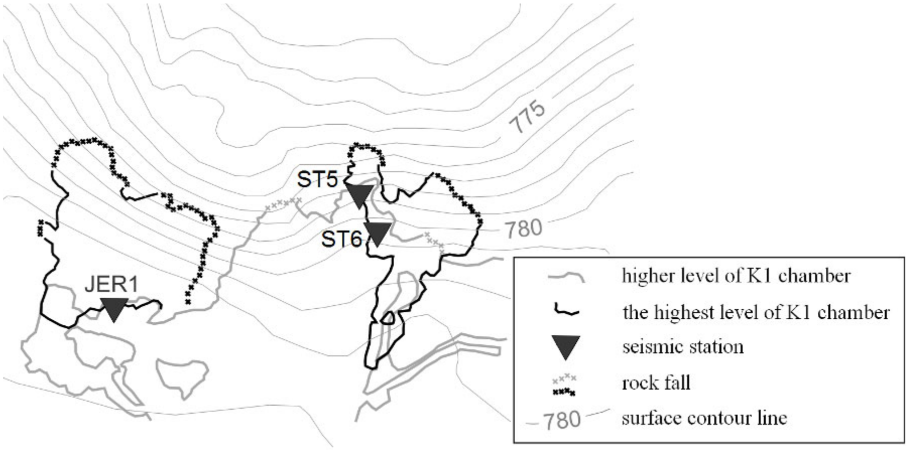


Fig. 6. Detailed situation in the surroundings of the ST5, ST6, and the JER1 stations.

chamber K1, the PGV ratio is much more different for all three components in comparison with PGV ratios at the other stations. The median of calculated PGV ratio is 1.0 at E component and only 0.65 at N and Z components. The station ST5 was situated in the part of the chamber K1, where a complex structure of pillars and hanging roofs is located. Above the station ST5 and in its surroundings there are located rock-falls leading from higher collapsed parts of the mine (Fig. 6). It may cause this specific character of the vibration effect in this place. One example of wave patterns for N components recorded during measurement A is in Fig. 7.

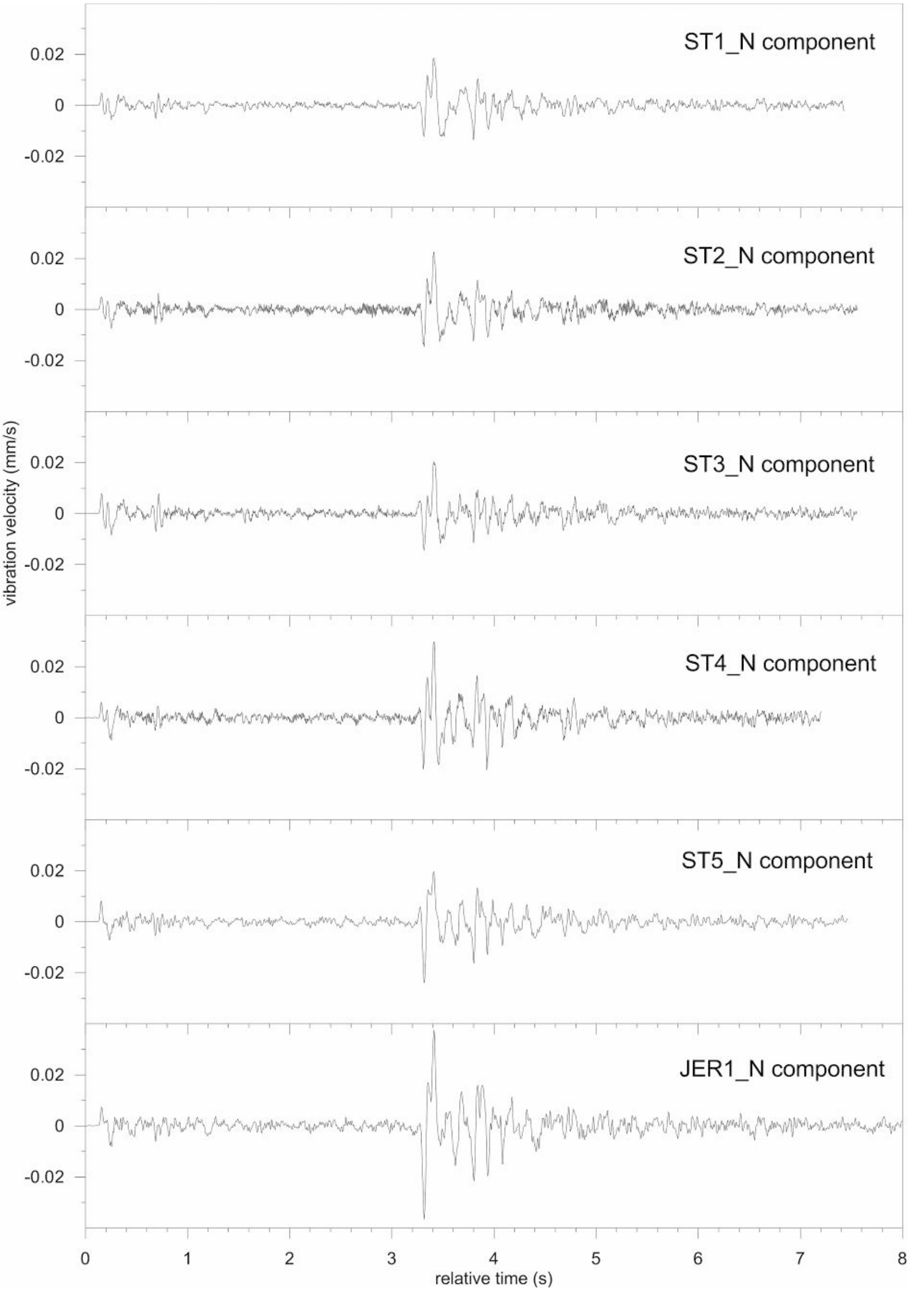


Fig. 7. Example of wave patterns of recorded earthquake (1 September 2011, M_L 2.5) for N components simultaneously measured on six seismic stations during short-term measurement A.

5.2 Short-term measurement B

More than 200 earthquakes with $M_L \leq 3.3$ were recorded at the permanent station JER1 and temporary station ST1 within the phase from 1 to 11 September 2011. The PGV ratio was calculated and analysed using box-whisker plots (Fig. 8). This graph demonstrates that the vibration effect is similar for all three components. The vibration effect at the adit level is also approximately half the effect at the station JER1. These results correspond to the re-

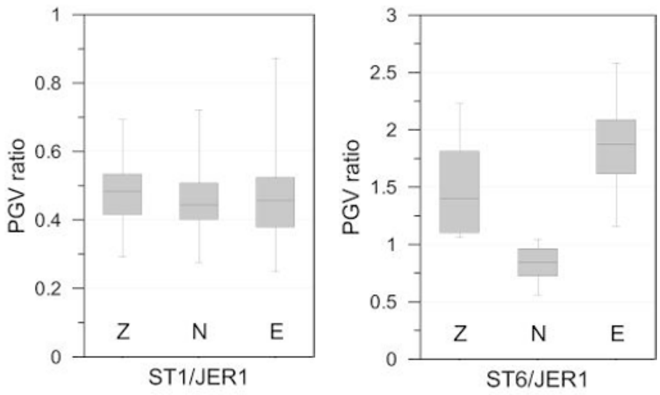


Fig. 8. The PGV ratios ST1/JER1 and ST6/JER1 for all components (data set B – left and data set C – right).

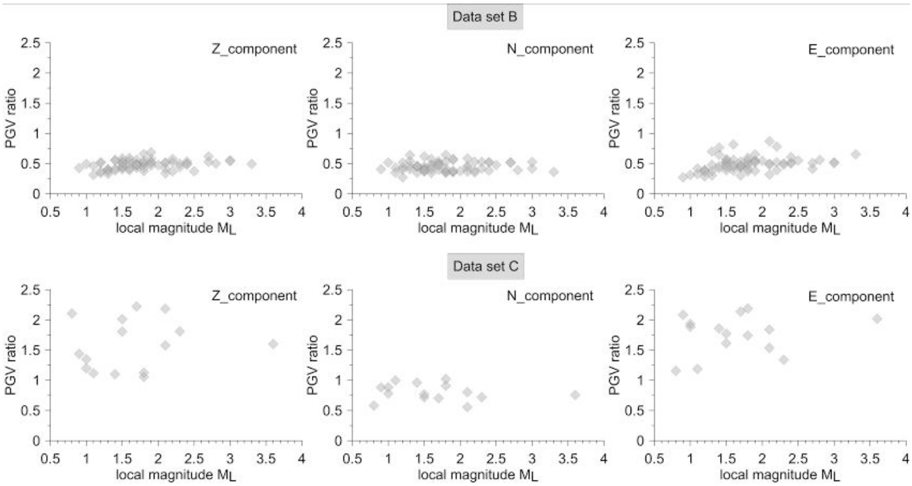


Fig. 9. The PGV ratios ST1/JER1 and ST6/JER1 for data set B and C depending on local magnitude.

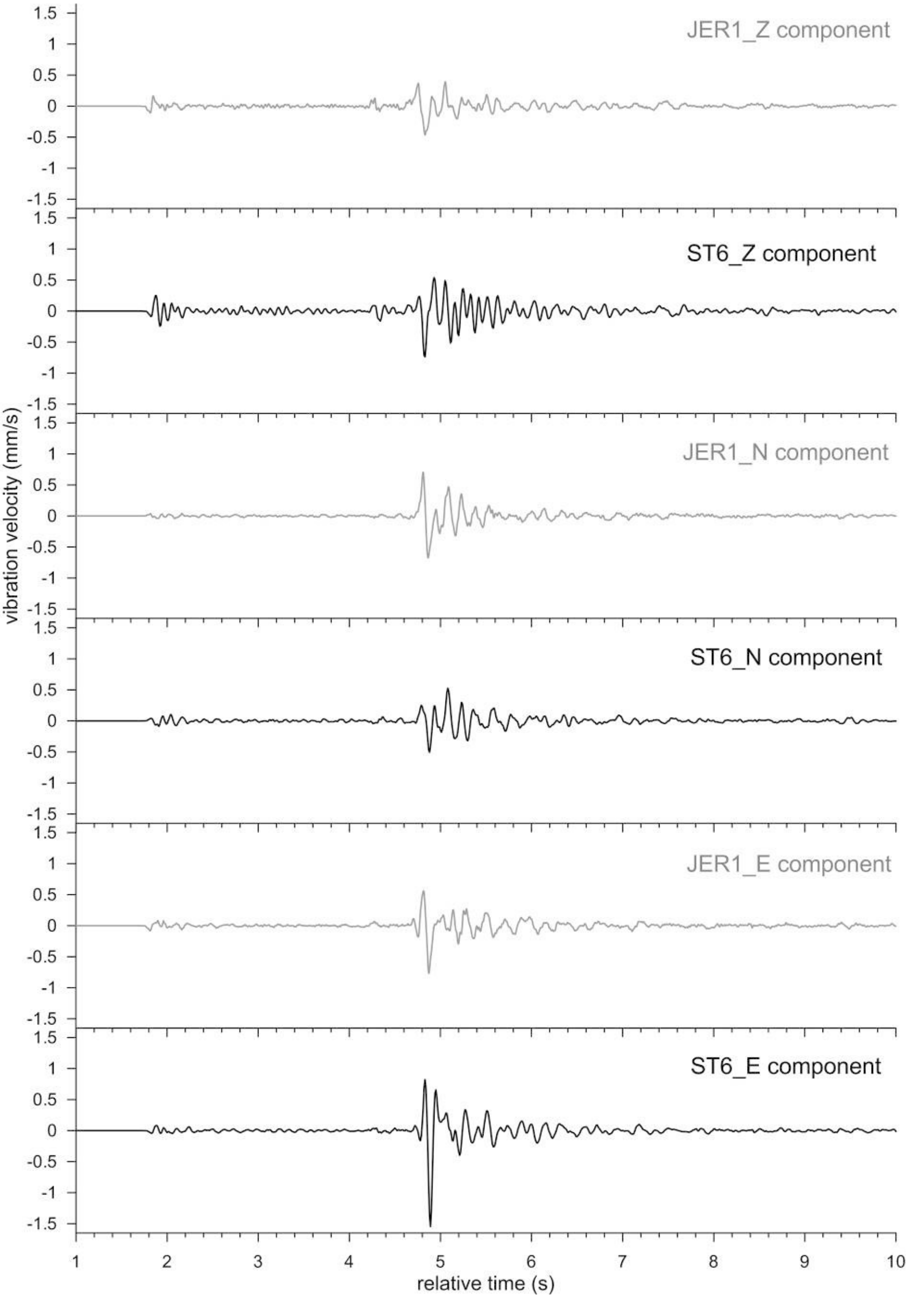


Fig. 10. Example of wave patterns of the recorded earthquake (3 August 2014, M_L 3.6) for Z, N and E components simultaneously measured on two seismic stations during short-term measurement C.

sults of a short-term measurement A, when only 10 earthquakes were elaborated. In order to be able to find out whether the calculated ratios are the same for earthquakes with different magnitudes, we have plotted the PGV ratio depending on local magnitude (Fig. 9). The information on local magnitude of earthquakes was possible to find in the map "Epicentres of earthquakes in the West Bohemia – Vogtland region" (http://www.ig.cas.cz/en/structure/observatories/webnet/map-of-epicenters/). This analysed data set contains 65 earthquakes ($M_L \leq 3.3$). The PGV ratio calculated for the stations JER1 and ST1 is generally the same for the whole evaluated range of local magnitudes M_L from 0.9 to 3.3.

5.3 Short-term measurement C

Almost 50 earthquakes with $M_L \leq 3.6$ were recorded at the permanent station JER1 and temporary station ST6 within the phase from 16 June to 28 August 2014. As well as for previous measurement, B, the PGV ratio was calculated and analysed using box-whisker plots (Fig. 8) and plotted depending on local magnitude (Fig. 9, selected data are presented here). The results in graph in Fig. 8 demonstrate that the vibration effect in the place of ST6 station is not the same for all three components compared to the JER1 station. PGV ratio is 1.4 and 1.8 for the Z and E components, respectively, but the PGV ratio is only 0.8 for the N component. The PGV ratio is generally the same for the whole evaluated range of local magnitudes. An example of wave patterns of all three components recorded simultaneously at JER1 and ST6 seismic stations is presented in Fig. 10 (earthquake with M_L 3.6).

6. DISCUSSION

According to the results of analysis, we can state that the vibration effect in the Jeroným Mine caused by earthquakes from the Nový Kostel focal zone is different for individual investigated parts of the mine complex. The lowest vibration effect was determined at the lowest accessible part at the adit level in the K2 chamber, *i.e.*, 53 m below the surface (ST1 seismic station). The most intensive vibration effect was detected on temporary seismic station ST6 in the higher part of K1 chamber, *i.e.*, 20 m below the surface.

It is known that the vibration effect underground is less than at the surface and that it decreases with increasing depth. Evaluation of vibration effect in different depths is often made, for example in geotechnical engineering, earthquake engineering and/or speleoseismology (Bokelmann and Gribovszki 2015, Szeidovitz *et al.* 2008). Hu and Xie (2004) investigated the variation of ground motion with depth for three different site conditions. They concluded that peak ground acceleration, velocity and displacement decrease with depth and the decline extent is more dramatic in shallower layers than that in deeper ones. In the paper, fitting curves calculated for subsurface/surface amplitude ratio are presented for investigated sites. For rock site, for example, the PGA (peak ground acceleration) on the surface is 1.2 times higher than at a depth of 100 m. For soil site, PGA on the surface is 1.8 times higher and for soil/rock site PGA is 3.3 times higher than at a depth of 100 m. They also found that the reduction of amplitude with depth is affected by the magnitude and site geology. Iwasaki *et al.* (1977) analysed acceleration measured in the boreholes at the different depths. Acceleration, with respect to depths, changes considerably with the change of soil conditions near the ground surface. Ratios of the surface acceleration to that at the deeper layer (110 to 150 m) are about 1.5 at a rocky ground, 1.5 to 3.0 at sandy ground and 2.5 to 3.5 at a very clayey ground.

Above the Jeroným Mine, no seismic station was placed on the surface during the measurements and, therefore, it was not possible to calculate PGV ratio between places underground and on the surface. Permanent station JER1 was used as the reference place, so the PGV ratio is equal to 1.0 at a depth of 30 m under the surface. Amplification and/or decreasing of vibration effect at other places were compared with this reference depth. The PGV ratio (median value from the data set A and C) depending on depth under the surface is presented in Fig. 11. The PGV at a depth of 53 m is approximately half of the PGV at a depth of 30 m below the surface. The PGV ratio calculated for the stations ST2, ST3, and ST4 is increasing with decreasing depth under the surface and it seems that the data are in linear correlation. Conversely, the PGV ratio calculated for the ST5 and ST6 stations is not in correlation with results from other stations. This effect may be caused, for example, by the local geological and geomechanical conditions at the place of ST5 and ST6 stations where complex structure of supporting pillars is situated and collapsed overburden is located (details on Fig. 6). It means that anomalous vibration effect may be regarded as a response of geological element in the mine. Therefore, it is necessary to perform detailed geomechanical evaluation of this place.

The results obtained from these seismological experiments provide important basis of seismic loading evaluation of this medieval mine. Damage to mines is more insignificant when they are located in highly competent, unweathered rock; the greatest damage occurs in mines found in loose unconsolidated or incompetent material. This is due to the effect of decreased vibration in competent rock; unconsolidated sediment is much more susceptible to damage caused by vibration. Large displacements occur primarily along pre-existing faults and fractures. Usually, more detailed analysis is required before seismic criteria can be formulated for the safety of people visiting the mine.

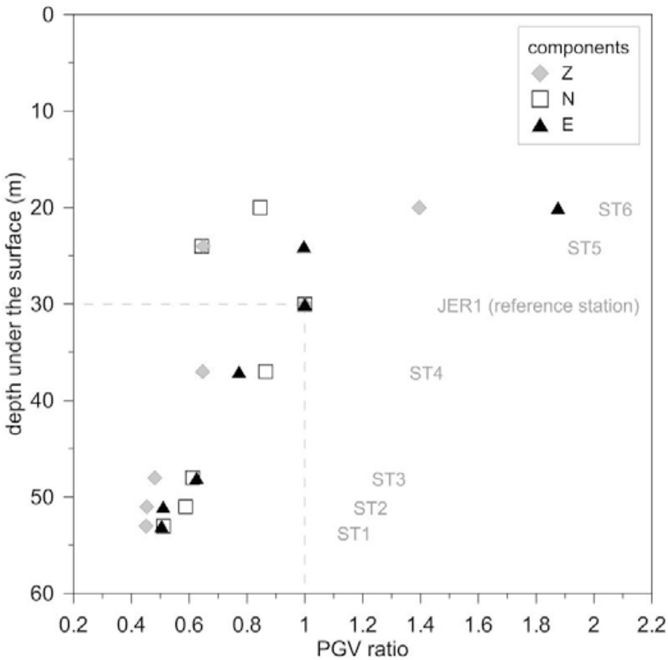


Fig. 11. The PGV ratios depending on depth under the surface; depth of 30 m below the surface represents the reference depth for PGV ratio calculation.

As it was mentioned above, maximum vibration velocity of 13 mms^{-1} expected in the Jeroným Mine at the place of JER1 station for an earthquake with local magnitude M_L 5.0 should not reach limit values for the first damage underground as stated in the literature (*e.g.*, Dowding and Rozen 1978). Nevertheless, this extrapolated value is valid only for the place of station JER1 at a depth of 30 m below the surface. Considering the results presented in Fig. 11 and the fact that the PGV decrease is more dramatic in shallower layers, especially in the weathered and collapsed parts of rock massive, the vibration effect can be more intensive in the highest levels of the mine and the expected maximum vibration velocity could reach the limit values for the first damage. Measurement on the station ST6 in 2014 confirmed this idea and it was documented that the vibration effect at the higher measured parts of K1 chamber is almost two times higher than the effect at the JER1 station in the east-west direction. Looking at the N component, the PGV value is lower than value measured at station JER1 (Fig. 11). This effect documents that the vibration field in underground spaces of the Jeroným mine is complex and anomalous vibration effect with increasing maximum PGV may occur at some places, especially in the highest levels near the collapsed parts and cracked pillars. It is not possible to perform the evaluation of vibration effect in this type of underground structure based on measurement at only one point. To confirm this idea, next measurement is necessary to be made at other places too, especially at the highest parts of the mine.

On the other side, results confirmed that the maximum expected vibration velocity will be approximately half the value of JER1 for chamber K2 and consequently all galleries and other workings at the adit level. The maximum expected vibration velocity values underestimate significantly limit values for the first damage caused by vibrations at these places. This information is very important from the stability point of view, because all the important galleries are located at the adit level – galleries for the entering all chambers, galleries for ventilation and dewatering of all accessible and/or inaccessible workings.

On the contrary, at the highest levels of the mine, it means chambers K3, K1, and K4, expected vibration velocities could be higher than the maximum expected PGV value of 13 mms^{-1}. This is very valuable information for stability assessment and also for planning of next geomechanical, geotechnical, and also seismological monitoring in this mine (Kaláb and Lednická 2016).

7. CONCLUSION

The paper presents the results of vibration effect evaluation in different parts of underground mine complex. Vibration effect was measured in different parts of shallow medieval Jeroným Mine, which is located in West Bohemia, Czech Republic. About 25 km from the mine, there is located the Nový Kostel focal zone, which is known as an area with weak seismic activity in the form of seismic swarms. During the 2011 and 2014 seismic swarms in the Nový Kostel focal zone, more than 250 earthquakes with local magnitude $M_L \leq 3.6$ were recorded in the underground spaces during performed short-term seismic measurements. One permanent and six temporary seismic stations were used for the measurement and data analysis. These stations were placed in different parts of the mine at the depth ranging from 20 to 53 m below the surface. Permanent station JER1 installed 30 m below the surface has been used for the seismic monitoring, for the evaluation of vibration effect in the mine and for the stability analysis since 2004. Results of experimental measurements performed using temporary seismic stations enabled to compare vibration effect at different parts of the mine. It was confirmed based on the obtained results that the vibration effect is higher at the highest places of the mine in comparison with vibration effect at the permanent station JER1. Based on previous studies, maximum vibration velocity for maximum expected earthquake of local magnitude M_L 5.0 was extrapolated based on long term monitoring at the permanent JER1 station that underestimates limit values for the first damage underground caused by vibrations.

Nevertheless, results obtained based on presented experimental measurement at different parts of the mine confirmed that new analysis has to be performed at the highest places of the mine where cracked and weathered supporting pillars and collapsed overburden are located. At these places, anomalous vibration effect can be estimated and the maximum expected vibration velocity could reach the limit values for the first damage underground.

From the stability point of view, it is necessary to add that the Jeroným Mine, as the whole complex of underground spaces, should be stable from the viewpoint of damage caused by vibrations. Maximum expected vibration velocity at the lowest parts of the mine at the adit level underestimates significantly the limit values for the first damage underground. But according to the results, at selected critical places, it is necessary to perform specific measurement and to analyse these places individually from the stability point of view, especially at shallow depths and near the collapsed parts. This information is very important for safety of people attending underground spaces and also for assessing the stability of these places and for subsequent stabilization of the most hazardous parts.

Acknowledgments. This research has been performed with the financial support of the long-term conceptual development support of research organisations RVO: 68145535 and, initially, of the Czech Science Foundation, No. 105/09/0089 “A prognosis of time-space changes in the stability of mine cavities of industrial heritage site, Jeroným Mine at Čistá”. The data acquisition was supported by the project of large research infrastructure CzechGeo/EPOS, Grant No. LM2010008.

References

Arosio, D., L. Longoni, M. Papini, and L. Zanzi (2013), Seismic characterization of an abandoned mine site, *Acta Geophys*. **61**, 3, 611-623, DOI: 10.2478/s11600-012-0090-0.

Beneš, V. (2011), Geophysical investigation above Jeroným mine, *Int. J. Explor. Geophys. Remote Sens. Environ. (EGRSE)* **18**, 1, 40-49 (in Czech).

Bokelmann, G., and K. Gribovszki (2015), Constraints on long-term seismic hazard from vulnerable stalagmites. **In:** *CSNI Workshop on "Testing PSHA Results and Benefit of Bayesian Techniques for Seismic Hazard Assessment", Pavia, Italy*, 1-7.

Dowding, C.H., and A. Rozen (1978), Damage to rock tunnels from earthquake shaking, *J. Geotech. Eng. Div*. **104**, GT2, 15-20.

Fischer, T., J. Horálek, J. Michálek, and A. Boušková (2010), The 2008 West Bohemia earthquake swarm in the light of the WEBNET network, *J. Seismol.* **14**, 4, 665-682, DOI: 10.1007/s10950-010-9189-4.

Fischer, T., J. Horálek, P. Hrubcová, V. Vavryčuk, K. Bräuer, and H. Kämpf (2014), Intra-continental earthquake swarms in West-Bohemia and Vogtland: A review, *Tectonophysics* **611**, 1-27, DOI:10.1016/j.tecto.2013.11.001.

IG CAS (1991), Institute of Geophysics of the Czech Academy of Sciences: West Bohemia Local Seismic Network, International Federation of Digital Seismograph Networks, Other/Seismic Network, DOI: 10.7914/SN/WB.

IG CAS (2000), *Studia Geophysica et Geodaetica*, Special issue, **44**, 2 and 4, Institute of Geophysics of the Czech Academy of Sciences, Prague, Czech Republic.

Hu, J., and L. Xie (2004), Variation of earthquake ground motion with depth, *Acta Seismol. Sin.* **18**, 1, 72-81, DOI: 10.1007/s11589-005-0008-x.

Iwasaki, T., S. Wakabayashi, and F. Tatsuoka (1977), Characteristics of underground seismic motions at four sites around Tokyo Bay. **In:** H.S. Lew (ed.), *"Wind and Seismic Effects", Proc. Eighth Joint Panel Conf. of the U. S. – Japan Cooperative Program in Natural Resources*, Nat. Bur. Stand. (U.S.), Sp. Publ. 477, U.S. Government Printing Office, Washington, III-41–III-56.

Kaláb, Z., and M. Lednická (2011), Seismic loading of medieval Jeroným Mine during West Bohemia swarm in 2008. **In:** A.F. Idziak and R. Dubiel (eds.), *Geophysics in Mining and Environmental Protection*, Geoplanet: Earth and Planetary Science, Springer-Verlag, Berlin-Heidelberg, 21-29, DOI: 10.1007/978-3-642-19097-1_3.

Kaláb, Z., and M. Lednická (2016), Long-term geomechanical observation in the Jeroným Mine, *Acta Geophys.* **64**, 5, DOI: 10.1515/acgeo-2016-0054.

Kaláb, Z., J. Knejzlík, R. Kořínek, and P. Žůrek (2006), Cultural monument Jeroným Mine, Czech Republic – Contribution to the geomechanical stability assessment, *Publs. Inst. Geophys. Pol. Acad. Sc.* **M-29**, 395, 137-145.

Kaláb, Z., J. Knejzlík, R. Kořínek, R. Kukutsch, M. Lednická, and P. Žůrek (2008), Contribution to experimental geomechanical and seismological measurements in the Jeroným Mine, *Acta Geodyn. Geomat.* **5**, 2, 213-223.

Kaláb, Z., M. Lednická, T. Kaláb, and J. Knejzlík (2015), Evaluation of vibration effect in shalow mine caused by natural and technical seismicity. **In:** *Proc. 15th Int. Multidisciplinary Scientific GeoConf. SGEM 2015, Science and Technologies in Geology, Exploration and Mining, Vol. III, 18-24 June 2015, Albena, Bulgaria*, 855-862, DOI: 10.5593/sgem2015B13.

Knejzlík, J., and Z. Kaláb (2002), Seismic recording apparatus PCM3-EPC, *Publs. Inst. Geophys. Pol. Acad. Sc.* **M-24**, 340, 187-194.

Kukutsch, R., P. Žůrek, and V. Hudeček (2011), Study of underground workings at small depths below the surface, *Int. J. Explor. Geophys. Remote Sens. Environ. (EGRSE)* **18**, 1, 83-91 (in Czech).

Lednická, M., and Z. Kaláb (2012), Evaluation of granite weathering in the Jeroným Mine using non-destructive methods, *Acta Geodyn. Geomater.* **9**, 2, 211-220.

Lednická, M., and Z. Kaláb (2013), Vibration effect of earthquakes in abandoned medieval mine, *Acta Geod. Geophys*. **48**, 3, 221-234, DOI: 10.1007/s40328-013-0018-4.

Lyubushin, A.A., Z. Kaláb, and M. Lednická (2014), Statistical properties of seismic noise measured in underground spaces during seismic swarm, *Acta Geod. Geophys*. **49**, 2, 209-224, DOI: 10.1007/s40328-014-0051-y.

Szeidovitz, G., I. Paskaleva, K. Gribovszki, K. Kostov, G. Surany, P. Varga, and G. Nikolov (2008), Estimation of an upper limit on prehistoric peak ground acceleration using the parameters of intact speleothems in caves situated at the western part of Balkan Mountain Range, North-West Bulgaria, *Acta Geod. Geophys. Hu*. **43**, 2-3, 249-266, DOI: 10.1556/AGeod.43.2008.2-3.13.

Tomíček, R. (2011), Historical and cultural review of the Jeroným Mine, *Int. J. Explor. Geophys. Remote Sens. Environ. (EGRSE)* **18**, 1, 14-22 (in Czech).

Žůrek, P., and R. Kořínek (2001/2002), Opening of the medieval Jeroným Mine in the Czech Republic to the public, *J. Min. Geol. Sci*. **40-41**, 51-72.

Received 22 July 2015
Received in revised form 28 December 2015
Accepted 23 February 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2264-2288
DOI: 10.1515/acgeo-2016-0015

Improved Data Preprocessing Algorithm for Time-Domain Induced Polarization Method with Digital Notch Filter

Shuang-Chao GE[1,2], Ming DENG[2], Kai CHEN[2], Bin LI[3], and Yuan LI[2]

[1]North University of China, Taiyuan, Shanxi, China; e-mail: GeShCh@nuc.edu.cn

[2]Key Laboratory of Geo-detection Ministry of Education,
China University of Geosciences, Beijing, China;
e-mail: dengming@cugb.edu.cn (corresponding author)

[3]Technical Department TSLC, Taiyuan, Shanxi, China; e-mail: msepaper@163.com

Abstract

Time-domain induced polarization (TDIP) measurement is seriously affected by power line interference and other field noise. Moreover, existing TDIP instruments generally output only the apparent chargeability, without providing complete secondary field information. To increase the robustness of TDIP method against interference and obtain more detailed secondary field information, an improved data-processing algorithm is proposed here. This method includes an efficient digital notch filter which can effectively eliminate all the main components of the power line interference. Hardware model of this filter was constructed and Vhsic Hardware Description Language code for it was generated using Digital Signal Processor Builder. In addition, a time-location method was proposed to extract secondary field information in case of unexpected data loss or failure of the synchronous technologies. Finally, the validity and accuracy of the method and the notch filter were verified by using the Cole-Cole model implemented by

SIMULINK software. Moreover, indoor and field tests confirmed the application effect of the algorithm in the fieldwork.

Key words: power line interference, digital notch filter, decay curve, apparent chargeability, SIMULINK.

1. INTRODUCTION

The induced polarization (IP) method was discovered more than 100 years ago by Conrad Schlumberger in 1913. The first measurements made using the IP method can be traced to the 1950s (Marshall and Madden 1959). After 60 years of development, two types of IP methods can now be used to collect data: spectral induced polarization (SIP) method and time-domain induced polarization (TDIP) method. TDIP measurement is characterized by low cost, high efficiency, and resistance to electromagnetic coupling. For these reasons, it has found wide use in mining exploration (Li *et al.* 2013), environmental and geohydrological detection (Gazoty *et al.* 2012, Martinho *et al.* 2006), and other geophysics detection applications. Hördt *et al.* (2007) estimated hydraulic conductivity in unconsolidated sediments from IP method to deduce the permeability. Dahlin *et al.* (2002) worked on the comparison between measurements with non-polarizable electrodes and stainless steel electrodes, and removed the spontaneous potential (SP) component from the record gained by the latter by subtracting the polarisation potential measured when no primary current and no IP signal were present. As a result, the observations cost of TDIP method could be cut down.

However, the disturbance rejection capability of existing TDIP instruments is insufficient. Use of TDIP is greatly limited by the effects of power line interference, various kinds of noises, and synchronous technology. Power line interference is a crucial source of noise in geophysical exploration with a fundamental frequency of 50 or 60 Hz (in China the fundamental frequency is 50 Hz) and a multi-harmonic sinusoidal disturbance caused by an unknown memory less nonlinearity. To remove power line interference, almost all TDIP instruments made in China (Fan and Yang 1979, Feng and Fu 1994, Wu and Dou 1996, Zhang 2009) are implemented using a hardware twin-T filter, which is not sufficient to reduce strong interference to a satisfactory level, especially the harmonic components. Moreover, this filter has electronic circuit inertia which will lead to distortion of the measured wave (Deng and He 1998).

Researchers provided different techniques to reduce this interference, including adaptive methods (Dhillon and Chakrabarti 2001, Kohno *et al.* 1990, Wan *et al.* 2006), empirical mode-decomposition methods (Nimunkar and Tompkins 2007, Su and Chuang 2013), and wavelet transformation filters (Galiana-Merino *et al.* 2013, Sahambi *et al.* 1997). Most of the above-mentioned methods were designed for electrocardiograph (ECG) signal processing. Butler and Russell (1993) provided two methods to suppress the interference in the seismoelectric record by subtracting a noise estimate, in which block subtraction used an interval with obvious power line interference but negligible nonharmonic components as the subtrahend, and sinusoid subtraction used sinusoids with the parameters appropriated from the record as the subtrahend. The latter method was improved to be more applicable by overcoming the limitations of window length (Butler and Russell 2003), but there are still some limitations that the fundamental frequency must be very accurately known. Xia and Miller (2000) provided a hum filter by using damped least-squares Levinson–Marquardt inversion to approximate the amplitudes and phases of multiple harmonic components simultaneously. This technology was very effective for power line interference with stable amplitude and frequency, but the instability of the interference during the measurement for a specific trace would affect the effectiveness. It may not accurately be determined by power line noise in a pre-first arrival window if there are other sources of noise. Pankratov and Geraskin (2010) suppressed the industrial noise using a Finite Impulse Response (FIR) filter for SIP measurement, the phase delay of which was negligibly small compared to the typical anomalous value of the phase difference. But the suppression of the harmonics was not described in their work. Warden *et al.* (2012) used the curvelets technology to distinguish the interfering frequency from the signals, and results showed that it was a perfect time-frequency local analysis tool. But the premise was to select an appropriate mother wavelet.

These methods gained well application effect in geophysical data processing, especially for signal extraction and harmonic suppression in seismoelectric record. However, there are limitations on the execution of above-mentioned methods for their restrictions. And, their interference suppression capabilities and amplitude-frequency properties are not appropriate for TDIP requirements. Because the observation of TDIP is constantly changing electric field, it is inaccessible to obtain appropriate subtraction for the primary field and secondary field or accurately estimate the parameters of power line interference. Besides, the complex arithmetic of these filtering methods may affect real-time performance, making a simple and efficient notch filter essential. A notch filter has a certain stop-band width which could cope with the frequency fluctuation, and stop-band attenuation could be very large. As for the undesirable effects of the notch filter, the signal used in TDIP method generally has low frequency, so the disadvantage is negligible. Thus, we designed a digital notch filter to restrain the power line interference for TDIP method. It offers the advantage of simplicity and applicability, and could be helpful to make effective use of the dynamic range of the recording system.

Another issue is that most of the previous TDIP measurements considered only polarizability and chargeability, and only some discrete data in small windows were collected, resulting in an incomplete description of the observed voltage decay. However, the IP decay curve contains much more information than chargeability (Tong *et al.* 2006). In addition, the shape of the IP decay curve contains spectral information of the IP phenomenon that can be extracted only by using the full decay in the inversion process (Fiandaca *et al.* 2012). Moreover, observation of entire time series is necessary to resolve the Cole-Cole parameters accurately (Swift Jr. 1973, Hönig and Tezkan 2007).

To address these deficiencies and shortcomings, an improved data-processing algorithm with a digital notch filter is proposed here. The notch filter is a FIR filter, easy to be implemented in hardware. Comparison tests showed that the parameters calculated by using the raw data without filtering were easily distorted by ambient noise and synchronization exception, but the results yielded by using the notch filter were significantly better than that without filtering, and more accurate to the theoretical values than the results obtained by using other filtering method. In practical tests, the complete decay curve and apparent chargeability were obtained in the situation of strong interference and synchronization failure by using the algorithm with this notch filter. A complete decay curve gives access to new applications of TDIP in environmental and hydrogeophysical investigations by extracting more detailed relaxation time spectrum and IP transients directly from TDIP (Tong *et al.* 2006, Fiandaca *et al.* 2012). This paper includes a detailed introduction to the digital notch filter design and modeling, the secondary field time-orientation algorithm, and de-noising algorithms. Finally, the effectivity of the filter and data-processing method were tested by a series of tests.

2. THEORY OF THE TDIP METHOD

Usually, TDIP uses rectangular current pulses of opposite polarity as an excitation source, in which pulses and pauses of the same duration are used with a 50% duty cycle (Fig. 1a). Immediately after the current is turned on, an induced potential, ΔV_2, is raised across the potential electrodes. After a charge-up effect, the primary voltage, ΔV_1, is measured just before the current is turned off to compute the direct current (DC) resistivity. When the current is turned off, the voltage drops to a secondary level and then decays with time during the relaxation period (Fig. 1b).

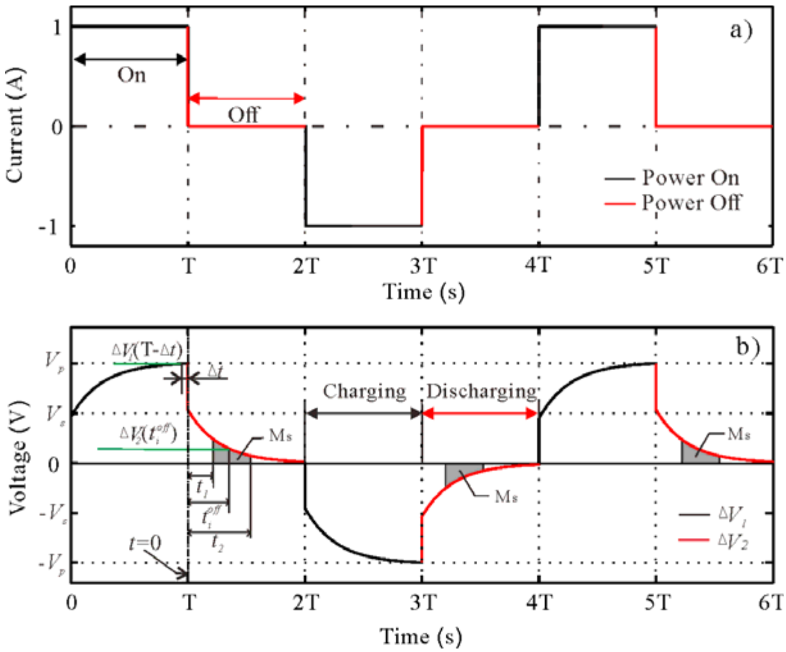


Fig. 1. The schematic of TDIP method: (a) excitation source for TDIP; (b) voltage curve for TDIP. The black and red line in panel (a), respectively, represents power-on and power-off period. The curves of the same colors in panel (b) are corresponding to primary field and secondary field. Some key time nodes are marked in (b).

The TDIP method measures the voltage decay induced by turning off the exciting current pulse and uses the characteristics of the decay to study the induced polarization, also known as time-domain chargeability. Seigel (1959) defined chargeability M_a as:

$$M_a = V_s / V_p \tag{1}$$

where V_s is the secondary field voltage immediately after the current is shut off and V_p is the effective primary voltage. In China, the apparent chargeability, η_s, is most commonly used:

$$\eta_s = \frac{\Delta V_2\left(t_i^{\mathrm{off}}\right)}{\Delta V_1(T - \Delta t)} \times 100\% \tag{2}$$

where $t = 0$ is the time node when the excitation current is turned off; t_i^{off} is a lapse time with respect to $t = 0$, typically close to 0.5 s; and $\Delta V_2(t_i^{\mathrm{off}})$ is the voltage of the secondary field at this time; $\Delta V_1(T - \Delta t)$ is the total field voltage, where T is the charging time, and Δt is a very short time before the excitation current is turned off so that the total field voltage is close to saturation voltage.

The precision of η_s is defined by $\Delta V_2(t_i^{\text{off}})$. However, the secondary voltage is difficult to measure accurately in the field, and as a result, modern IP instruments no longer measure the instantaneous value of $\Delta V_2(t_i^{\text{off}})$, but the mean value over the time window $[t_1\ t_2]$, $\frac{\int_{t_1}^{t_2}\Delta V_2(t)dt}{(t_2-t_1)}$. Therefore, another expression of apparent chargeability M_s can be expressed as:

$$M_s = \frac{\int_{t_1}^{t_2}\Delta V_2(t)dt}{\left(t_2-t_1\right)\cdot\Delta V_1(T-\Delta t)}\times 100\% \tag{3}$$

where, t_1 and t_2 are lapse times with respect to $t = 0$ (as shown in Fig. 1b). According to the Newmont standard, $t_1 = 0.45$ s and $t_2 = 1.1$ s (Swift Jr. 1973).

3. DATA-PROCESSING ALGORITHM

Pankratov and Geraskin (2010) put forward some reference processing algorithms for IP data processing including detrend and distortion suppression. However, their methods were designed for SIP method, and mainly focused on frequency-domain characteristics. To suppress the interference and obtain more complete IP information in time-domain, a data-processing algorithm is proposed in this paper (Fig. 2). Using a digital notch filter, segmentation, and overlapping averaging, a TDIP decay curve and apparent chargeability can be obtained. As part of this process, several key issues must be addressed: an efficient digital notch filter, exact secondary-field time-range determination, random noise suppression, SP removal technology, IP decay-curve extraction, and the apparent chargeability calculation.

3.1 Digital notch filter

According to the characteristic of the power line interference that its energy is concentrated mainly on the fundamental frequency and the odd harmonics, the zero location of the new notch filter should be as in Fig. 3, where $\omega_0 = 2\pi f_0/f_s$, and f_0 is the fundamental frequency and f_s is the sampling rate. For TDIP measurement, f_s was commonly 120 Hz. But with the increase of the system's throughput, f_s is raised to 2400 Hz.

The approach used here to design this notch filter is first to design a comb filter and then to implement a notch filter by taking an all-pass filter and subtracting the comb filter. The all-pass filter, which is intended for

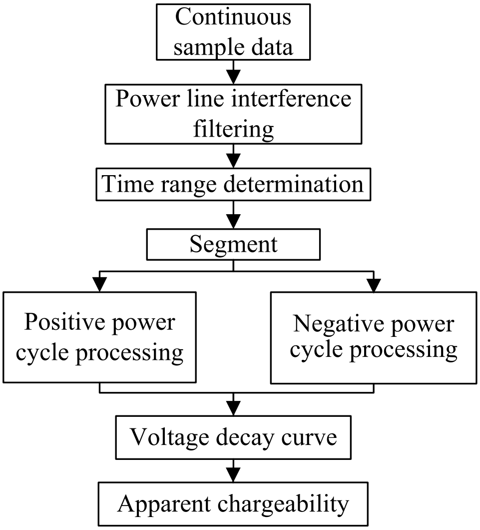


Fig. 2. TDIP measurement algorithm flowchart.

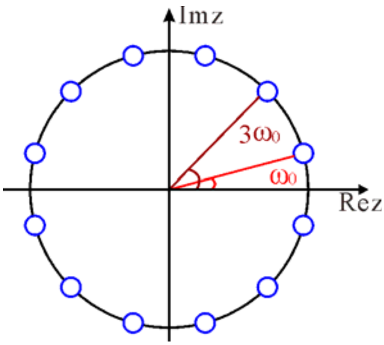


Fig. 3. Zero locations of the notch filter. The blue cycles represent the angular frequency of the odd harmonics. The red line represents the angular frequency ω_0 of the fundamental and the brown line represents $3\omega_0$ of the third harmonic, and so on.

signal delay purposes, is used to cancel the phase delay caused by the comb filter.

The comb filter is designed based on a narrow-band band-pass filter proposed by Hu (2003). The amplitude-frequency characteristic of this band-pass filter can be expressed as:

$$H_{bp}(z) = \frac{1+z^{-M}}{\left(1-z^{-1}e^{j\omega_0}\right)\left(1-z^{-1}e^{-j\omega_0}\right)} \tag{4}$$

where $\omega_0 = 2\pi f_0/f_s$, j is a square root of -1, and z is the delay unit which is equal to $e^{j\omega}$; and M is the order of the comb filter, in other words, the number of the comb filter poles.

From Eq. 4, the filter permits only frequencies close to the fundamental frequency to pass. Therefore, the comb filter is designed by setting poles at odd harmonic frequencies, $f_{\text{odd}} = (2i-1)\omega_0$, $i = 1, 2, \ldots, n/2$, and $n = f_s/f_0$) (Fig. 3). Hence, the comb filter amplitude-frequency characteristic is:

$$H_{cf}(z) = \frac{1+z^{-M}}{\prod_{i=1}^{\frac{n}{4}}\left[\left(1-z^{-1}e^{\mathrm{j}(2i-1)\omega_0}\right)\times\left(1-z^{-1}e^{(-1)\times\mathrm{j}(2i-1)\omega_0}\right)\right]} = \frac{1+z^{-M}}{1+z^{-n/2}} \tag{5}$$

Because only the odd harmonics are considered, the scope of i is from 1 to $n/4$. There is no odd harmonic interference except the fundamental frequency in the Nyquist band when $n < 4$. In that case, general band-stop filter could be used to deal with the interference. Winder (2002) described the design method of the band-stop filter in detail, so we will not cover it again.

The comb filter is designed based on pole zero cancellation method, so the fundamental pole ω_0 is also the zero point of the filter. Thus, we substitute $\omega_0 = 2\pi f_0/f_s$ into $z = e^{-jM\omega_0}$. And next substituting z into the numerator of Eq. 5 gives:

$$1+e^{-j2\pi M\frac{f_0}{f_s}} = 0 \Rightarrow \mathrm{e}^{-j2\pi M\frac{f_0}{f_s}} = -1 \Rightarrow j2\pi Mf_0/f_s = (2d+1)\pi \Rightarrow M = (2d+1)\times n/2, \tag{6}$$

where d is an integer. So, the frequency response of this comb filter is:

$$\begin{aligned} H_{cf}\left(e^{j\omega}\right) &= \frac{1+e^{-j\omega M}}{1+e^{-j\omega\frac{n}{2}}} = \frac{1+\cos(\omega M)+j\sin(\omega M)}{1+\cos\left(\frac{n\omega}{2}\right)+j\sin\left(\frac{n\omega}{2}\right)} \\ &= \frac{2\left[\cos\left(\frac{\omega M}{2}\right)\right]^2 + j2\cos\left(\frac{\omega M}{2}\right)\sin\left(\frac{\omega M}{2}\right)}{2\left[\cos\left(\frac{\omega n}{4}\right)\right]^2 + j2\cos\left(\frac{\omega n}{4}\right)\sin\left(\frac{\omega n}{4}\right)} \\ &= e^{-j\omega\left(\frac{M}{2}-\frac{n}{4}\right)}\frac{2\cos\left(\frac{\omega M}{2}\right)}{2\cos\left(\frac{\omega n}{4}\right)} = e^{-j\omega d\frac{f_s}{2f_0}}\frac{2\cos\left[(2d+1)\frac{\omega f_s}{4f_0}\right]}{2\cos\left(\frac{\omega f_s}{4f_0}\right)} \end{aligned} \tag{7}$$

According to the nature of the trigonometric function, the extreme value of Eq. 7 is $(2d+1)$. So Eq. 5 is normalized to:

$$H_{cf}\left(z\right)=\frac{1+z^{-(2d+1)\frac{f_s}{2f_0}}}{\left(2d+1\right)\times\left(1+z^{-\frac{f_s}{2f_0}}\right)} \tag{8}$$

The corresponding all-pass filter can be expressed as:

$$H_{ap}\left(z\right)=z^{-d\frac{f_s}{2f_0}}. \tag{9}$$

Finally, the required notch filter is calculated by subtracting the comb filter from the all-pass filter as follows:

$$H_{nf}\left(z\right)=z^{-d\frac{f_s}{2f_0}}-\frac{1+z^{-(2d+1)\frac{f_s}{2f_0}}}{\left(2d+1\right)\times\left(1+z^{-\frac{f_s}{2f_0}}\right)} \tag{10}$$

To achieve better amplitude-frequency characteristics and reduce ripple, a multi-level filter (k-order) is used, and therefore Eq. 10 becomes:

$$H_{nf}\left(z\right)=\left(z^{-d\frac{f_s}{2f_0}}\right)^k-\left[\frac{1+z^{-(2d+1)\frac{f_s}{2f_0}}}{\left(2d+1\right)\times\left(1+z^{-\frac{f_s}{2f_0}}\right)}\right]^k=\left(z^{-d\frac{f_s}{2f_0}}\right)^k-\left\{\frac{1-\left(-z^{-\frac{f_s}{2f_0}}\right)^{(2d+1)}}{\left(2d+1\right)\times\left[1-\left(-z^{-\frac{f_s}{2f_0}}\right)\right]}\right\}^k$$
$$=\left(z^{-\frac{f_s}{2f_0}}\right)^{dk}-\left[\frac{1}{\left(2d+1\right)}\sum_{i=0}^{2d}\left(-z^{-\frac{f_s}{2f_0}}\right)^i\right]^k \tag{11}$$

As a result, we developed a multi-frequency notch-filter with FIR structure which could suppress all of the odd harmonics and the fundamental frequency. An example is shown in Fig. 4, with $f_s = 1000$ Hz and $f_0 = 50$ Hz. This notch-filter's amplitude-frequency characteristic is determined by d and k (Fig. 5).

As shown in Figs. 4 and 5, d permits main control of the stop-band characteristics including the attenuation amplitude and the bandwidth, the bigger the d the smaller the attenuation and the narrower the stop-band width, and k mainly determines the ripple of the passband, the bigger the k the smaller the ripple. Specifically, the relation of the above-mentioned indexes with *k versus d* is shown in Fig. 6.

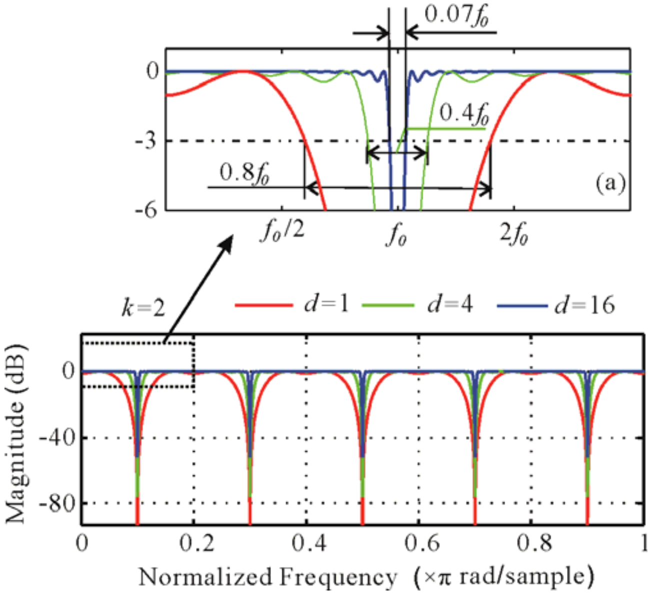


Fig. 4. Frequency response of the notch filter with fixed *k*. Panel (a) is zoomed image of the section within the dotted box and it shows the bandwidth of the stop-band at -3 dB.

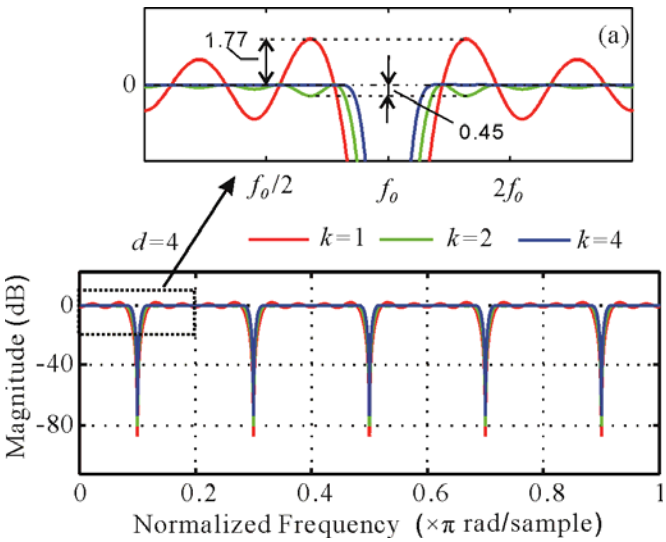


Fig. 5. Frequency response of the notch filter with fixed *d*. Panel (a) is zoomed image of the section within the dotted box and it shows the ripple strength of the passband.

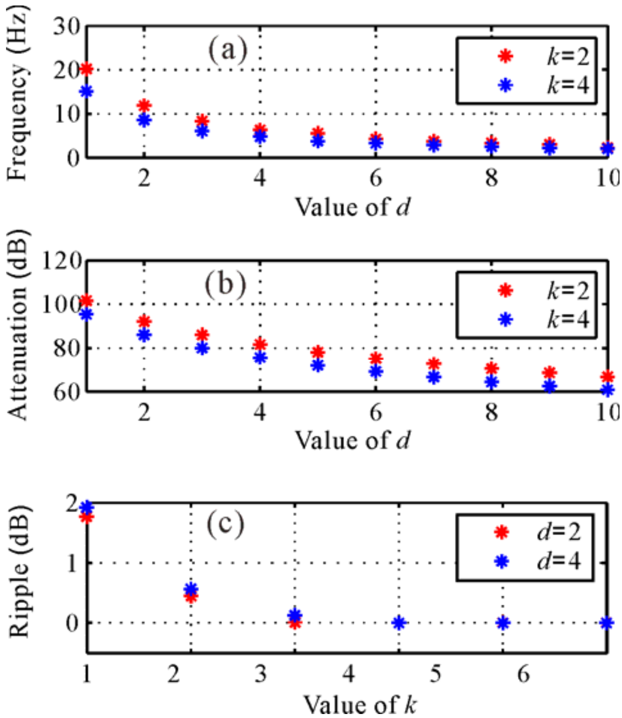


Fig. 6. Impact of the key parameters on the indexes of the notch filter. Panels (a) and (b): impact of *d* on the indexes of the notch filter with fixed *k*; (a) the bandwidth at -3 dB; (b) the attenuation amplitude at the fundamental frequency; (c) impact of *k* on the ripple of the notch filter with fixed *d*.

The choice of *d* should take into account of the interference suppression and undesirable effects of the notch filter. And *k* is not the bigger the better, because the length of the filter will increase exponentially with *k* and *d*, especially *k*, which leads to a rapid increase of hardware resource consumption. After all these comprehensive consideration, *d* and *k* were suggested as 4 and 2 for the notch filter used in TDIP. This notch filter has no low-frequency decay and phase delay, and all odd harmonics can be filtered with stop-band attenuation ≈ 80 dB. In addition, a certain stop-band width can deal with small fluctuations in the power line interference. This is a FIR filter with the advantages of stability and linear phase. The filter was implemented in both C++ code and MATLAB code so that it could be used in various situations. Furthermore, a hardware model of the filter was constructed in Digital Signal Processor (DSP) Builder (Fig. 7). DSP Builder is a product for performing Model-Based Design targeting Altera Field Programmable Gate Array (FPGA). By using Altera DSP Builder, algorithm

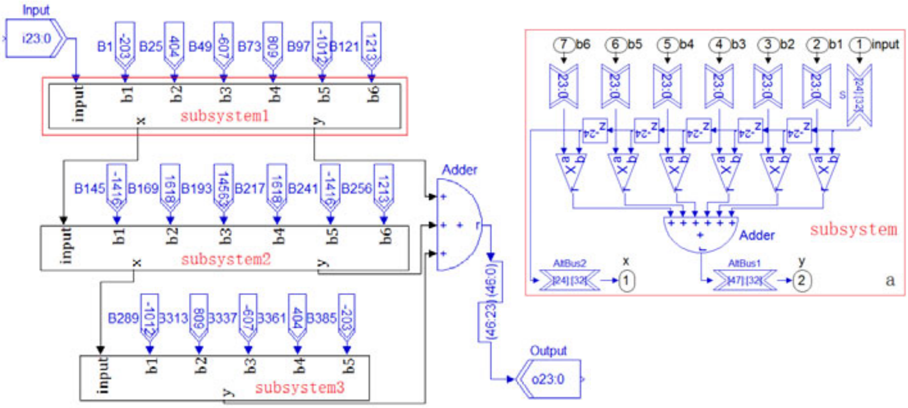


Fig. 7. Hardware model of the notch filter. f_s = 2400 Hz and f_0 = 50 Hz. The input data and output data are adjustable, and we set them to 24 bit here. Panel (a) is the internal structure of the subsystem.

and system designers can shorten DSP design cycles by creating the hardware representation of a DSP design in an algorithm-friendly development environment as well as share a common development platform. We utilized the DSP builder to implement a sample project for the filter and then exported it to the FPGA and compiled in Vhsic Hardware Description Language (VHDL) to obtain an intellectual property core (IP core) which could be integrated into the host. Beyond TDIP data processing, this notch filter can also be applied to other electromagnetic prospecting problems.

3.2 Noise reduction and SP removal

In addition to the power line interference, random noise and SP can also severely interfere with TDIP measurement. Because in this method TDIP data are collected for several successive cycles, it is possible to improve the signal to noise ratio (SNR) using segmentation and digital averaging.

Before segmentation, the time range of each power cycle should be determined. In most cases, synchronous technologies can provide an accurate temporal location (Deng *et al.* 2006), but they will not work properly under adverse circumstances. To overcome this limitation, an evaluation method for time location was proposed, called the "shifting average difference method (SAD)", as described below.

The first step is to determine the shift number N_{order}. This parameter is determined by the nature of the notch filter used in the process, and the filter program in Section 3.1 will automatically produce it. Hence, a n_{order}-point moving-average filter ($n_{\text{order}} = N_{\text{order}} - 1$) is first used to filter sequence x starting at the first data point to obtain x_1, where x is the previous segment (length is Tf_s, where T is the charging time) of the data processed by the notch filter. Second, the first-order difference of x_1, denoted by $\mathbf{dx}_1$, is calculated. The next step is to find the location of the maximum absolute value of $\mathbf{dx}_1$, denoted by t_1. Then one-step left shift is performed for x each time, and the above three steps are repeated. After norder iterations, the corresponding $t_2, t_3, \ldots, t_{N\text{order}}$ are calculated. And then the first-order difference of the sequence $\{t_1, t_2, t_3, \ldots t_{N\text{order}}\}$ is calculated, denoted by $\mathbf{d_t}$. The location where $\mathbf{d_t}$ begins to be convergent is denoted by c, and t_c is the jump edge. Based on comprehensive consideration of certain constraint conditions, including the jump-edge directionality of different power cycles and the periodicity of the secondary field, the initial time of the first secondary field, denoted by t_{sf1}, can be obtained. Finally, using t_{sf1}, f_s, and the period for each power cycle, the time information of the primary and secondary fields can be obtained. An algorithm flowchart of this method is shown in Fig. 8. To implement this procedure, it was coded in C++ and in MATLAB, respectively, generating two different but equivalent codes to suit different applications.

On this basis, the filtered data are segmented by positive or negative power cycle according to the time information calculated earlier, and then the positive segments are overlaid and averaged, as well as the negative segments. During this process, the useful signal is stacked in the same phase, but the interference components are stacked in different phases, further increasing SNR. This method yields two charge-discharge curves with opposite power direction. For SP, existing IP instruments usually combine manual

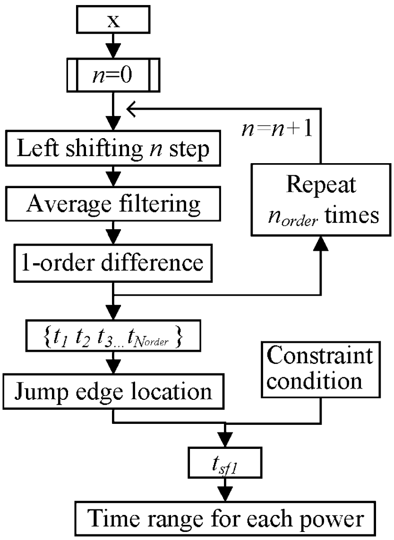


Fig. 8. Flowchart of time location. t_{sf1} is the initial time of the first secondary field.

and automatic compensation. This method will impact the amount of useful information in the presence of residual ΔV_2. Using the DC feature of SP, the charge-discharge curves of the positive and negative power cycles are averaged to remove SP from the secondary field. Finally, the secondary field decay curve can be obtained. In this way, random noise is decreased, and SP is also removed.

3.3 Decay curve and apparent chargeability

The approach described above yields a voltage decay curve for a complete power cycle. To improve the reliability of the results, the test values of repeated measurements can be averaged. Thus, a complete TDIP decay curve is obtained, composed of a complete primary field ΔV_1 curve and secondary field ΔV_2 curve.

In addition, the apparent chargeability η_s at different delay times can also be obtained. To do this, the average values of ΔV_2 in a time box (adjustable) can be calculated as a delay of the increase in the decay curve and then divided by the average value of the second half of ΔV_1 curve. In this way, the apparent chargeability at different points in time can be obtained.

4. RESULTS

4.1 Simulation test

The method described above was verified using the Cole-Cole model implemented in Simulation and Link (SIMULINK). SIMULINK, developed by MathWorks, is a graphical programming environment and widely used in automatic control and digital signal processing for multi-domain simulation and Model-Based Design. Models are represented graphically in SIMULINK as block diagrams. One of the primary advantages of employing SIMULINK for the analysis of dynamic systems is that it allows us to quickly analyze the response of complicated systems that may be prohibitively difficult to analyze analytically.

In general, the mathematical equations representing a given system that serve as the basis for a SIMULINK model can be derived from physical laws. In this section, we derived a mathematical model for TDIP measurement and then implemented that model in SIMULINK, as shown in Fig. 9. The supply current consisted of three parts: $I = I_0 + I_n + I_{sp}$, where I_0 was a forward and reverse pulse, 2 A with 1:1 duty cycle, with each power supply lasting 2 s; I_n was a fixed amount of random noise and power line interference with a fundamental frequency of 50 Hz, and I_{sp} was DC (equivalent of SP). The sampling rate was 2400 Hz, and the total sampling time was 48 s. The first step was to use I as an input stimulus current and the raw voltage data were acquired (shown in Fig. 10a). Obviously, the useful signal was

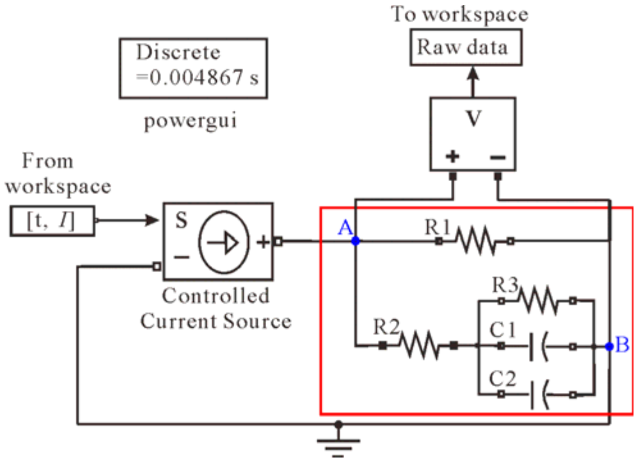


Fig. 9. SIMULINK model for TDIP measurement. The sample data was the potential difference between points A and B. The TDIP result was a reflection of the IP effect of the model in the red box, where R1 = 10 Ω, R2 = 40 Ω, R3 = 80 Ω, and C1 = C2 = 5E-3 F.

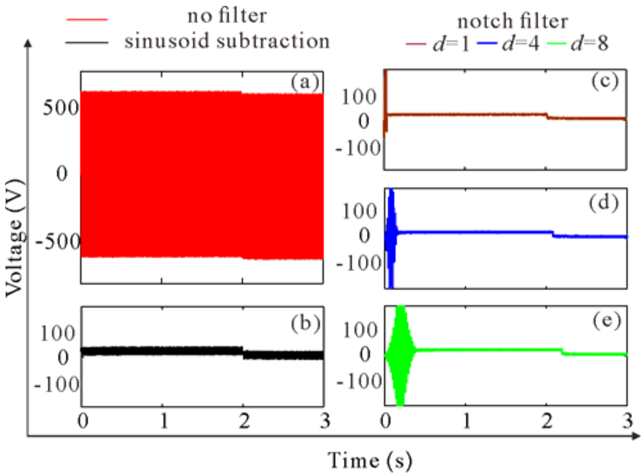


Fig. 10. Time-domain waves. Parameter k for the three notch filters was fixed as 2, and b was variable for filter performance analysis.

completely submerged by the powerful power line interference, and therefore the top priority was to remove this interference. The raw data were first filtered using two methods: (i) method 1: sinusoid subtraction (other filtering methods were not discussed due to their algorithmic complexity or precondition), as shown in Fig. 10b; and (ii) method 2: the notch filter described in Section 3.1 with $k = 2$ and $d = 1, 4, 8$, as shown in Fig. 10c-e.

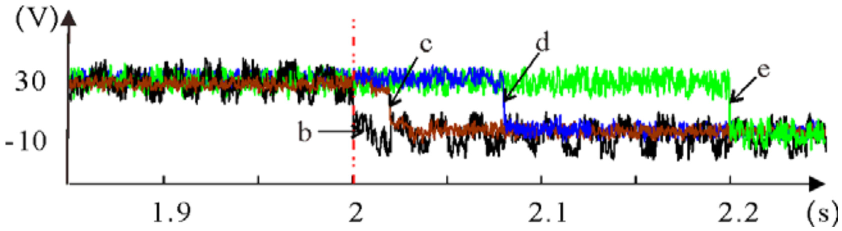


Fig. 11. Detail description of the time-domain data: panels (b) through (e) show the same curves as Fig. 10b-e.

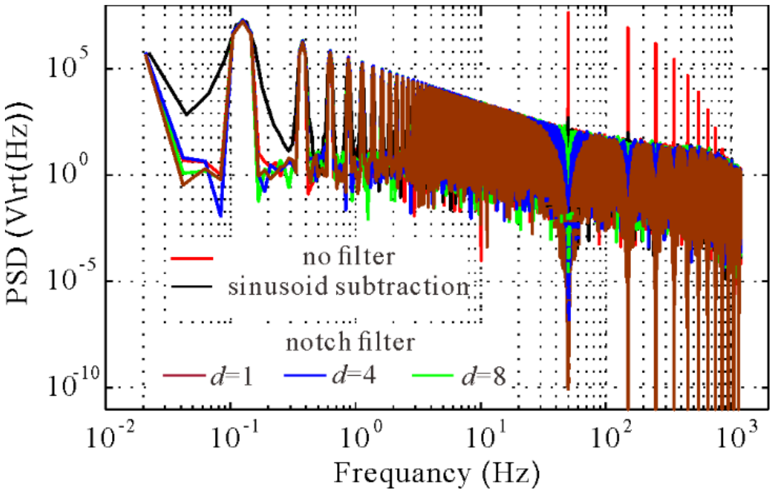


Fig. 12. Frequency-domain waves. The PSD were calculated by Fast Fourier Transform (FFT) algorithm, with length of 115 200 and hanning window.

As shown in Fig. 10, with the increase of *b*, the nonstationary process was extended and the filtering effect reduced. But too small d would lead to boundary confusion within the primary field and the secondary field, as shown in Fig. 11. So, $k = 2$ and $d = 4$ were selected as the optimal paremeters for the notch filter. As for sinusoid subtraction, this approach has the virtue of early stability and no phase delay. But its interference suppression ability appeared to be weaker than the notch filter and it could not cope with the frequency fluctuation of the interference, as shown by the Power Spectral Density (PSD) curves in Fig. 12.

By analyzing the frequency spectra of the time-domain data (Fig. 12), it becomes clear that the notch filter has successfully removed the power line interference while preserving the remaining spectral content of the excitation signal. The next step was to extract the time information of each power cycle using the SAD method. Next, the random noise and SP mixed into the fil-

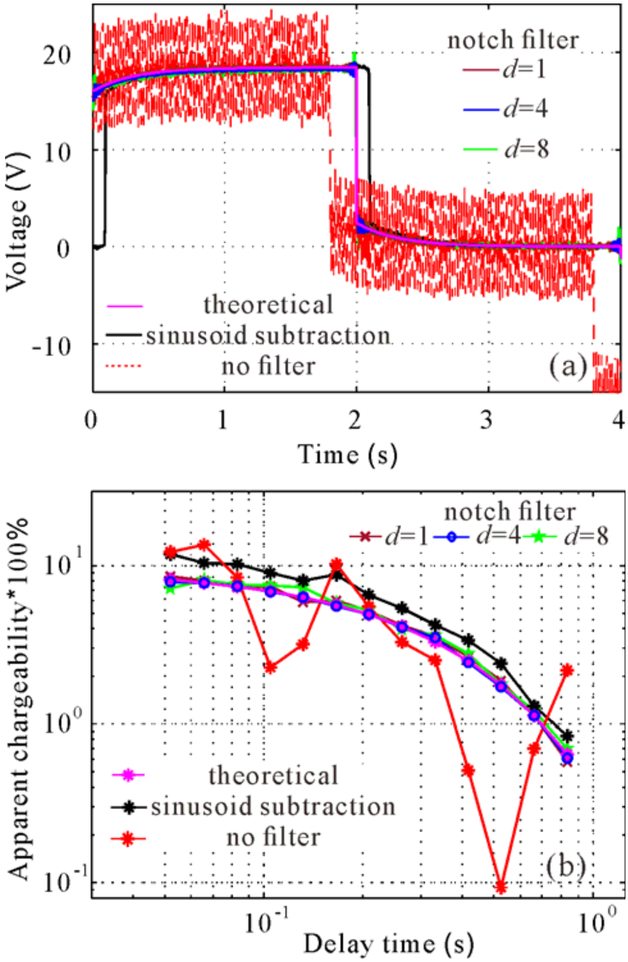


Fig. 13. Simulation results: (a): TDIP voltage decay curves; (b) comparison of the apparent chargeability gained by different methods. The pink line in panel (a) is the theoretical decay curve of this model which is collected by using I_0 as the excitation signal. The apparent chargeability curves in different colors in panel (b) are corresponding to the decay curves of the same color in (a).

tered data x were processed. Finally, after low-pass filtering and calculation, the model's charge-discharge curve was obtained (Fig. 13a) as well as the apparent chargeability η_s (Fig. 13b).

The simulation results show that the secondary field information calculated by using the selected notch filter is closer to the theoretical values, which proves the accuracy and effectiveness of the filter and the proposed method.

To simulate adverse circumstances under which the synchronous technologies cannot work well or data loss may occur (Deng *et al.* 2008), a random piece of data was removed from the actual data set at the beginning of the complete time series. In this case, the SAD method could determine the time range of the secondary field exactly. Therefore, the effective decay curve and the apparent chargeability could still be obtained.

4.2 Indoor test

Indoor test was carried out on the campus lawn at China University of Geoscience (Beijing). The charging and discharging time was both 2 s and power cycle was 8 s, with the emission current 5 A and f_s of 2400 Hz. The tests lasted about 30 min and three-channel data (with the electrodes parallel with each other) for 40 power cycles were collected. A segment of the raw data was shown in Fig. 14.

During this experiment, there was a lag time about 0.09 s between the receiver and the transmitter due to some hardware malfunction, as shown in Fig. 14 marked by terror. Traditional method and the new method were used in data processing, and the decay curve and chargeability were shown in Figs. 15 and 16.

The delay time brought a mismatch between the results calculated by two different methods. The starting time of the secondary field in the traditional method was earlier than the actual value (2 s). But the new algorithm got the correct time information for the decay curve.

The results of the three channels were about the same, so the curves with same symbols but different colors were overlapped on each other. But there were large differences between the results obtained by the two methods.

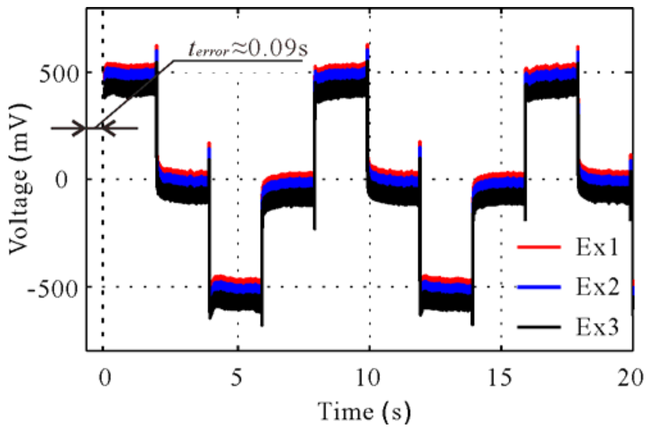


Fig. 14. Raw data of the lawn test. Colors: red, blue, and black represent the data collected by channel 1, channel 2, and channel 3 (Ex1, Ex2, and Ex3).

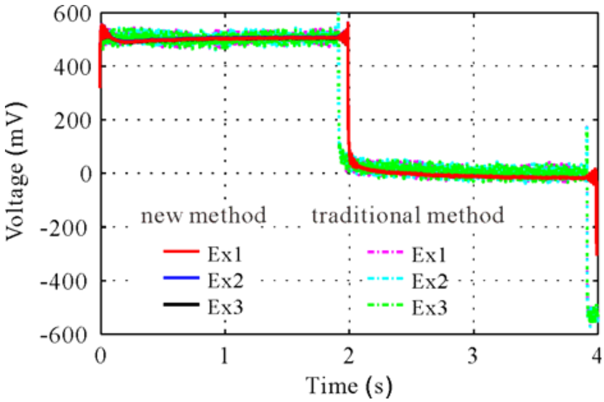


Fig. 15. Decay curve of the lawn test. The solid curves were calculated by the new method, and the dotted curves were analogue results of the traditional method for comparison. The traditional method could not output complete decay curve, because that it just collected partial data of the secondary field. Colors: red, blue, and black represent the results obtained by the data collected by channel 1, channel 2, and channel 3.

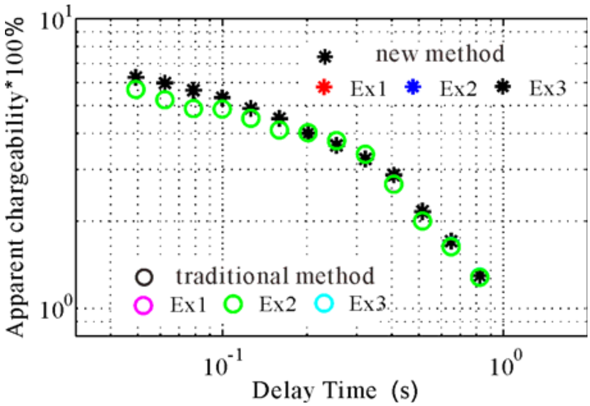


Fig. 16. Apparent chargeability of the indoor test. The curves of red, blue, and black represent the results obtained by the data collected by channel 1, channel 2, and channel 3. Symbols: results obtained by using traditional method (o) and the data-processing algorithm proposed in this paper (*), respectively.

4.3 Field test

Field tests were conducted in Xilin Gol Grassland, Inner Mongolia, which is a sparsely populated region. Results were shown in Fig. 17.

The spikes in Fig. 17 were caused by over-charging. They have seriously affected the observations, so we removed these abnormal data by using curve fitting, similar to the approach taken in Pankratov and Geraskin

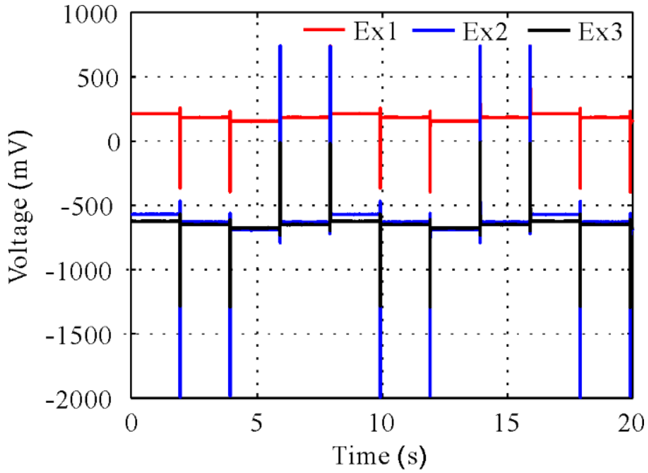


Fig. 17. Raw data of the field test. The curves of red, blue, and black represent the results obtained by the data collected by channel 1, channel 2, and channel 3 (Ex1, Ex2, and Ex3, with the electrodes parallel with each other).

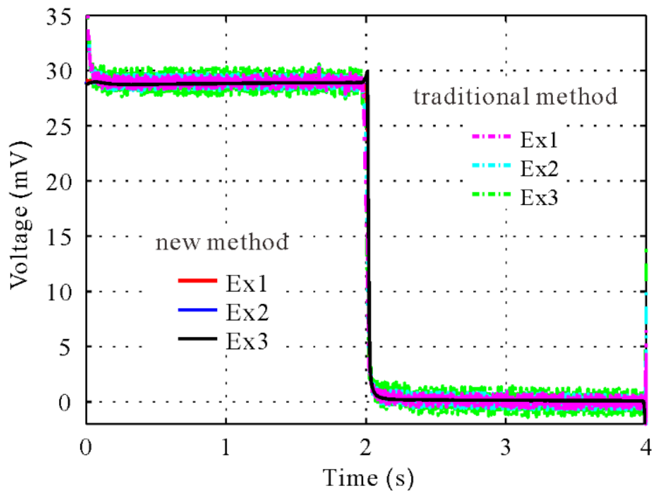


Fig. 18. Decay curve of the field test. The solid curves were calculated by the new method and the dotted curves were analogue results of the traditional method respectively. Ex1, Ex2, and Ex3 represented the results obtained by the data collected by channel 1, channel 2, and channel 3.

(2010). But fitting criterions were different. In our algorithm, the primary field and the secondary field were processed by cubic spline interpolation and tendency fitting separately, and fitting lengths were twice of the overcharging. At the same time, the self-potential was rather strong, but the power

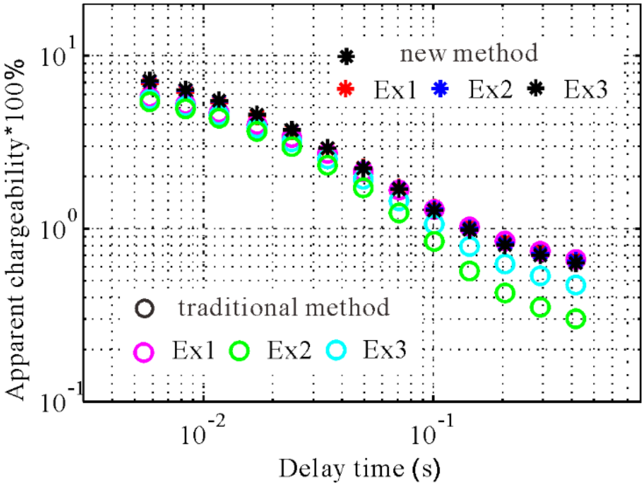


Fig. 19. Apparent chargeability of the field test. Symbols: results obtained by using traditional method (o) and the data-processing algorithm proposed in this paper (*), respectively.

line interference was negligible. So the filtering step was overleaped in the following data processing. Traditional method and the new method without notch filter were used. The curve fitting was also used for both of them to improve the readability of the results. The decay curve and apparent chargeability were shown in Figs. 18 and 19.

5. DISCUSSION

For the simulation test, by analyzing the time domain and frequency domain wave of the raw data and the filtered data, both of the two methods have successfully removed the power line interference. By comparison, sinusoid subtraction has preserved the remaining spectral content of the signal, including the frequencies close to the power line interference; moreover, it has the advantages of early stability and no phase delay. But the interference attenuation is obviously lower than that brought by the notch filter. Indeed, the attenuation distortion at frequencies in the vicinity of the notch did not affect the TDIP result directly, but insufficient interference suppression did. As a result, both the decay curve and the apparent chargeability obtained by using the digital notch filter were closer to the theoretical values of the model than by using sinusoid subtraction. Hence, the results obtained by the non-filtered data were seriously distorted.

In the lawn experiment, conventional method based on time control was fail to obtain good results in the case of a synchronization failure and the environment interference, but the new algorithm described in Section 3 worked well and valid IP parameters were obtained.

As for the field test, the difference of the raw data collected by the three channel were disturbed by the SP and the inconsistency of channels. In addition, the time-domain data were filled with spikes induced by over-charging. Nevertheless, effective decay curves and apparent chargeability were obtained by the new algorithm, with the results of different channels almost consistent with each other. In contrast, there were obvious differences among the IP results of different channels calculated by the traditional method, and the result of channel 1 which had the minimum self-noise was closest to the new algorithm's results.

6. CONCLUSIONS

To address various problems existing in TDIP measurement, including power line interference, incomplete decay curves, and synchronous technology limitations, an improved algorithm has been proposed here to increase the interference rejection capability and obtain more complete information on the secondary field. The first step was to design an effective digital notch filter for power line interference which is available for multiple data-processing fields and to code it in C++ and MATLAB separately. In addition, a hardware model of this filter was constructed with $f_s = 2400$ Hz and $f_0 = 50$ Hz in DSP Builder, and an IP core was obtained. Next, segmentation, overlapping, and averaging methods were used to suppress the random noise and remove SP. As part of this process, the "shifting average difference method" was proposed to determine the time range of the secondary field. By this method, it was possible to determine exactly the jump edge of the secondary field in the case of incomplete synchronization of the receiver and the transmitter. Finally, a series of tests verified that this method can obtain approximately complete secondary field information and apparent chargeability at different decay times in a powerful noise environment. The notch filter can greatly improve the SNR and is suitable for TDIP measurement. The agreement of the calculated and theoretical results proved the effectiveness and accuracy of the notch filter and the algorithm. Indoor and field tests indicated that the new algorithm described in Section 3 could improve the anti-interference ability of TDIP method significantly, and it was also help to overcome the restriction of the adverse field conditions and the effect of the instruments' inherent differences. By using this algorithm, we can get a complete decay curve as well as general TDIP parameters from the raw data which was seriously polluted by various interferences.

Acknowledgments. This research is sponsored by 863 Program (2012AA09A201), National Natural Science Foundation of China

(61531001), and the open fund (No. GDL1411) of the Key Laboratory of Geo-detection (China University of Geosciences, Beijing), Ministry of Education. We thank MathWorks for the MATLAB code used for filter design and simulation. We are truly grateful to the editors' and reviewers' advices and suggestions on this paper.

References

Butler, K.E., and R.D. Russell (1993), Subtraction of powerline harmonics from geophysical records, *Geophysics* **58**, 6, 898-903, DOI: 10.1190/1.1443474.

Butler, K.E., and R.D. Russell (2003), Cancellation of multiple harmonic noise series in geophysical records, *Geophysics* **68**, 3, 1083-1090, DOI: 10.1190/1.1581080.

Dahlin, T., V. Leroux, and J. Nissen (2002), Measuring techniques in induced polarisation imaging, *J. Appl. Geophys*. **50**, 3, 279-298, DOI: 10.1016/S0926-9851(02)00148-9.

Deng, M., and Z.X. He (1998), Data Acquisition characteristics and data acquisition system of differentially-normalized method, *Geoscience – J. Graduate School Chin. Univ. Geosc*. **12**, 3, 442-446 (in Chinese).

Deng, M., F. Liu, Q. Zhang, and K. Chen (2006), Long-span and multi-point synchronizing data acquisition for seafloor magnetotelluric based on union of marine and land, *Sci. Technol. Rev*. **24**, 10, 28-32 (in Chinese).

Deng, M., W. Wei, Q. Zhang, and M. Wang (2008), Real-time backup technique for reducing data loss in seafloor magnetotelluric data acquisition*, J. Cent. South Univ. (Science and Technology)* **39**, 2, 362-367 (in Chinese).

Dhillon, S.S., and S. Chakrabarti (2001), Power line interference removal from electrocardiogram using a simplified lattice based adaptive IIR notch filter. **In:** *Proc. 23rd Ann. Int. Conf. IEEE*, DOI: 10.1109/IEMBS.2001.1019561.

Fan, S.J., and X.G. Yang (1979), SDJ-1 electric logging device, *J. Shandong Eng. Col*. **4**, 94-104 (in Chinese).

Feng, Y.J., and Z.H. Fu (1994), DJD6-1 multichannel IP instrument, *Geosci. Instr.* **3**, 33-36 (in Chinese).

Fiandaca, G., E. Auken, A.V. Christiansen, and A. Gazoty (2012), Time-domain-induced polarization: Full-decay forward modeling and 1D laterally constrained inversion of Cole-Cole parameters, *Geophysics* **77**, 3, E213-E225, DOI: 10.1190/geo2011-0217.1.

Galiana-Merino, J.J., D. Ruiz-Fernandez, and J.J. Martinez-Espla (2013), Power line interference filtering on surface electromyography based on the stationary wavelet packet transform, *Comput. Meth. Prog. Biomed*. **111**, 2, 338-346, DOI: 10.1016/ j.cmpb.2013.04.022.

Gazoty, A., G. Fiandaca, J. Pedersen, E. Auken, A. Christiansen, and J. Pedersen (2012), Application of time domain induced polarization to the mapping of lithotypes in a landfill site, *Hydrol. Earth Syst. Sci.* **16**, 6, 1793-1804, DOI: 10.5194/hess-16-1793-201.

Hönig, M., and B. Tezkan (2007), 1D and 2D Cole-Cole-inversion of time-domain induced-polarization data, *Geophys. Prospect*. **55**, 1, 117-133, DOI: 10.1111/j.1365-2478.2006.00570.x.

Hördt, A., R. Blaschek, A. Kemna, and N. Zisser (2007), Hydraulic conductivity estimation from induced polarisation data at the field scale – the Krauthausen case history, *J. Appl. Geophys*. **62**, 1, 33-46, DOI: 10.1016/j.jappgeo.2006.08.001.

Hu, G.S. (2003), *Digital Signal Processing: Theory, Algorithms and Implementation*, 2nd ed., Tsinghua University Press, Beijing, 648 pp. (in Chinese).

Kohno, R., H. Imai, M. Hatori, and S. Pasupathy (1990), An adaptive canceller of cochannel interference for spread-spectrum multiple-access communication networks in a power line. Selected areas in communications, *IEEE J. Sel. Area* **8**, 4, 691-699, DOI: 10.1109/49.54465.

Li, J.H., P.R. Lin, B.L. Xu, Q.K. Meng, and D. Li (2013), Study on multifunctional electromagnetic prospecting technology in mineral exploration, *Adv. Mat. Res.* **734**, 178-182, DOI: 10.4028/www.scientific.net/AMR.734-737.178.

Marshall, D.J., and T.R. Madden (1959), Induced polarization, a study of its causes, *Geophysics* **24**, 4, 790-816, DOI: 10.1190/1.1438659.

Martinho, E., F. Almeida, and M. Senos Matias (2006), An experimental study of organic pollutant effects on time domain induced polarization measurements, *J. Appl. Geophys*. **60**, 1, 27-40, DOI: 10.1016/j.jappgeo.2005.11.003.

Nimunkar, A.J., and W.J. Tompkins (2007), EMD-based 60-Hz noise filtering of the ECG. **In:** *Proc. 29th Ann. Int. Conf. IEEE EMBS* 2007, DOI: 10.1109/IEMBS.2007.4352688.

Pankratov, O.V., and A.I. Geraskin (2010), On processing of Controlled Source Electromagnetic (CSEM) data, *Geol. Acta Int. Earth Sci. J.* **8**, 1, 31-49, DOI: 10.1344/105. 000001514.

Sahambi, J., S. Tandon, and R. Bhatt (1997), Quantitative analysis of errors due to power-line interference and base-line drift in detection of onsets and offsets in ECG using wavelets, *Med. Biol. Eng. Comput*. **35**, 6, 747-751, DOI: 10.1007/BF02510988.

Seigel, H.O. (1959), Mathematical formulation and type curves for induced polarization, *Geophysics* **24**, 3, 547-565, DOI: 10.1190/1.1438625.

Su, M.-L., and K.-S. Chuang (2013), An ECG signal enhancement based on improved EMD. **In:** *PIERS Proceedings, Taipei, Taiwan*.

Swift Jr., C. (1973), The L/M parameter of time-domain IP measurements – A computational analysis, *Geophysics* **38**, 1, 61-67, DOI: 10.1190/1.1440334.

Tong, M., L. Li, W. Wang, and Y. Jiang (2006), A time-domain induced-polarization method for estimating permeability in a shaly sand reservoir, *Geophys. Prospect*. **54**, 5, 623-631, DOI: 10.1111/j.1365-2478.2006.00568.x.

Wan, H., R. Fu, and L. She (2006), The elimination of 50 Hz power line interference from ECG using a variable step size LMS adaptive filtering algorithm, *Life Sci. J.* **3**, 4, 90-93.

Warden, S., S. Garambois, P. Sailhac, L. Jouniaux, and M. Bano (2012), Curvelet-based seismoelectric data processing, *Geophys. J. Int.* **190**, 3, 1533-1550, DOI: 10.1111/j.1365-246X.2012.05587.x.

Winder, S. (2002), *Analog and Digital Filter Design*, 2nd ed., Newnes, Amsterdam.

Wu, J.G., and H. P. Dou (1996), DCX-1 active electric field differential instrument, *Petrol. Inst*. **10**, 1, 21-24 (in Chinese).

Xia, J., and D. Miler (2000), Design of a hum filter for suppressing power-noise in seismic data, *J. Environ. Eng. Geophys*. **5**, 2, 31-38, DOI: 10.4133/JEEG5.2.31.

Zhang, D.J. (2009), Circuit design of late-model IP detection system, M.Sc. Thesis, China University of Geosciences, Beijing, China (in Chinese).

Received 15 June 2015
Received in revised form 2 March 2016
Accepted 21 March 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2289-2304
DOI: 10.1515/acgeo-2016-0113

Interpretation of Gravity and Gamma-Ray Spectrometry Data in Low Temperature Hydrothermal Systems, Southeastern Part of Fukuoka, Japan

Jun NISHIJIMA and Yasuhiro FUJIMITSU

Department of Earth Resources Engineering, Faculty of Engineering,
Kyushu University, Fukuoka, Japan;
e-mail: nishijima@mine.kyushu-u.ac.jp

Abstract

The Hakata hot springs area is located in Fukuoka City, which is in the southwestern part of Japan. Gamma-ray and gravity surveys were conducted to understand the relationship between the low-temperature hydrothermal systems and geophysical data of the area. The depth of the reservoir basement, which was derived from gravity data, gradually deepens toward the east; it includes some steep depth gradients in the Hakata hot springs area. High intensities of gamma-rays were detected around these gradients. In addition, higher hot spring temperatures and flow rates can be observed in this area. These results indicate that some part of the level of the basement where the hot springs are concentrated is a part of the Kego Fault and is similar to the fracture zone created by past activities of the fault. Moreover, these steep depth gradients act as a path for hot spring water from the deeper side of the granitic body to the surface.

Key words: gravity, gamma-ray, Hakata hot springs, Kego Fault, hydrothermal system.

Ownership: Institute of Geophysics, Polish Academy of Sciences;

1. INTRODUCTION

Fukuoka City, which has the largest population on Kyushu Island, is situated in the southwestern part of Japan (Fig. 1). The Hakata hot springs area is one of the low- to moderate-temperature hot spring areas on the southeastern end of Fukuoka and is associated with an active fault (Kego Fault). Some 11 wells were drilled in the late 1960s to access the hot springs (Table 1). The wells reach depths of 150 m and produce hot spring temperatures of 27°C to 50°C (Matsushita *et al.* 1971). The hot spring temperatures are relatively high compared to the other hot springs in the city, but no deeper geothermal resource is expected to be discovered (Karakida *et al.* 1994).

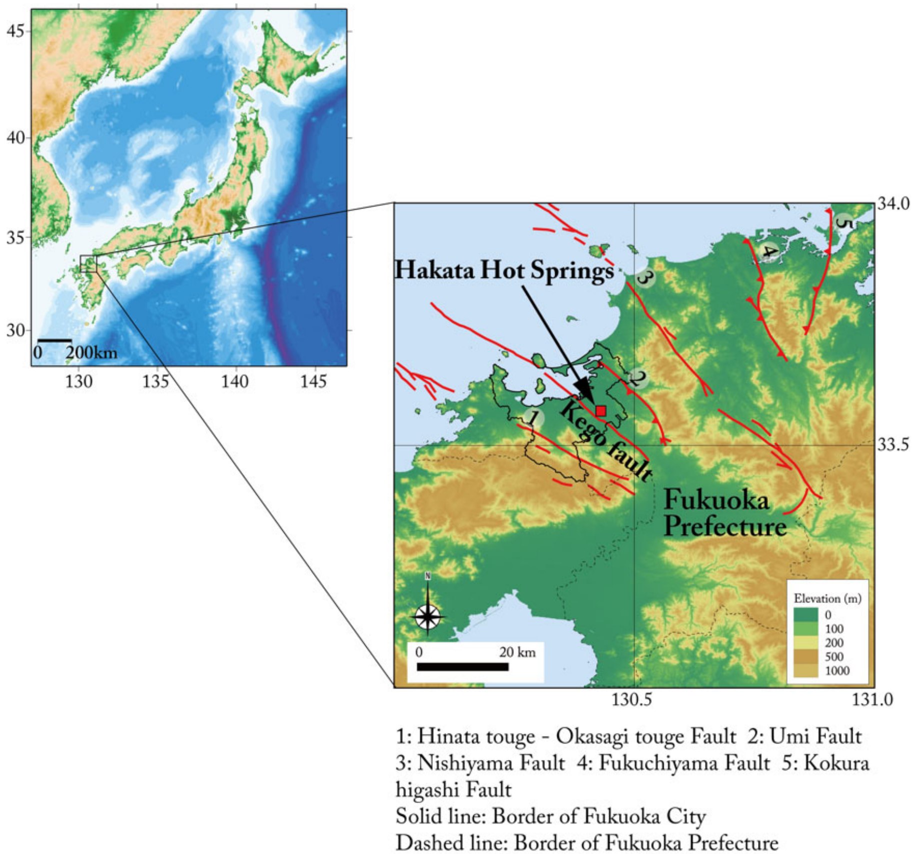


Fig. 1. Location of the Hakata hot springs area, southwestern part of Japan; modified from HERP (2013). NW-SE trending normal and left-lateral strike-slip faults are developed throughout the area.

Table 1

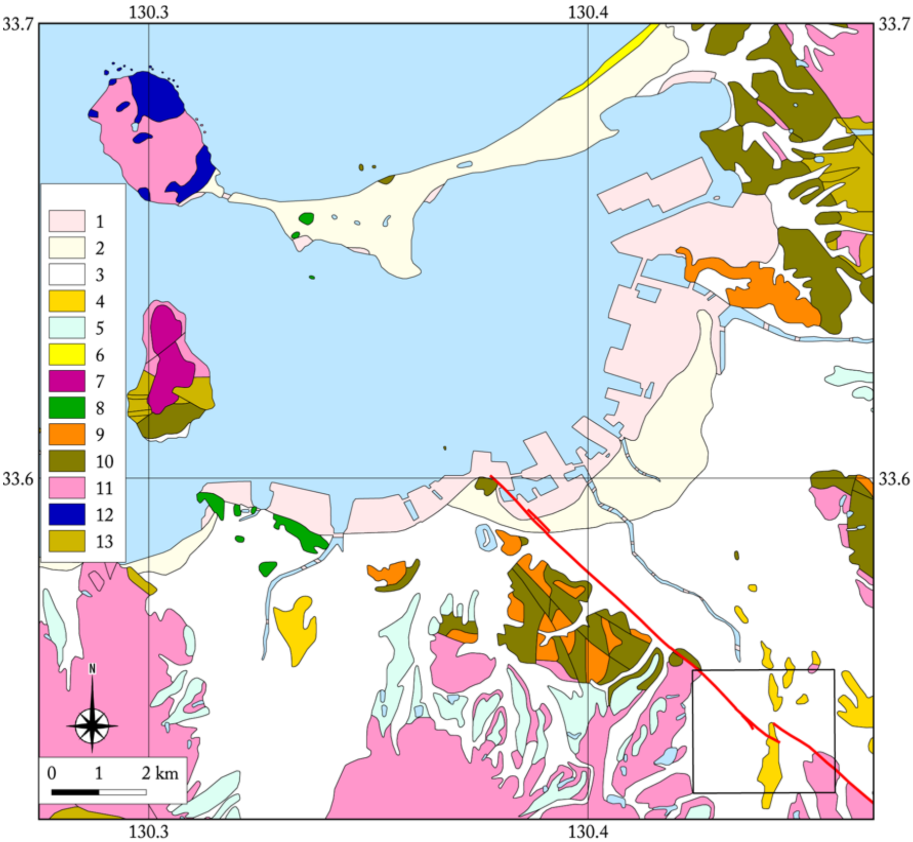
Characteristics of hot spring wells in the Hakata hot spring area (Matsushita *et al.* 1971)

Well ID	AB	S1	S2	SE	MI	KK	A	TK-A	T-F	SI	FU
Depth [m]	72	70	100	78	100	150	100	150	150	130	100
Temperature [°C]	49.0	43.0	37.5	46.0	31.0	42.0	27.0	46.0	34.3	33.1	43.0
Flow rate [kg/s]	1.7	0.9	2.0	0.8	0.4	1.8	0.3	2.0	0.7	–	2.0
Borehole geology	○					○					

A strong earthquake (M_{JMA} 7.0) occurred offshore to the west of Fukuoka Prefecture on 20 March 2005. The distribution of hypocenters and seismogenic stress as defined by aftershocks suggested that the 2005 West Offshore Fukuoka Prefecture earthquake occurred on the fault that runs in the NW extension of the Kego Fault (Shimizu *et al.* 2006). The results of gravity gradient interpretation techniques (Saibi *et al.* 2008) showed that the Hakata hot springs area was strongly connected with the NW-SE active fault system that is named the Kego Fault. MEXT (2013) reported the precise location of the Kego Fault based on the comprehensive survey of the active fault in Fukuoka city, passing through this study area (Fig. 2).

Most active urban faults are covered by unconsolidated Quaternary sediments. Dense buildings and other facilities constructed in the study area prevented the use of electromagnetic and seismic methods. There are some geophysical approaches, however, that can detect the fractures and faults in urban areas (Sultan *et al.* 2012, Xu *et al.* 2015). A gravity survey is one of the effective methods for the detection of a fault structure related to a path of hot water (Abdelzaher *et al.* 2011, Represas *et al.* 2013). Additionally, a gamma-ray survey is an efficient method for fracture identification. Subsurface fractures previously have been associated with elevated uranium concentrations due to the movement of uranium in subsurface hot water circulation (McCay *et al.* 2014). The results of a gamma-ray survey in geothermal and hot spring areas (Mogi and Okada 1990a, b) suggested that the high gamma-ray intensities of ^{214}Bi and ^{208}Tl are strongly related to the presence of an upflow of hot water.

The over pumping of hot spring water is causing a shortage of hot spring resources. To avoid such an exhaustion of a hot spring, it is necessary to understand the underground structure and hydrothermal system. Several surveys have been conducted in the Hakata hot spring area. Yamashita *et al.* (1965) used an electrical resistivity method for the groundwater survey in this area. The low resistivity zone was detected at a depth of more than 30 m.



1: Reclaimed land 2: Qurtanary sand dune deposit 3: Quartanary alluvium 4: Aso-4 pyroclastic flow deposit 5: Quartanary gravel and sand 6: Qurtanary Paleo-sand dune deposit 7: Neogene Basalt 8: Paleogene sandstone and mudstone 9: Paleogene sandstone and conglomerate 10: Paleogene sandstone, mudstone, conglomerate and coal 11: Cretaceous granite 12: Cretaceous plutonic rock 13: Paleozonic metamorphic rock

—— Kego Fault (MEXT, 2013) ☐ Study area

Fig. 2. Geological map of the study area and its surroundings. Modified from GSJ AIST (2014). The red line and the black rectangle show the location of the Kego Fault (MEXT 2013) and study area, respectively.

Matsushita *et al.* (1971) interpreted this low resistivity zone as a hot spring reservoir based on the observation of a stratigraphic column. The flow path of the hot spring water and the reservoir structure, however have been little investigated. In addition, few geophysical studies have attempted to understand low-temperature hydrothermal systems such as this hot spring area. This study was carried out to understand the relationship between the hot spring water flow in the low-temperature hydrothermal system and the geophysical survey results (gravity and gamma-ray surveys).

2. GEOLOGICAL SETTINGS

Northern Fukuoka is located in the W-E compression field caused by the subduction of the Philippine Sea Plate under the Eurasian Plate (Shimizu *et al.* 2006). There are therefore many NW-SE trending active faults in the northern part of Fukuoka Prefecture (Fig. 1). The Headquarters for Earthquake Research and Promotion (HERP 2007) reported that the Kego Fault is a NW-SE-trending left-lateral strike-slip fault with an overall length of 27 km. According to Kimura *et al.* (2013), the strike-slip basin structure is characterized by west-to-south westward tilting and bounded by the Kego Fault on its southwest side. This basin of the basement is covered by Middle Pleistocene to Holocene deposits.

Figure 2 shows the geological map of Fukuoka city (GSJ AIST 2014). The geological setting in Fukuoka is comprised of Sangun metamorphic rocks (Paleozoic), Late Mesozoic granitic rocks, Paleogene sandstone, Neogene basaltic rocks, and Quaternary sediments (Karakida *et al.* 1994). In this map, the gravitational basement, which is high in density, is considered Mesozoic granitic rocks.

3. GAMMA-RAY SPECTROMETRY

The gamma-ray survey was carried out using the portable multi-channel spectrum analyzer E-560A and a NaI (Tl) scintillation detector (made by NAIG Co. Ltd., Japan) at 214 points. The diameter and height of detector is 4 inches. The detector unit was set on the ground; the analyzer measures 1024 channels between 0.03 and 3.00 MeV. The spacing between stations varies between 50 and 100 m, depending on the accessibility. Three radioactive elements, ^{40}K, ^{214}Bi (1.76 MeV), and ^{208}Tl (2.61 MeV), were detected as an indicator of upflows of hot water. A data processing method based on Mogi and Jinguuji (1993) was used; it included the error estimation of observation. The intensity of each element was defined as the net count in 300 s.

Aswathanarayana (1985) pointed out that the net count is affected by surface geology and meteorological conditions. The two net count ratios, ^{214}Bi/^{40}K and ^{208}Tl/^{40}K, were therefore adopted for the index of gamma-ray intensity. These indexes are shown in Figs. 3 and 4, classified into the following four categories using an average (m) and a standard deviation (SD).

$$\text{Cat. 1 (high)} > m + SD > \text{Cat. 2} > m > \text{Cat. 3} > m - SD > \text{Cat. 4 (low)}$$

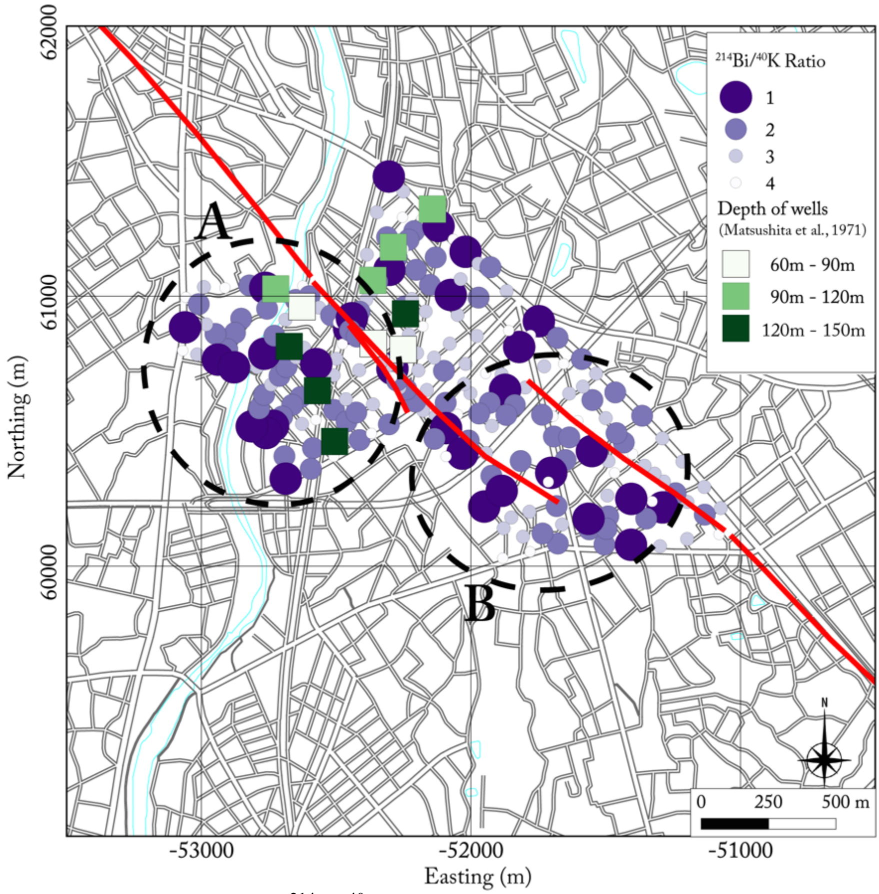


Fig. 3. Distribution of the ^{214}Bi/^{40}K ratio. The coordinates in Figs. 3-6 are Japan Plane Rectangular Coordinate System II. Two high ^{214}Bi/^{40}K intensity areas (A and B) were detected along the Kego Fault. Mean: 0.16, Standard deviation: 0.04. Cat. 1 > 0.2 > Cat. 2 > 0.16 > Cat. 3 > 0.12 > Cat. 4.

Two high ^{214}Bi/^{40}K and ^{208}Tl/^{40}K intensity areas, which correspond to the Kego Fault, were identified. The northwestern high intensity area was located in and around the hot spring wells. These results indicated that ^{222}Rn and ^{220}Rn, the parent elements of ^{214}Bi and ^{208}Tl, respectively, come to the surface together with hot water. The difference between these two elements is their half-life, as the half-life of ^{208}Tl (3 min) is shorter than that of ^{214}Bi (19.9 min). ^{208}Tl is accumulated rapidly at the location where hot water nears the surface.

Although there is no hot spring well in the southeastern area where the high intensity ^{214}Bi and ^{208}Tl are detected, it is possible that hot water is moving up towards the surface in that area.

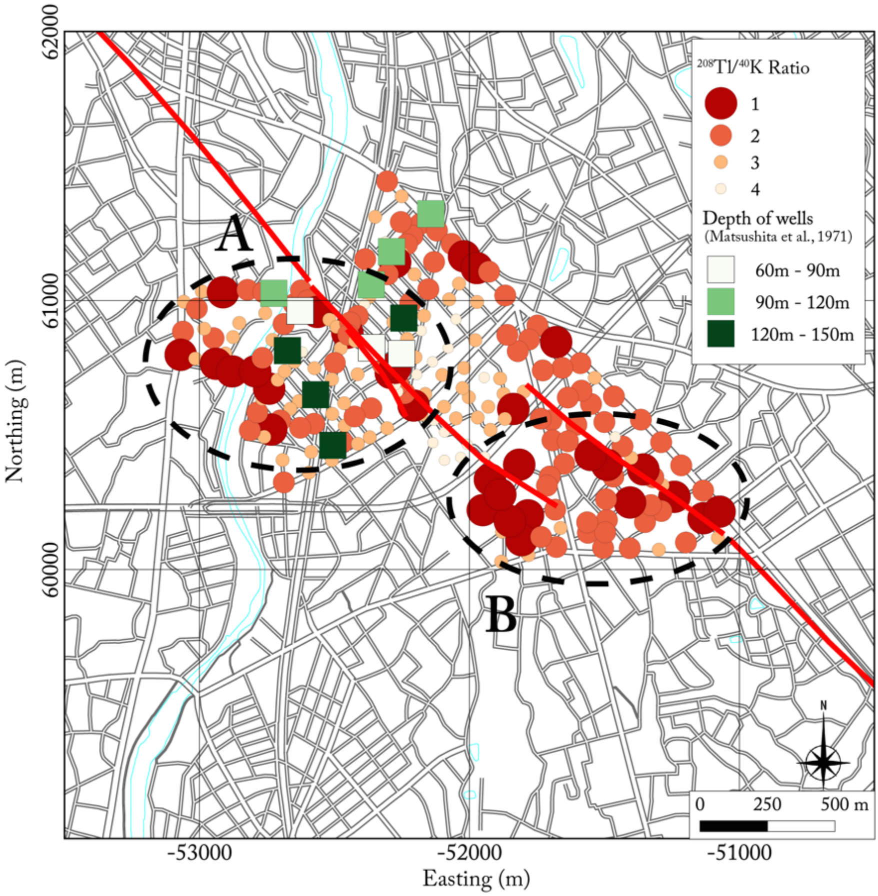


Fig. 4. Distribution of the ^{208}Tl/^{40}K ratio. Two high ^{208}Tl/^{40}K intensity areas (A and B) were detected along the Kego Fault. Mean: 0.08, Standard Deviation: 0.02. Cat. 1 > 0.1 > Cat. 2 > 0.08 > Cat. 3 > 0.06 > Cat. 4.

4. GRAVITY DATA

The gravity survey was carried out using Scintrex CG-3, CG-3M, and CG-5 gravimeters. The intervals of the survey points were set as several tens to two hundred meters. The measured gravity data was reduced to complete the Bouguer anomaly, applying the necessary corrections such as height, tidal, free-air, Bouguer, and terrain conditions. The assumed density of 2470 kg/m^3 was determined by an objective Bayesian approach (Murata 1993) for the Bouguer and terrain corrections.

The complete Bouguer anomaly map is depicted in Fig. 5. The Bouguer anomalies are all positive, ranging from 14.1 to 17.4 mgal, decreasing in the southeastern regions of the map area (Fig. 5). The regional anomaly, caused

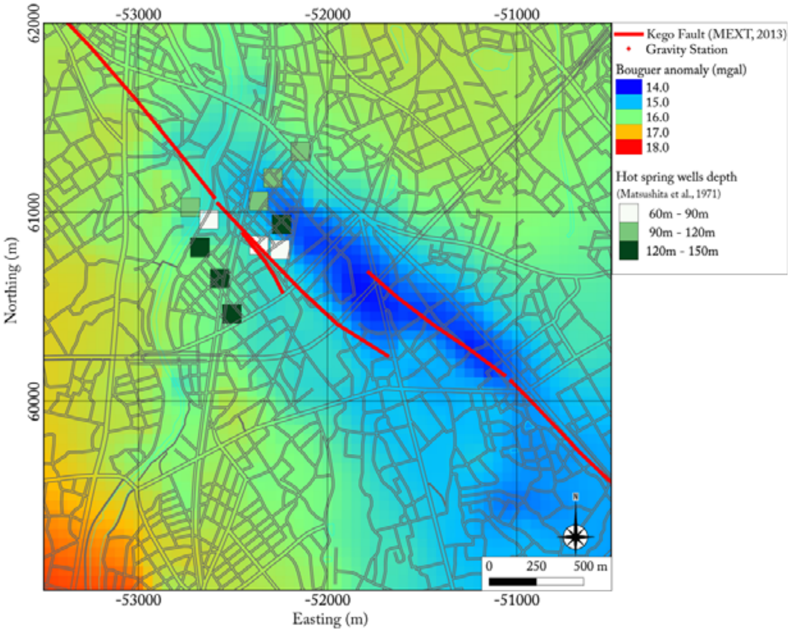


Fig. 5. Bouguer anomaly with an assumed density of 2470 kg/m^3. Red circles show the location of the gravity station.

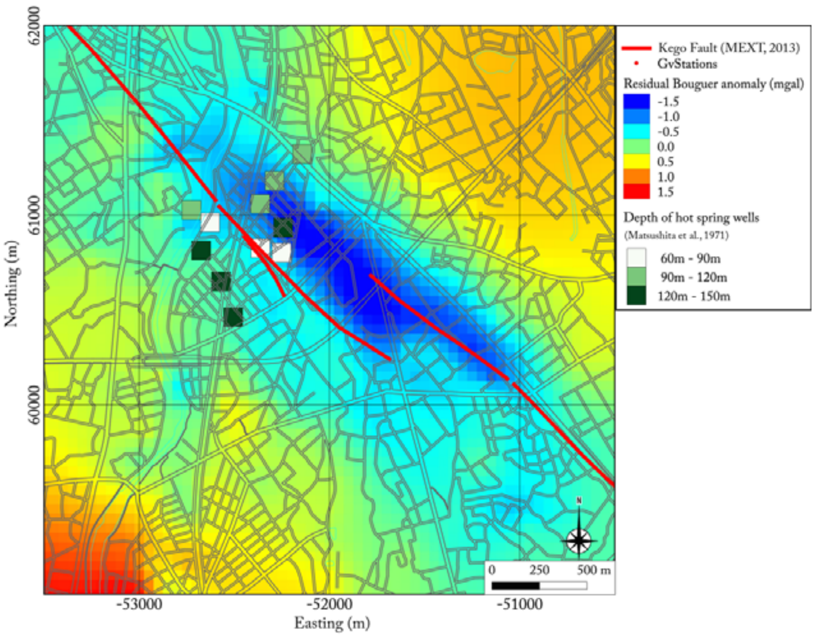


Fig. 6. Residual Bouguer anomaly map. The regional anomaly calculated by first-order polynomial fitting was removed from Fig. 5.

by the deep granite structure, was isolated using the first-order polynomial trend.

Figure 6 clearly indicates a low Bouguer anomaly extension that has a NW-SE direction and coincides with the Kego Fault, as was further confirmed by a trenching survey (Shimoyama *et al.* 2005). The Kego Fault (MEXT 2013) is on the western edge of the low Bouguer anomaly (Fig. 6).

In addition, the location of this low Bouguer anomaly is consistent with a depression of basement rock based on the stratigraphic column (Kimura *et al.* 2013). Most wells are on the northwestern edge of this low anomaly and the surrounding area. Moreover, a high-low-high pattern can be found from the southwest part to the northeast part of the study area. Cretaceous granite and Pleistocene pyroclastic flow deposits correspond to these high Bouguer anomalies.

5. GRAVITY MODELING AND DISCUSSION

According to the geological map (Karakida *et al.* 1994) and the stratigraphic column (Matsushita *et al.* 1971) shown in Fig. 7, it is possible that the survey area can be approximated by two geological layers: Quaternary sediments and Cretaceous granitic basement rock. In order to determine the depth and detailed shape of the basement rock, including fault structure, 3-D gravity modeling was conducted. Distribution of the basement rock was approximated as an aggregation of rectangular prisms whose horizontal dimensions are the same as those of the input residual Bouguer anomaly grid (Cordell and Henderson 1968, Rao *et al.* 1999). The grid adjusted the depth of the basement rock and the prismatic cells that formed a three dimensional (3D) model. This was done to minimize the difference between the obtained gravity data and the model's output. The thicknesses of the prisms represent the depth, from surface to the basement, and a constant density contrast of -400 kg/m^3 was assigned to all prisms. That is, the density of the Quaternary sediments was set at 2200 kg/m^3, while that of the basement rock was 2600 kg/m^3 (GSEAK 1981). Since gravity modeling is non-unique, the depth of granite in the two hot spring wells (Fig. 7) was used to constrain the model. The horizontal size of the cells was set as 50 m and a westward coordinate was set at 41 degrees.

The outcome of the 3D gravity basement analysis can be seen in Fig. 8. Panel (a) shows the depression with a NW-SE trend on the eastern side of the Kego Fault. On the eastern side of the depression, a continuous steep gradient in the level of the gravity basement was detected. The depth of the basement is reducing to the west, including some steep gradients in the level of basements on the western side of the Kego Fault. These depth gradients have a strike of NW-SE and the most eastern step corresponds with the Kego

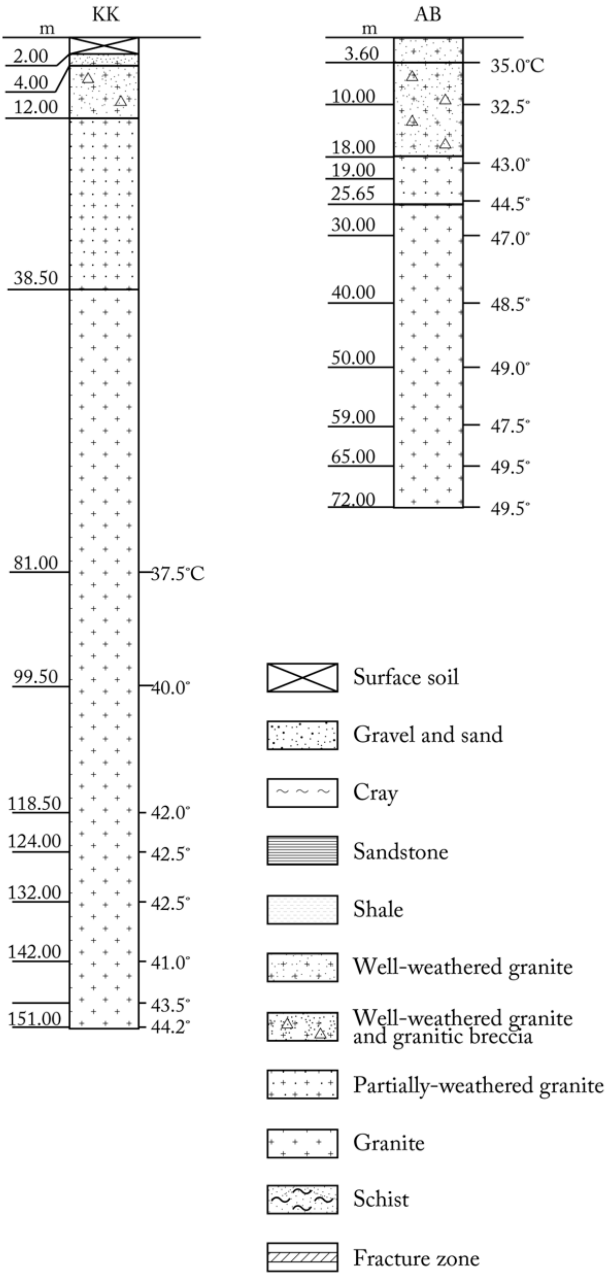


Fig. 7. Stratigraphic column of two wells (well ID: KK and AB) in the Hakata hot spring area (Matsushita *et al.* 1971). The depth of granite in these wells was used to constrain the three-dimensional gravity modeling.

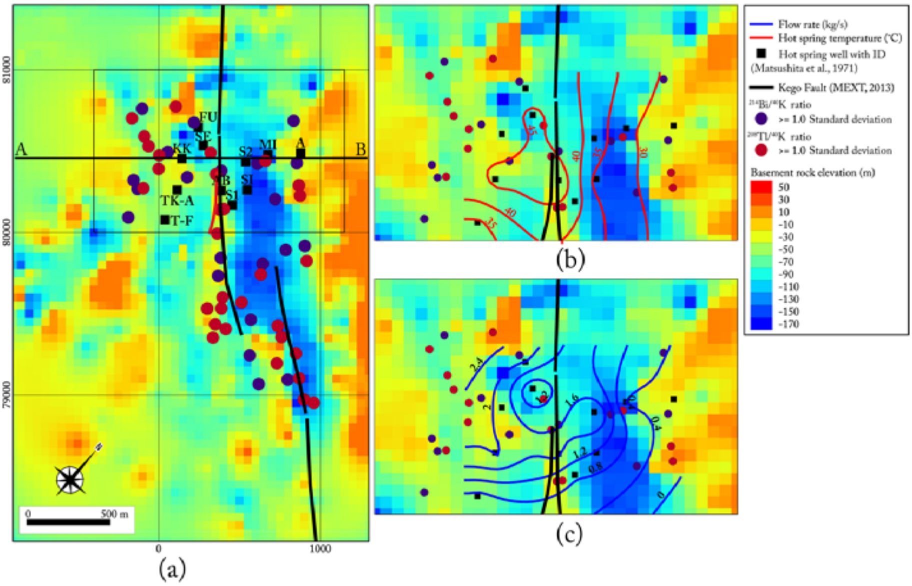


Fig. 8. (a) Gravity basement in meters above sea level with density contrast of 400 kg/m^3. Black rectangle shows the area of panels (b) and (c). (b) Focused map of the Hakata hot springs in comparison with the distribution of hot spring temperatures based on Table 1. (c) Focused map of the Hakata hot springs in comparison with the distribution of the hot spring flow rates [kg/s] based on Table 1. The black line and black rectangle show the location of the Kego Fault (MEXT 2013) and the hot spring wells with well ID, respectively. The coordinate is turning westward at 41 degrees.

Fault. Almost all the hot spring wells are located near these steep depth gradients. Additionally, most of the high-intensity gamma-ray results agree with the depth gradients in the gravity basement.

Figures 8b and c show a comparison of the distributions of hot spring temperatures and flow rates, respectively, based on the hot spring wells data (Table 1). Figure 8b indicates that the high-temperature area is located on and to the west of the Kego Fault. This high-temperature area coincides with high gamma-ray intensities and the steps or steep depth gradients of the gravity basement. In contrast, the hot spring temperature decreases toward eastern side of the Kego Fault. Figure 8c shows that the high flow rate area is located in the western part of the Kego Fault. This high flow rate of hot water can be observed in almost the same location as the high-temperature area; it corresponds to the small depression of the gravity basement. The high gamma-ray intensities correspond to high flow rate area. In comparison with the distribution of the hot spring temperatures, the extent of the high

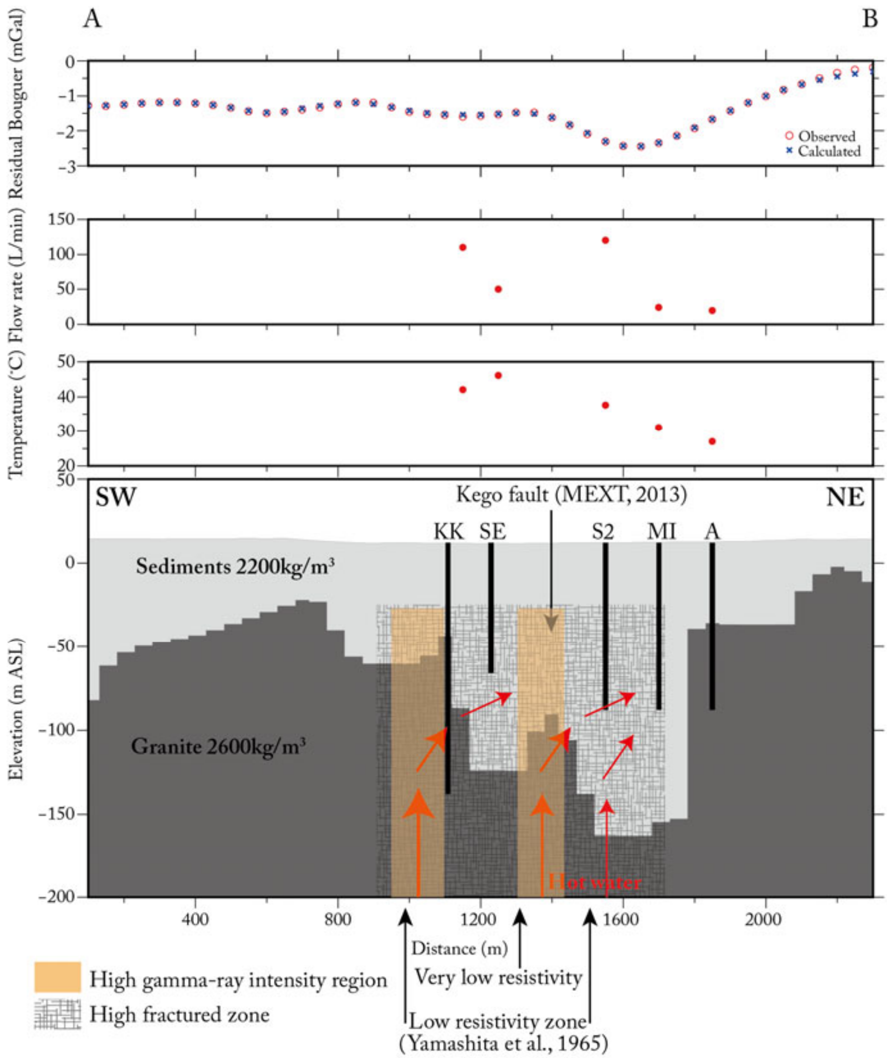


Fig. 9. Profile of line AB in Fig. 8a. The hot water flows up to the near surface through the fault structure. The density of the granite is assumed to be 2600 kg/m^3 and that of the sediments to be 2200 kg/m^3. The main flow path of the hot spring is considered a part of the Kego Fault. KK, SE, S2, MI, and A indicate the hot springs well IDs, which are located near the profile.

flow rate area is narrow. Considering these results, it can be presumed that the hot spring water is coming from deeper part of the high flow rate area and flows toward the eastern and southern areas laterally through the Kego Fault.

Figure 9 shows a cross-section of A-B in Fig. 8a. The depth of basement gradually deepens toward the east; it includes steep depth variations in the Hakata hot springs area. High gamma-ray intensities (^{214}Bi and ^{208}Tl) were detected around these areas. Furthermore, the high-temperature hot springs and high flow rates were observed above these steps.

Results of the electrical resistivity observations (Yamashita *et al.* 1965) indicated that a low resistivity zone (less than 10 Ohm-m) was detected at a depth of more than 30 m in the Hakata hot springs area. Matsushita *et al.* (1971) interpreted this low resistivity zone as a hot spring reservoir by using the stratigraphic column of the hot spring wells.

Taking these results into account, a portion of the steep difference in the level of the gravity basement where the hot springs are concentrated is a part of the Kego Fault. This is similar to a fracture zone created by past activities of the fault. Moreover, the fracture zone around these steep depth gradients acts as a path for hot spring water to travel from the deeper side of the granitic body. It is therefore thought that the hot spring water comes up through this fracture zone from a deeper side in the high flow rate area and then flows toward the east and southeast at a shallow depth.

6. CONCLUSION

This paper presented the analysis of gamma-ray and gravity surveys over the Hakata hot springs area. Two high-intensity gamma-ray areas (^{214}Bi/^{40}K and ^{208}Tl/^{40}K) that coincide with the Kego Fault were detected. The northwestern high-intensity area is located in and around the high temperature and high flow rate area. The distribution of the 3D gravity basement depths enabled the delineation of the interface between the basement rocks and the sediments. In addition, some steep depth gradients of the gravity basement, including the Kego Fault, were detected. We have reported that the results of gamma-ray and gravity surveys indicated that the differences in the level of the basement where the hot spring wells concentrate correspond to a part of the Kego Fault. Fracture zones were created by the past activities of the fault; these enable hot spring water to come up to the surface from the deeper side of the granitic body. It can be presumed that the hot spring water flows laterally to the east and south through the shallow part of the Kego Fault in relation to the distributions of hot spring temperatures and high flow rates. Although there is no hot spring well in the southeastern area where similar underground structures were detected, it is possible that hot water is moving up toward the surface from those structures.

These geophysical survey methods are likely to remain appropriate in the future, as they were quick and cost-effective ways to evaluate the hydrothermal system of this hot springs area. These methods allowed not only an understanding of the underground structure of low-temperature hydrothermal systems, but also an assessment of the sustainability of hot spring usage. To understand the sustainability of hot spring usage more fully, a further study that would involve geothermal modeling and numerical simulation of hydrothermal systems should be conducted.

References

Abdelzaher, M., J. Nishijima, G. El-Qady, E. Aboud, O. Masoud, M. Soliman, and S. Ehara (2011), Gravity and magnetotelluric investigations to elicit the origin of Hammam Faraun hot spring, Sinai peninsula, Egypt, *Acta Geophys.* **59**, 3, 633-656, DOI: 10.2478/s11600-011-0006-4.

Aswathanarayana, A. (1985), *Principles of Nuclear Geology*, Balkema, Rotterdam.

Cordell, L., and R.G. Henderson (1968), Iterative three-dimensional solution of gravity anomaly data using a digital computer, *Geophysics* **33**, 4, 596-601, DOI: 10.1190/1.1439955.

GSJ AIST (2014), Seamless digital geological map of Japan 1:200,000, Jan 14, 2014 version, Geological Survey of Japan, National Institute of Advanced Industrial Science and Technology.

GSEAK (1981), Fukuoka Jibanzu, Geotechnical data of subsoil in Fukuoka, Geological Survey Enterprises Association Kyushu, Japan (in Japanese).

HERP (2007), Long-term evaluation of the Kego active faults, Headquarters for Earthquake Research Promotion, Japan, available from: http://www.jishin.go.jp/main/chousa/13feb_chi_kyushu/k_5.pdf (accessed: 19 March 2007) (in Japanese).

HERP (2013), Long-term evaluation of the active faults in Kyushu, Headquarters for Earthquake Research Promotion, Japan, available from: http://www.jishin.go.jp/main/chousa/13feb_chi_kyushu/k_honbun.pdf (accessed: 1 February 2013) (in Japanese).

Karakida, Y., S. Tomita, S. Shimoyama, and K. Chijiwa (1994), Geology of the Fukuoka district. Quadrangle series, Scale 1:50000, Geological Survey of Japan, Tsukuba, 192 pp. (in Japanese with English abstract).

Kimura, K., Y. Kou, and Y. Hanashima (2013), Subsurface geologic structures of the Quaternary deposits underlying the Fukuoka plain, Fukuoka Prefecture, western Japan. **In:** *Seamless Geoinformation of Coastal Zone*, Coastal zone around Fukuoka, Digital Geological Map S-3, Geological Survey of Japan, AIST (in Japanese with English abstract).

Matsushita, H., T. Miki, and A. Yamashita (1971), An overturned structure observed in the southern part of Fukuoka city, *Sci. Rep. Shimabara Volcano Observ. Fac. Sci. Kyushu Univ.* **7**, 1-8 (in Japanese with English abstract).

McCay, A.T., T.L. Harley, P.L. Younger, D.C.W. Sanderson, and A.J. Cresswell (2014), Gamma-ray spectrometry in geothermal exploration: State of the art techniques, *Energies* **7**, 4757-4780, DOI: 10.3390/en7084757.

MEXT (2013), Report of intensive survey and observation on the Kego fault zone, Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan, 25-29 (in Japanese).

Mogi, T., and M. Jinguuji (1993), Active fault survey using gamma-ray spectrometer – Automatic appraisal analysis of data obtained from NaI(Tl) scintillation detector, *J. Seismol. Soc. Japan* 2nd ser., **46**, 1-8 (in Japanese with English abstract).

Mogi, T., and S. Okada (1990a), Gammra-ray spectra survey in geothermal area, *J. Geothermal Res. Soc. Japan* **12**, 3, 295-308, DOI: 10.11367/grsj1979.12.295 (in Japanese with English abstract).

Mogi, T., and S. Okada (1990b), Gamma-ray spectra survey in Futsukaichi hot springs, Fukuoka Prefecture: The reliability of spectre data, *J. Japan Soc. Eng. Geol.* **31**, 11-19 (in Japanese with English abstract).

Murata, Y. (1993), Estimation of optimum average surficial density from gravity data: an objective Bayesian approach, *J. Geophys. Res.* **98**, B7, 12097-12109, DOI: 10.1029/93JB00192.

Rao, P.R., and K.V. Swamy, and I.V. Radhakrishna Murthy (1999), Inversion of gravity anomalies of three-dimensional density interfaces, *Comput. Geosci.* **25**, 8, 887-896, DOI: 10.1016/S0098-3004(99) 00051-5.

Represas, P., F.A.M. Santos, J. Ribeiro, J.A. Ribeiro, E.P. Almeida, R. Gonçalves, M. Moreira, and L.A. Mendes-Victor (2013), Interpretation of gravity data to delineate structural features connected to low-temperature geothermal resources at Northeastern Portugal, *J. Appl. Geophys.* **92**, 30-38, DOI: 10.1016/j.jappgeo.2013.02.011.

Saibi, H., J. Nishijima, T. Hirano, Y. Fujimitsu, and S. Ehara (2008), Relation between structure and low-temperature geothermal systems in Fukuoka city, southwestern Japan, *Earth Planets Space* **60**, 8, 821-826, DOI: 10.1186/BF03352833.

Shimizu, H., H. Takahashi, T. Okada, T., Kanazawa,Y. Iio, H. Miyamachi, T. Matsushima, M. Ichiyanagi, N. Uchida, T. Iwasaki, H. Katao, K. Goto, S. Matsumoto, N. Hirata, S. Nakao, L. Uehira, M. Shinohara, H. Yakiwara, N. Kame, T. Urabe, N. Matsuwo, T. Yamada, A. Watanabe, K. Nakahigashi, B. Enescu, K. Uchida, S. Hashimoto, S. Hirano, T. Yagi, Y. Kohno, T. Ueno, M. Saito, and M. Hori (2006), Aftershock seismicity and fault structure of the 2005 West Off Fukuoka Prefecture Earthquake (MJMA7.0) derived from urgent joint observations, *Earth Planets Space* **58**, 12, 1599-1604, DOI: 10.1186/BF03352668.

Shimoyama, S., N. Iso, T. Matsuda, T. Ichihara, N. Chida, M. Okamura, T. Mogi, S. Suzuki, H. Ochiai, S. Nagasawa, H. Imanishi, F. Kawabata, H. Yakabe, M. Ooteki, and K. Matsuura (2005), Trenching study at Yakuin site across the Kego fault, Fukuoka City, West Japan, *Active Fault Res.* **25**, 117-128, (in Japanese with English abstract).

Sultan, S.A., F.A.M. Santos, and T. Arafa-Hamed (2012), Delineating active faults by using integrated geophysical data at northeastern part of Cairo, Egypt, *NRIAG J. Astron. Geophys.* **1**, 33-44, DOI: 10.1016/j.nrjag.2012.11.004.

Xu, C., H. Wang, Z. Luo, J. Ning, and H. Liu (2015), Multilayer stress from gravity and its tectonic implications in urban active fault zone: A case study in Shenzhen, South China, *J. Appl. Geophys.* **114**, 174-182, DOI: 10.1016/j.jappgeo.2015.01.017.

Yamashita, A., A. Okawa, K. Ota, and Y. Ueki (1965), On the electrical prospecting for ground water in Fukuoka city, *J. Min. Inst. Kyushu* **33**, 407-414 (in Japanese with English abstract).

Received 24 May 2015
Received in revised form 20 January 2016
Accepted 5 April 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2305-2321
DOI: 10.1515/acgeo-2016-0103

Natural Radioactivity at the Sin Quyen Iron-Oxide-Copper-Gold Deposit in North Vietnam

Dinh Chau NGUYEN[1], Phon Le KHANH[3], Paweł JODŁOWSKI[2], Jadwiga PIECZONKA[1], Adam PIESTRZYŃSKI[1], Hao Duong VAN[3], and Jakub NOWAK[2]

[1]AGH University of Science and Technology (AGH-UST), Faculty of Geology, Geophysics and Environmental Protection, Krakow, Poland;
e-mail: Nguyen.Chau@fis.agh.edu.pl

[2]AGH University of Science and Technology (AGH-UST), Faculty of Physics and Applied Computer Science, Krakow, Poland

[3]University of Mining and Geology (UMG), Hanoi, Vietnam

Abstract

The field radiometric and laboratory measurements were performed at the Sin Quyen copper deposit in North Vietnam. The field gamma-ray spectrometry indicated the concentration of uranium ranging from 5.5 to 87 ppm, thorium from 5.6 to 33.2 ppm, and potassium from 0.3 to 4.7%. The measured dose rates ranged from 115 to 582 nGy/h, the highest doses being at the copper ore. Concentrations in the solid samples were in the range of 20-1700 Bq/kg for uranium, 20-92.7 Bq/kg for thorium, and 7-1345 Bq/kg for potassium. The calculated doses were from 22 to 896 nGy/h; both measured and calculated dose rates are mostly related to uranium. Concentrations of radium in water samples were below 0.17 Bq/L. Uranium in water samples was significantly higher than the hydrogeological background; the maximum of 13 Bq/L was at the waste zone pool, but neither radium nor uranium were present in tap water. Radon concentration in the dwelling air was from 42 to 278 Bq/m^3 for ^{222}Rn and from 8 to 193 Bq/m^3 for ^{220}Rn. The estimated committed dose rates

Ownership: Institute of Geophysics, Polish Academy of Sciences;

were principally related to ^{222}Rn concentration and ranged from 1.1 to 8.1 mSv/y.

Key words: IOCG deposit, natural radionuclides, dose rate.

1. INTRODUCTION

Iron oxide copper gold (IOCG) deposits are usually characterized by a broad spectrum of various minerals. Apart from iron and copper bearing minerals such as, *e.g.*, magnetite, pyrite and chalcopyrite, there are also minerals containing gold, silver, rare Earth elements and natural radionuclides (Baker 2005, Corriveau *et al.* 2007, Hunt *et al.* 2005, Piestrzyński 1989, Sillitoe 2003, Williams *et al.* 2005, Zhao and Zhou 2011). Mining of the ore from such a deposit usually leads to an increased concentration of natural radioactive elements in the mined material which may cause a significant increase of radiation in the environment and therefore it would be a health hazard to the miners and to population of the region. To investigate this effect, both field *in situ* measurements and laboratory analyses of the collected samples were performed. The field survey included determination of natural radionuclides' concentration through gamma-ray spectrometry and measurements of the absorbed dose rates through gamma-ray dosimetry; concentration of radon in the air of dwellings in the vicinity of the mining area was measured using the track detectors. The field measurements and solid sample collection were performed mainly along the profile crossing the strike line of geological structure, at the active copper mining area and at the waste dumps, reservoir sediments and floatation products. Water samples were collected from the surface water pools at the open pit mine, waste dumps and from: Red River, the private wells and tap water. All the collected samples were analyzed for natural radionuclides. On the basics of the obtained data, gamma-ray dose rates and annual committed doses were estimated.

2. REGION OF INVESTIGATION

The Sin Quyen IOCG deposit in Lao-Cai North Vietnam is localized on the right of the Red River, which is a natural Vietnam-China boundary in this region (Fig. 1a). The deposit was discovered in 1961, owing the field radiometric and magnetic surveys. The area and copper wealth of this deposit amount to above 120 ha and 52 million tons, respectively. The principal rocks occurring at the Sin Quyen copper deposit are Paleozoic metamorphic gneiss, sandstone, mudstone, intrusive granitoid and skarn; they occur nearly vertically and trend in the NW-SE direction (Fig. 1b). The ore bodies are often met at contact zones between intrusives and sediments as lens that are several tens meters long and dozen meters thick. Since 2006 this IOCG deposit had been mined as an open pit mine. Close to the copper mine there is

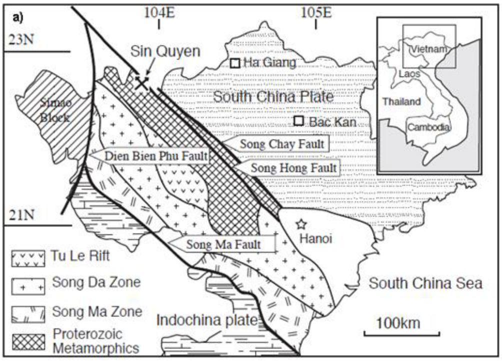


Fig. 1a. Sketch of tectonics of the Northern Vietnam and location of Sin Quyen Deposit (Ishihara *et al.* 2011, McLean 2001).

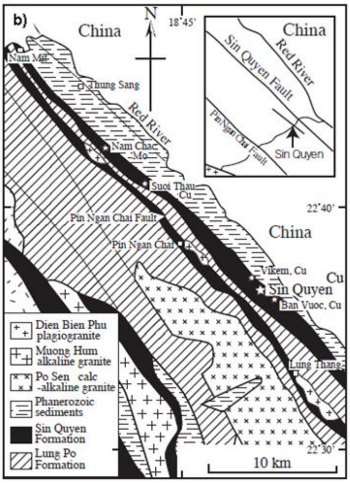


Fig. 1b. Geological map of the Sin Quyen deposit (Ishihara *et al.* 2011, Bui *et al.* 2004, McLean 2001).

floatation plant and melting factory. The annual spoil rock excavation, ore exploitation and copper metal product amount to over 6 million cubic meters, one million tons and near 12 000 tons, respectively (Le *et al.* 2015).

3. MEASUREMENT METHODS

3.1 Field measurement and sample collection

The field radiometric measurements were performed using the 512 channels potable spectrometer GF5 with NaI(Tl) scintillation crystal of the Czech Gamma Surveyor Company™; at each measurement point the detector was placed one meter above the Earth surface. The measured data were given on display as the concentration of K [%], eU [ppm], eTh [ppm], and gamma absorbed dose rate D [nGy/h]. For each point the measurements were performed three times and the final results and their adequate uncertainty were estimated using the data obtained. The coordinates of the measurement points, the average concentrations and uncertainty of uranium, thorium and potassium are summarized in Table 1.

The solid samples were collected from the selected points, such as the mining waste dumps and the sediment of floatation reservoir. The samples represented the rocks and ores occurring at the studied region. Additionally, some samples were collected from the interesting places, such as high dose rate or/and high content of the natural radionuclides and so on. The mass of the solid samples varied from 0.5 to 1.0 kg, and every sample was placed in the plastic bag. The geographic coordinates of the solid sampling places are summarized in Table 2.

Table 1

Potassium, radium, and thorium concentrations and gamma adsorbed dose rates measured by a potable gamma spectrometer

Point No.	Coordinate		K [%]	U [ppm]	Th [ppm]	Dose [nGy/h]
	latitude	longitude				
1	22°36′45″63	103°48′54″24	2.35±0.25	12.4±3.6	16.8±1.8	143±18
2	22°36′39″00	103°48′54″27	2.59±0.11	8.30±0.61	16.9±0.4	125±6
3	22°36′38″54	103°48′53″99	2.69±0.09	8.30±0.32	14.7±1.2	120±5
4	22°36′37″55	103°48′52″21	3.24±0.16	8.00±0.75	18.1±0.7	134±4
5	22°36′37″35	103°48′51″97	2.96±0.06	7.37±0.41	15.8±0.4	120±2
6	22°36′37″45	103°48′51″76	2.95±2.1	11.8±1.9	19.9±2.0	156±7
7	22°36′35″73	103°48′52″54	3.56±0.6	9.37±0.35	23.7±3.5	159±18
8	22°36′35″18	103°48′52″44	2.88±0.06	22.9±1.1	17.5±1.1	212±4
9	22°36′35″75	103°48′50″51	2.93±0.32	5.5±1.5	13.0±2.7	103±19
10	22°36′43″29	103°48′49″50	4.52±0.79	10.1±2.5	27.5±3.9	185±33
11	22°36′33″64	103°48′51″19	3.66±0.07	8.40±1.0	22.0±2.7	151±5
12	22°36′33″64	103°48′50″60	2.20±0.23	13.5±1.5	28.6±2.2	177±3
13	22°36′33″47	103°48′49″06	2.77±0.01	57.2±0.4	26.8±0.3	428±2
14	22°36′32″91	103°48′48″54	4.68±0.25	7.2±0.7	15.8±0.2	142±7
15	22°36′34″31	103°48′45″34	1.55±0.15	28.8±1.9	21.0±1.0	236±10
16	22°36′36″48	103°48′43″96	2.91±0.21	87.1±2.5	19.5±2.4	581±11
17	22°36′35″64	103°48′44″77	2.03±0.30	62.8±2.3	21.5±3.1	437±5
18	22°36′36″15	103°48′43″19	2.31±0.11	55.3±1.8	22.9±1.3	402±8
19	22°36′33″76	103°48′48″92	2.82±0.06	66.6±3.5	22.4±1.7	472±25
20	22°36′33″53	103°48′48″81	2.42±0.13	23.7±3.4	16.5±1.7	208±19
21	22°36′33″49	103°48′48″46	3.93±0.15	11.2±0.7	18.4±0.8	161±5
22	22°36′33″46	103°48′48″43	3.63±0.04	10.1±1.9	16.0±0.2	145±10
23	22°36′33″36	103°48′48″29	4.18±0.21	8.15±1.1	15.0±0.5	139±5
24	22°36′33″20	103°48′48″40	4.12±0.25	7.5±0.3	14.6±0.8	133±1
25	22°36′32″91	103°48′48″47	4.7±0.29	6.0±0.4	6.0±0.1	111±2
26	22°36′32″91	103°48′48″29	3.56±0.21	12.1±1.6	18.4±2.0	162±11
27	22°36′34″10	103°48′46″64	1.90±0.19	38.2±3.5	33.2±5.4	325±4
28	22°36′33″10	103°48′48″05	1.88±0.09	25.4±1.3	11.9±1.2	200±4
29	22°36′32″84	103°48′48″01	3.97±0.35	10.1±0.1	17.9±0.6	145±6
30	22°36′31″96	103°48′47″18	3.28±0.30	12.6±2.4	25.0±3.3	177±12
31	22°36′32″48	103°48′47″39	3.36±0.29	10.5±0.7	14.0±0.5	139±9
32	22°36′32″11	103°48′46″72	1.88±0.15	7.8±1.1	14.2±1.2	105±8
33	22°36′31″52	103°48′46″03	2.70±0.15	24.0±0.2	18.5±0.1	218±2
34	22°36′31″69	103°48′45″96	3.27±0.06	9.20±1.1	20.3±0.4	146±10
35	22°36′31″46	103°48′45″61	3.26±0.29	8.9±1.1	24.8±6.5	156±9
36	22°36′30″93	103°48′45″41	1.63±0.03	10.4±0.8	19.4±0.7	129±6
37	22°36′30″54	103°48′45″09	3.24±0.22	5.70±0.7	14.6±0.6	112±8
38	22°37′05″25	103°48′43″85	0.76±0.07	24.4±1.8	7.4±0.9	167±11
39	22°37′04″51	103°48′45″33	0.38±0.05	30.5±0.9	5.6±0.7	193±6
40	22°37′24″71	103°48′38″05	4.0±0.6	24.0±2.4	15.0±0.7	227±18
41	22°37′27″47	103°48′28″39	2.98±0.20	29.0±0.7	19.9±1.1	253±4
Min:			1.55	5.5	6.0	103
Max:			4.70	87.1	33.2	581
Mean:			3.04	19.8	19.0	199

Table 2

^{40}K, ^{226}Ra, ^{232}Th contents of the solid samples measured by gamma spectrometer coupled with HPGe detector and estimated gamma absorbed dose rates

Sample	Coordinate		^{40}K [Bq/kg]	^{226}Ra [Bq/kg]	^{232}Th [Bq/kg]	Dose [nGy/h]
	latitude	longitude				
W25	22°36′34″42	103°48′59″38	697±21	114±3.4	41.5±1.2	106±3
W13	22°36′34″02	103°48′58′65	937±28	40.5±1.2	51.1±1.5	88±3
W15	22°36′32′05	103°48′59″19	632±19	128±4	51.7±1.6	116±4
W21	22°36′30″71	103°48′59″59	795±24	84.5±2.6	58.6±1.8	107±3
W26	22°36′29″94	103°48′56″09	466±14	640±19	65.4±2.0	354±11
W20	22°36′32″80	103°48′51″48	637±19	1699±51	46.0±1.4	839±25
W27	22°36′32″54	103°48′51″90	417±13	163±5	33.3±1.0	113±3
W12	22°36′29″94	103°48′56″09	557±17	280±9	83.4±2.5	116±6
W23	22°36′32″54	103°48′51″90	382±11	740±22	90.6±2.7	202±12
W07	22°36′32″61	103°48′51″62	632±20	401±12	62.8±1.9	412±8
W08	22°36′32″60	103°48′52″26	484±15	671±20	75.7±2.3	249±11
W17	22°36′31″01	103°48′54″79	964±29	49.6±1.5	66.1±2.0	376±3
W24	22°36′29″85	103°48′57″18	710±21	57.7±1.7	57.2±1.7	102±3
W16	22°36′30″16	103°48′55″32	770±23	44.8±1.4	37.7±1.2	90±2
W31	22°36′28″56	103°48′53″79	556±17	144±5	92.7±2.8	75±5
W02	22°36′30″71	103°48′59″06	115±5	41.7±1.3	26.1±1.0	145±1
W28	22°36′29″18	103°48′54″59	7±2	30.0±1.0	29.8±1.0	40±1
W22	22°36′29″80	103°48′55″18	490±15	68.3±2.0	66.4±2.0	32±3
W18	22°36′30″00	103°48′55″36	581±18	66.6±2.1	49.5±1.5	92±3
W19	22°36′29″61	103°48′55″08	528±16	58.0±1.7	81.7±2.5	84±3
W30	22°36′38″01	103°50′45″41	690±21	33.4±1.0	26.4±1.0	98±2
W29	22°36′28″49	103°48′53″55	390±12	76.9±2.2	42.0±1.3	60±2
W05	22°36′30″18	103°48′57″28	91±5	302±10	20.1±1.0	77±5
W01	22°36′30″18	103°48′58″40	123±5	389±12	25.8±1.0	155±6
W11	22°36′29″94	103°48′56″09	136±5	296±9	26.2±1.0	200±5
W10	22°37′27″54	103°48′38″65	678±20	583±17	64.7±1.8	336±10
W09	22°37′24″97	103°49′13″06	536±16	553±17	66.1±1.9	317±10
W04	22°37′27″99	103°48′28″39	1345±40	47.1±1.4	64.1±2.0	116±3
W33	22°37′25″23	103°49′13′69	326±11	20±5	31.8±1.0	42±3
W03	22°36′46″69	103°49′06″21	752±23	49.3±1.5	77.5±2.3	100±3
W06	22°37′02″57	103°48′50″91	663±20	71.4±2.2	61.2±1.9	97±3
W14	22°37′01″74	103°48′52″39	843±25	202±6	54.5±1.7	161±5
Min:			7	20	20	32
Max:			1345	1699	92	839
Mean:			560	254	54	161

Table 3

Measured concentrations of some selected ions and ^{238}U, ^{234}U, ^{226}Ra, and ^{228}Ra in the water samples

Sample	Coordinate		pH	TDS [g/L]	Ion concentration [mg/L]							^{226}Ra [mBq/L]	^{228}Ra [mBq/L]	^{234}U [mBq/L]	^{238}U [mBq/L]
	latitude	Longitude			Na^+	Ca^{2+}	Fe^{2+}	Cu^{2+}	HCO_3^-	Cl^-	SO_4^{2-}				
SQ-1	22°36′35″31	103°49′00″88	3.4	2.6	4.6	244	192.6	9.4	≤0.5	3.1	1850	67±5	70±10	12700±1500	13100±1600
SQ-2	22°36′32″80	103°48′51″48	7.6	1.5	61.5	280	0.034	0.004	152	15.6	851	83±5	≤30	180±19	330±37
SQ-3	22°36′29″94	103°48′56″09	7.9	1.5	33.9	319	0.003	0.004	142	7.8	908	81±4	60±7	180±20	320±36
SQ-4	22°36′32″61	103°48′51″62	7.9	1.8	57.5	347	≤0.002	0.005	246	13.1	1020	166±20	163±22	168±18	343±38
SQ-5	22°37′25″67	103°47′21″77	8.0	0.39	3.5	46.8	≤0.002	0.0003	272	4.7	5.5	≤5	≤30	≤0.5	≤0.5
SQ-6	22°37′24″97	103°49′13″06	7.8	0.988	12.1	243	≤0.002	0.004	55.8	7.8	612	≤5	≤30	97.4±10.4	158±18
SQ-7	22°37′27″54	103°48′38″65	3.6	0.946	45.4	196	0.103	0.009	≤0.5	13.7	622	≤5	≤30	77.8±8.3	96.3±7.0
SQ-8	22°37′25″23	103°49′13″69	3.6	0.960	45.1	201	1.367	0.007	≤0.5	15.6	624	20±3	≤30	26.3±2.9	62.8±7.2
SQ-9	22°37′02″57	103°48′50″91	7.2	0.487	19.6	84.3	0.027	0.009	82	9.4	237	≤5	≤30	1134±120	1237±140
SQ-10	22°36′46″69	103°49′06″21	7.3	0.095	3.20	11.7	0.033	0.014	30.0	4.70	13.1	≤5	≤30	29.2±3.1	50.0±5.4
SQ-11	22°37′18″14	103°49′16″79	7.0	0.098	3.37	11.9	0.071	0.007	38.0	3.10	14.2	≤5	≤30	22.9±2.3	45.2±5.0

The water samples were collected using a specific plastic cup and placed into 2 dm^3 plastic bottle. Before collecting, the cup and bottle were cleaned up using distilled water and rinsed twice by the investigated water. The coordinates of the water collecting places are presented in Table 3. All the water and solid samples were sent to AGH University of Science and Technology in Kraków, Poland.

The ^{222}Rn and ^{220}Rn concentrations in the air were measured for 22 houses surrounding the Sin Quyen IOCG deposit using Radosys detector. The Radosys detector is composed of two CR-39 track detectors and two diffusion chambers, one detector for each chamber. The detectors are designed for detecting the ^{222}Rn and ^{220}Rn activity; at every house the Radosys detec-

Table 4

Concentrations of ^{222}Rn and ^{220}Rn in the dwelling air [Bq/m^3] and annual dose rate [mSv/a] inside and outside of house

Point No.	Inner			Outer		
	^{222}Rn [Bq/m^3]	^{220}Rn [Bq/m^3]	Dose [mSv/a]	^{222}Rn [Bq/m^3]	^{220}Rn [Bq/m^3]	Dose [mSv/a]
1	200	60	5.38	185	19	4.77
2	165	40	4.38	123	33	3.28
3	146	10	3.74			
4	127	42	3.44	194	13	4.96
5	278	36	7.21	105	77	3.08
6	194	28	5.05	139	16	3.59
7	207	33	5.40	110	9	2.82
8	145	15	3.74	43	117	1.74
9	193	23	4.99	185	19	4.77
10	214	17	5.49	99	95	3.03
11	138	30	3.65	177	88	4.54
12	207	15	5.30	159	36	4.21
13	191	22	4.94	214	38	5.61
14	191	22	4.94	140	16	3.62
15	139	8	3.55	163	13	4.18
16	151	18	3.91	155	33	4.09
17	163	13	4.18	196	15	5.02
18	237	22	6.10	208	22	5.36
19	206	40	5.42	175	74	4.82
20	123	13	3.17	192	14	4.92
21	42	193	2.14	181	27	4.71
22	247	21	6.34	156	23	4.06
Min:	42	8	2.14	43	9	1.74
Max:	247	193	7.21	214	117	5.61
Mean:	177.4	32.8	4.65	157	38	4.15

tor was hung at a height of the 1.5 to 1.8 m from the Earth surface and at a distance of 2 m from door and walls. The detectors were exposed for a period of three months (from November 2014 to February 2015). The coordinates of the investigated houses are shown in Table 4.

3.2 Laboratory methods

3.2.1 Determination of uranium, thorium and potassium in a solid sample

For gamma spectrometric measurement, the samples were milled and then dried in an oven at a temperature of 120 °C for 24 h to ensure that moisture is completely removed, then accurately weighed and packed in radon-impermeable aluminums cylindrical Marinelli beakers of 720 ml capacity and sealed to prevent escape of radon gas. The tightly sealed samples were left for at least 22 days to reach secular equilibrium between the ^{222}Rn and ^{226}Ra (Jodlowski and Kalita 2010).

The activity concentrations of ^{40}K, ^{226}Ra, and ^{232}Th were measured by gamma spectrometer coupled with a semiconductor HPGe detector (Canberra GX4020) with relative efficiency of 42% and resolution of 1.9 keV for 1332 keV line. The gamma spectrometer was calibrated using the IAEA reference materials RGU, RGTH, RGK as standard sources for the ^{226}Ra, ^{232}Th, and ^{40}K, respectively.

The gamma lines of 1001 keV from ^{234m}Pa and 609, 1120, and 1765 keV from ^{214}Bi were used to determine the activity concentration of ^{238}U and ^{226}Ra, while those of ^{232}Th were determined from the gamma lines of 911 and 969 keV from ^{228}Ac, and 583 and 2614 keV from ^{208}TI. For ^{40}K, its activity concentration was determined from its 1461 keV gamma line. The maximum counting time for samples was 40 hours to minimize uncertainty to less than 3% for low-activity samples. The self gamma absorption resulted from the difference of density of the solid samples and standard ones was introduced follow the method described by Jodlowski (2006). The determined concentrations of ^{226}Ra, ^{232}Th, and ^{40}K are summarized in Table 2.

3.2.2 Determination of chemical composition and radium and uranium isotopes in a water sample

Chemical composition

The chemical composition was analyzed using an ICP-AES PerkinElmer Optima 7300 DV spectrometer, calibrated with a multi-element standard solution of the Merck® company. The induced couple plasma instrument worked with a cooling argon flow of 14 L/min, a reflected RF power of 1350 W, both auxiliary gas and nebulizer flow rates of 1.0 L/min, a sample intake of 0.8 mL/min. The limit of determination depended on the individual element and ranged from a few to tens ppb with 3% of uncertainty.

Radium isotopes

The radium isotopes were precipitated from the water sample of one litter together with barium element as the sulphate compound. Then the obtained sample was cleaned up and separated from the other isotopes using the procedure described by Nguyen *et al.* (1997). The final precipitate was placed in the glass vial 22 ml of capacity and mixed with 12 ml of gel scintillation cocktail produced by Perkin Elmer™ company and measured using the α/β 1414 Wallac Liquid Scintillation Counter®. To eliminate the background radiation originating from the chemical reagents, cosmic and electronic noise, the background sample from the distilled water was prepared together with the series of the investigated water samples. Every sample was measured for two hours daily until the day when the expected equilibrium between ^{226}Ra and its short-lived products was established (above 21 days). The contents of ^{226}Ra and ^{228}Ra in the measured water sample were estimated on the basis of the dependence of the net measured count rates in the α and β channels on the time elapsed from the precipitation radium in water sample. The limit of detection is equal to 5 and 30 mBq per sample for ^{226}Ra and ^{228}Ra, respectively. The uncertainty of the concentration of ^{226}Ra in water sample was calculated using the uncertainty of the net count rates in alpha channel of the water sample and uncertainty of the chemical efficiency obtained from the ^{226}Ra standard solution sample in accordance with the propagation law. The uncertainty of ^{228}Ra was calculated using the uncertainties of the net count rates measured in both alpha and beta channels as well as the uncertainty of the chemical efficiency. The calculation of the uncertainty for ^{226}Ra and ^{228}Ra was described in detail by Nguyen *et al.* (1997).

Uranium isotopes

The determination of the uranium isotopes in the water sample followed the method described by Skwarzec (1997). The uranium was precipitated from the water sample of one litter together with manganese oxide as ammonium uranyl compound. The uranium in the precipitate was separated from the other elements using HCl solution and ion exchange column with Dowex 100-200 mesh. Finally, the uranium in the obtained elute was again precipitated using neodymium chloride and placed onto the plastic membrane filter, 0.1 μm of porosity. The obtained sample was measured using Canberra alpha spectrometer with semiconductor detector of the PIPS type. The contents of ^{238}U and ^{234}U in the water sample were estimated using the measured net count rates in the peaks of the adequate isotopes and net count rate in the peak of the ^{232}U − an isotopic tracer. The known amount of ^{232}U was added into water sample at the beginning of the preparation procedure. The concentration, uranium, and radium isotopes in water samples are presented in Table 3. The uncertainties of the concentration of ^{234}U and ^{238}U in water sample were calculated using the uncertainties of the net count rates of the peaks belonging to the ^{234}U and ^{238}U isotopes respectively, and uncertainty of the count rates of the peak belonging to the ^{232}U and uncertainty of the ^{232}U standard solution in accordance with the propagation law. The limit of detection for both isotopes (^{234}U and ^{238}U) amounted to 0.5 mBq per sample.

Radon isotopes

After exposition, the track detectors were collected and transferred to the National Atomic Energy Agency of Vietnam, where the detectors were chemically treated and the ^{222}Rn and ^{220}Rn activities were determined following the Radosys procedures (Radosys 2013). The determined concentrations of radon isotopes in the air of the dwellings are summarized in Table 4.

4. RESULTS AND DISCUSSION

4.1 Field gamma spectrometric and gamma absorbed dose rates

The concentrations of uranium, thorium, potassium in the rocks and soil as well as gamma dose rates measured at the points of the profile are summarized in Table 1. The minimum, maximum, average and median values amounted to 0.4, 4.7, 3.0, and 3.0% for potassium; 3.3, 87.1, 20.2, and 11.2 ppm for uranium; 5.4, 33.2, 17.9, and 18 ppm for thorium and 115, 581, 196, and 159 nGy/h for gamma absorbed dose rate. The average concentrations of all mentioned elements are higher than average concentrations of the adequate element in the Earth crust, which are equal to 2.6%, 3.5 and 10 ppm for potassium, uranium and thorium respectively (Lange 1972). The average gamma absorbed dose rate at the deposit is above three times higher than the worldwide average natural dose to human (59 nGy/h) (UNSCEAR 2000). The relation between the measured gamma absorbed dose rates is mostly related to the uranium concentration (Fig. 2).

4.2 Laboratory gamma spectrometric measurements

Table 2 summarizes the measured activity concentrations of ^{40}K, ^{226}Ra, and ^{232}Th in the solid samples. The minimum, maximum, average and median values amounted to 7, 1345, 560, and 569 Bq/kg for ^{40}K, 20, 1699, 255, and 99 Bq/kg for ^{226}Ra and 20, 92.7, 54, and 56 Bq/kg for ^{232}Th. The average specific activity of ^{226}Ra is six times higher than that of the Earth crust (40 Bq/kg) (UNSCEAR 2000).

The gamma absorbed dose rate was calculated based on the measured activity concentrations of ^{226}Ra, ^{232}Th, and ^{40}K in the solid samples using the following formula (UNSCEAR 2000):

$$D[\mathrm{nGy/h}] = 0.041\cdot \mathrm{K} + 0.462\cdot \mathrm{Ra} + 0.604\cdot \mathrm{Th}\ , \tag{1}$$

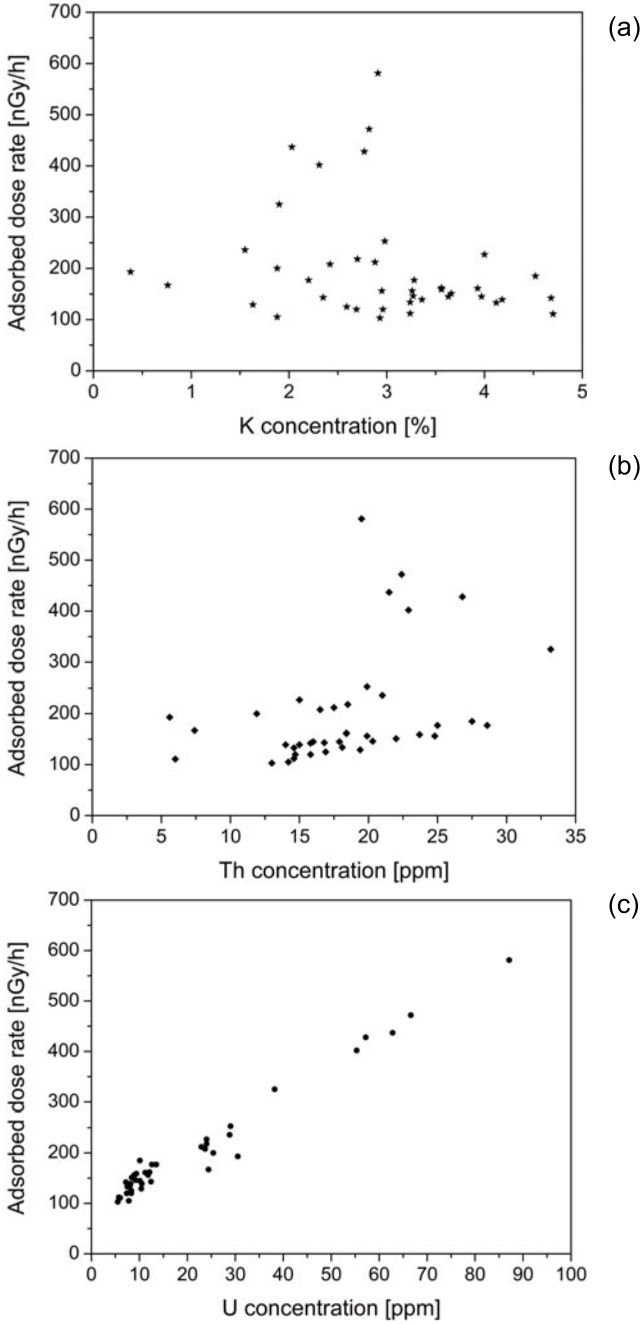


Fig. 2. Relation between the gamma absorbed dose rates and potassium (a), thorium (b) and uranium (c) measured by portable gamma spectrometer.

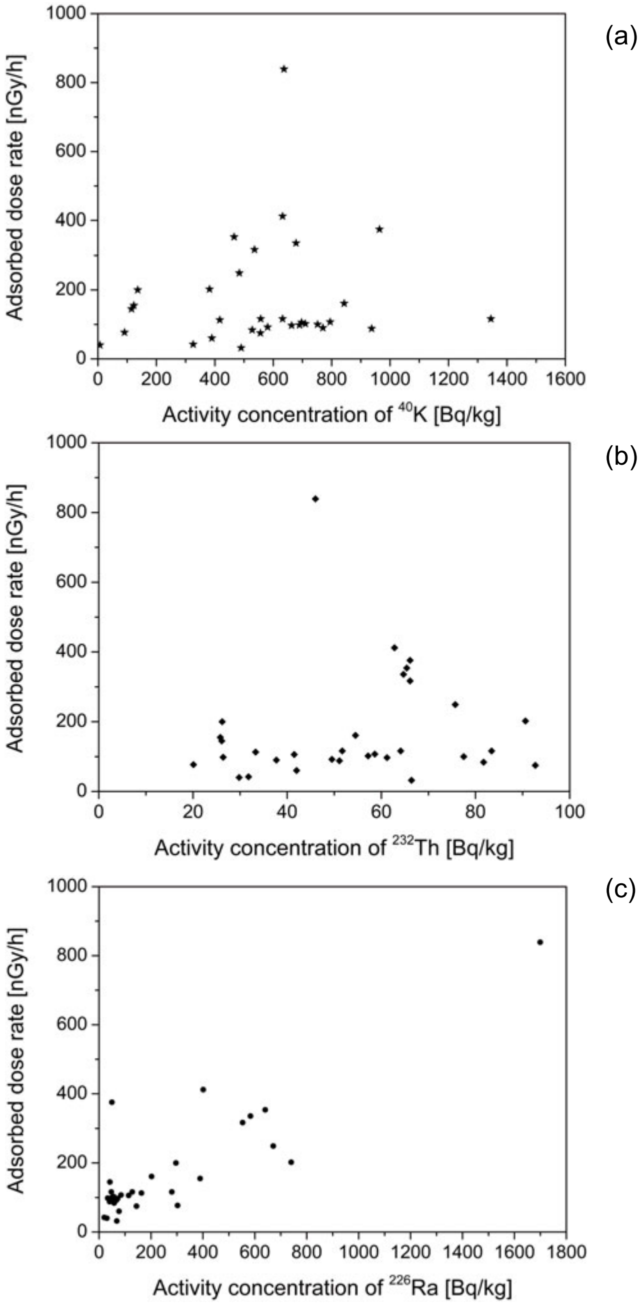


Fig. 3. Relation between estimated gamma absorbed dose rates and potassium (3), thorium (3) and uranium (3) measured by gamma spectrometer coupled with HPGe detector.

where K, Ra, and Th are specific activities of ^{40}K, ^{226}Ra, and ^{232}Th of the sample, expressed in Bq/kg.

The calculated gamma absorbed dose rate varied from 22 to 896 nGy/h with 173 and 103 nGy/h of average and median, respectively (Table 2). The relations between estimated gamma absorbed dose rate and the activity concentrations of ^{40}K, ^{226}Ra, and ^{232}Th are presented in Fig. 3.

Comparing Fig. 2 with adequate Fig. 3 one can see that: (i) the curves of the dependences of the gamma absorbed dose rates on the activity concentration of potassium, uranium, and thorium measured in the field are similar to those made on the basis of the data obtained from laboratory gamma-ray spectrometry measurements; (ii) the ranges of both gamma absorbed dose rates measured at the field and that calculated are comparable; (iii) the natural radioactive anomaly principally originates from the uranium occurrence in the ore and rock formations.

4.3 Chemical composition and uranium and radium isotopes in water

Analyzing the chemical data from Table 3, the investigated waters can be classified into several groups: (i) the waste zone water; (ii) mine water; (iii) water from the floatation reservoirs and private pool; (iv) river, and (v) tap water. The waste zone water (SQ-1) is acid (pH ~ 3.4) and rich in Fe^{2+} (192.6 mg/L), Cu^{2+} (9.4 mg/L), SO_4^{2-} (1.85 g/L), and ^{238}U (12.7 Bq/L). This can be explained by long stagnation time in the contact with solid waste and atmospheric air. This condition enables to leach the iron, copper, sulphur and uranium from the pyrite, chalcopyrite and uraninite into the water. The mechanism by which pyrite is oxidized in water can be presented as follow:

$$FeS_2 + 3.5\,O_2 + H_2O = Fe^{2+} + 2\,SO_4^{2-} + 2H^+. \tag{2}$$

The mine waters (SQ-2, SQ-3, SQ-4) are characterized by the high concentrations of radium, uranium and SO_4^{2-}. The water migrating along the fissures from the hill into the mine leaches the elements from the rocks. The pool and floatation reservoir waters (SQ-6, SQ-7, SQ-8) are acid (pH = 3.6), and the HCO_3^- concentration is very low; therefore, the waters can be regarded as the industrial water. The Red River water (SQ-9, SQ-10) is characterized by low mineralization, but it is contaminated with uranium. The significant difference in the chemical composition of tap water in comparison to the water from the Red River is the proof of its origin not from the Red River at the studied region.

4.4 Radon isotopes

Table 4 summarizes the concentrations of ^{222}Rn and ^{220}Rn in the air inside and outside of the dwellings. The minimum, maximum, average and median

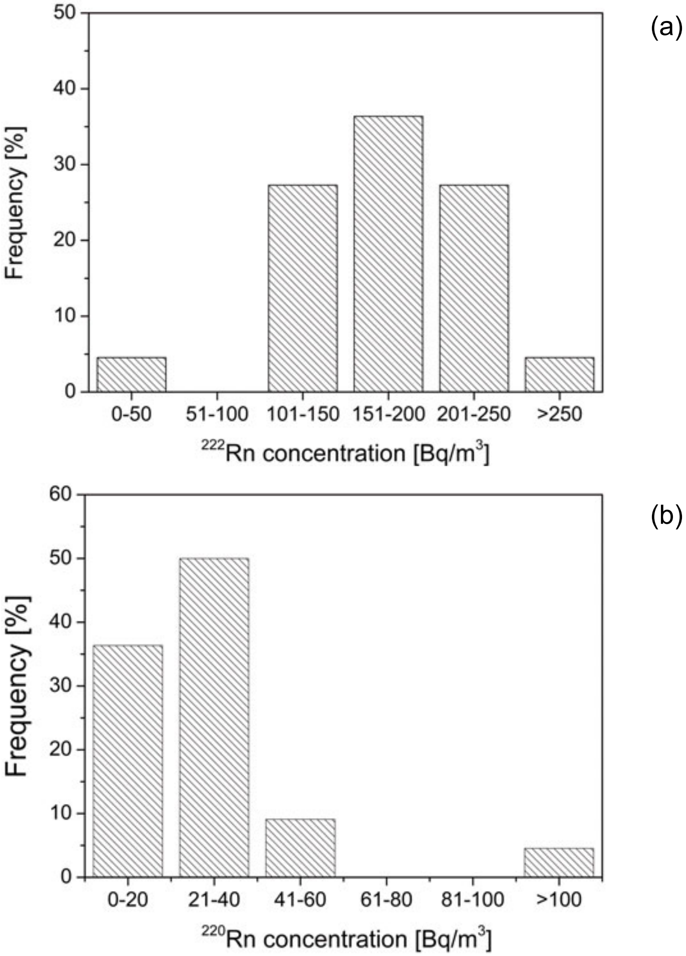


Fig. 4. Distribution of the concentrations of ^{222}Rn (a) and ^{220}Rn (b) measured in dwelling air.

values amounted to 42, 278, 180, and 192 Bq/m^3 for ^{222}Rn and 8, 193, 34, 22 Bq/m^3 for ^{220}Rn. The concentrations of ^{222}Rn and ^{220}Rn in the air outside of the houses are similar to those inside, because the houses are often open due to tropical climate in Vietnam. The average concentration of ^{222}Rn at the Sin Quyen region is near 5 times higher than the world average (37 Bq/m^3). The contribution of the houses where the ^{222}Rn concentration was higher than 150 Bq/m^3 or ^{220}Rn lower than 40 Bq/m^3 amounted to above 70% (Fig. 4). The annual committed dose rate originated from the inhalation of ^{222}Rn and ^{220}Rn and their progeny inside and outside of the dwellings were calculated using the following Eqs. 3 and 4, respectively (UNSCEAR 2000):

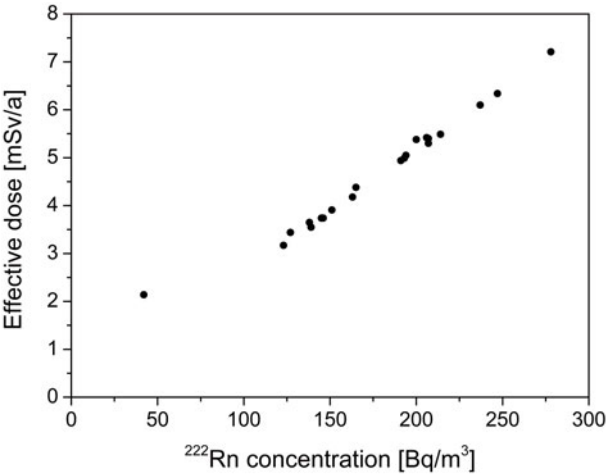


Fig. 5. Relation between estimated annual committed doses and ^{222}Rn concentration in dwelling air.

$$D\,[\mathrm{mSv/a}] = 7000 \times (0.4\times\mathrm{Rn} + 0.02\times\mathrm{Tn})/1000\ , \tag{3}$$

$$D\,[\mathrm{mSv/a}] = 1800 \times (0.4\times\mathrm{Rn} + 0.02\times\mathrm{Tn})/1000\ , \tag{4}$$

where Rn and Tn are the concentrations expressed in Bq/m^3 of ^{222}Rn and ^{220}Rn, respectively.

The estimated annual effective dose from radon and its progeny inhalation inside the dwellings at the Sin Quyen region ranged from 1.1 to 8.1 mSv/a with 4.45 mSv/a of average; this value is three times higher than the world average (1.5 mSv/a). The ^{222}Rn contribution in the committed dose rate amounted to above 95%, and there is a linear correlation between the mentioned parameters (Fig. 5).

5. CONCLUSION

The Sin Quyen IOCG deposit is an elevated natural radiation area; the levels of all investigated radionuclides in the rocks, water and dwelling air are significantly higher than the adequate world averages. The average uranium concentrations measured by field survey and laboratory gamma spectrometers amount to 20.2 ppm and 225 Bq/kg, respectively, being nearly six times higher than the average concentration in the Earth crust (3.5 ppm, 40 Bq/kg). The concentration of thorium obtained by field and laboratory measurements amounts to 17.9 ppm and 54 Bq/kg, respectively; these values are one and haft time higher than the Earth average (10 ppm, 40 Bq/kg). The field measured and calculated gamma absorbed dose rates amount to 196 and

173 nGy/h, while the dose rate from the natural radionuclides in the Earth crust and cosmic ration ranges from 50 to 59 nGy/h (UNSCEAR 2000).

The data of uranium in water samples show that all the surface waters are clearly contaminated with uranium isotopes; the phenomena resulted from the transformation of uranium ion from 4+ state to 6+ one in the contact with atmospheric air. The contents of radium isotopes in the surface waters are low; this fact can be connected with weak leaching of the radium from the rocks into the meteoric water.

The average radon concentration in the dwelling air of investigated region amounts to 169 Bq/m^3, being three times higher than the average world one. The annual effective dose rate from radon and its progeny inhalation amounts to 4.45 mSv of the average. The total annual dose rate composed of the gamma radiation and radon in dwelling air is equal to 4.90 mSv.

There is a linear dependence between the gamma absorbed dose rate and uranium contents for the both field and laboratory gamma spectrometric measurements. The committed dose rate is also linearly depended on the ^{222}Rn concentration in the dwelling air. These facts confirm that the radiation anomaly at the Sin Quyen IOCG deposit is mainly generated by high uranium contents in the rocks and ores of this region.

Acknowledgment. The authors are grateful to UST-AGH Kraków for financial support, grants Nos. 11.11.140.320, 11.11.220.01, and University of Mining and Geology (UMG), Hanoi, Vietnam, grant No. 01/2012/ HD-HTQTSP.

References

Baker, T. (2005), IOCG deposits – The exploration and Research Challenges, available from: http://www.minerals.pir.sa.gov.au.

Bui, P.M., V.H. Nguyen, V.K. Phan, and D.T. Tran (2004), Geology and mineral resources of Kim Binh Lao Cai sheet. Scale 1:200 000, Dept. Geol. Mineral., Hanoi, Vietnam.

Corriveau, L., L. Ootes, H. Mumin, V. Jackson, V. Bennet, J.F. Cremer, B. Rivard, and I. McMartin (2007), Alteration vectoring to IOCG(U) deposits in frontier volcano-plutonic terrains, Canada. **In:** B. Milkereit (ed.), *Proc. Exploration '07: Fifth Decennial Int. Conf. on Mineral Exploration*, 1171-1177.

Hunt, J., T. Baker, and D. Thorkelson (2005), Regional-scale Proterozoic IOCG-mineralized breccia systems: examples from the Wernecke Mountains, Yukon, Canada, *Miner. Deposita* **40**, 492-514.

Ishihara, S., H. Hirano, M. Hoshino, N.C. Pham, T.D. Pham, and A.T. Tran (2011), Mineralogical and chemical characteristics of the allanite-rich copper and iron ore from the Sin Quyen mine, northern Vietnam, *Bull. Geol. Surv. Jap.* **62**, 5/6, 197-209.

Jodlowski, P. (2006), Self-absorption correction in gamma-ray spectrometry of environmental samples – an overview of methods and correction values obtained for the selected geometries, *Nukleonika* **51**, *Supl*. 2, S21-S25.

Jodlowski, P., and S. Kalita (2010), Gamma-ray spectrometry laboratory for high-precision measurements of radionuclide concentrations in environmental samples, *Nukleonika* **55**, 2, 143-148.

Lange, R. (1972), *Geochemical Tables*, Edition Leipzig.

Le, K.P., T.S. Nguyen, A. Piestrzynski, J. Pieczonka, P. Jodłowski, D.C. Nguyen, V.L. Nguyen, V.N. Nguyen, and H.H. Do (2015), Dispersion of radioactive materials due to processing, mining activities in Sin Quyen Copper mine, Lao-Cai, North Vietnam. **In:** *Materials of the National Conference of the 70th Anniversary of Independence of Vietnam, 5-6 November 2015, Hanoi, Vietnam*, 253-261 (in Vietnamise).

McLean, R.N. (2001), The Sin Quyen iron oxide-copper-gold-rare earth oxide mineralization of North Vietnam. **In:** T.M. Porter (ed.), *Hydrothermal Iron Oxide Copper-Gold and Related Deposits: A Global Perspective*, Vol. 2, PGC Publishing, Adelaide, 293-301.

Nguyen, D.C., J. Niewodniczański, J. Dorda, A. Ochoński, E. Chruściel, and I. Tomza (1997), Determination of radium isotopes in mine waters through alpha- and beta-activities measured by liquid scintillation spectrometry, *J. Radioanal. Nucl. Chem*. **222**, 1, 1-2, 69-74, DOI: 10.1007/BF02034249.

Piestrzyński, A. (1989), Uranium and thorium in copper ore deposits on the fore-Sudetic monocline (SW Poland), *Miner. Pol*. **20**, 1, 41-53.

Radosys (2013), Radosys User Manual of CR-39, Radosys Kft.

Sillitoe, R.H. (2003), Iron oxide-copper-gold deposits: an Andean view, *Miner. Deposita* **38**, 7, 787-812, DOI: 10.1007/s00126-003-0379-7.

Skwarzec, B. (1997), Radiochemical methods for the determination of polonium, radiolead, uranium and plutonium in environmental samples, *Chem. Anal*. **42**, 1, 107-113.

UNSCEAR (2000), Annex B – exposures from natural radiation sources, UNSCEAR Report 2000, 84-156.

Williams, P.J., M.D. Barton, D.A. Johnson, L. Fontbote, A. De Haller, G. Mark, N.H.S. Oliver, and R. Marschik (2005), Iron oxide copper-gold deposits: geology, space-time distribution, and possible modes of origin, *Economic Geology 100th Anniv. Vol.*, 371-405.

Zhao, X.F., and M.F. Zhou (2011), Fe–Cu deposits in the Kangdian region, SW China: a Proterozoic IOCG (iron-oxide–copper–gold) metallogenic province, *Miner. Deposita* **46**, 7, 731-747, DOI: 10.1007/s00126-011-0342-y.

Received 5 January 2016
Received in revised form 1 March 2016
Accepted 5 May 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2322-2336
DOI: 10.1515/acgeo-2016-0092

Very Low Frequency Electromagnetic Induction Surveys in Hydrogeological Investigations; Case Study from Poland

Szymon ORYŃSKI[1], Marta OKOŃ[1], and Wojciech KLITYŃSKI[2]

[1]Institute of Geophysics, Polish Academy of Sciences, Warszawa, Poland;
e-mails: sorynski@igf.edu.pl (corresponding author), msliwa@igf.edu.pl

[2]AGH University of Science and Technology,
Faculty of Geology, Geophysics and Environment Protection, Kraków, Poland,
e-mails: gpklityn@geol.agh.edu.pl, gpklityn@gmail.com

Abstract

In 2011, a geophysical survey was carried out in the surroundings of the Jagiellonian University in Cracow, using a Very Low Frequency method. The measurements were designed to determine the reason of frequent flooding of the lowest level of the building. The main objective of the study was to find out from where and in which way the rainwater seeps into the building and how this problem can be solved in the least invasive manner. The aim of geophysical methods was also to provide necessary information that will enable the construction of a hydrogeological model of the local environment. The interpretation revealed the presence of a sandy gutter surrounded by impermeable clay. There is a big resistivity contrast between those layers. Their location and approximate dimensions were determined.

Key words: Very Low Frequency method, groundwater research, electromagnetism, shallow geophysics, near surface.

1. INTRODUCTION

The aim of this study is to examine the extent to which the Very Low Frequency method can be useful in mapping shallowly buried geological structures. VLF EM induction method is used to determine the electric properties of the rock in the superficial zone. Then, on this basis, there is a possibility to establish the type of material present in the examined environment. The method is used in two modes: VLF-R (electric) and VLF (magnetic). A qualitative interpretation method was used to interpret VLF EM data to identify anomalous zones. Research has been done using the VLF-3 Scintrex instrument. The receiving frequency has been set at 21.75 kHz (HWU, Table 1). We have analysed a case which theoretically is on the border of applicability of the method. We tried to detect shallowly buried, watered sand trough in a low resistivity environment. The justification of application VLF method in this case is very simple. There is a big resistivity contrast between layers occurring there (Kleczkowski *et al.* 1994). Clays and loams have very low electrical resistivity (Table 2), but sands (even saturated by water) are characterized by higher resistivity (Table 3). The resistivities of different rocks vary within wide limits. The saline water, which saturates rocks, clays and sands is characterized by ionic conduction, and their resistivities are relatively low (Antoniuk *et al.* 2003). The increase of the water salinity causes a significant increase in its conductivity (Plewa and Plewa 1992). The main impact on the resistivity of biphasic rock has the component of lower resistivity. In this case, VLF EM as a geophysical method should provide information about the hydrogeological conditions for technical and construction purposes.

2. VERY LOW FREQUENCY METHOD

The VLF EM method belongs to the electromagnetic methods applied in near-surface geophysics. Its main application is the mapping of vertical discontinuities such as faults, flexures and ore mineralization zones (Saydam 1981). It is used also in hydrogeological studies and detecting deposits of particular rock material (Chandra 2015).

The VLF method uses electromagnetic waves of very low frequency in terms of radio waves. From the standpoint of geoelectrical methods, these frequencies are considered to be quite high. The sources of such waves are strong military radio stations used to maintain communication with submarines. There are 11 major VLF transmitters distributed around the world (Reynolds 2011). They are emitting waves in the frequency range from 15 to 30 kHz. Electromagnetic field is polarized cylindrically around the antenna. Registered electromagnetic waves can be divided into:

- Ionospheric waves, which once or multiply bounce off the ionosphere and the surface of the earth. The most common are singly (first sky hop) and doubly (second sky hop) reflected waves;
- Spatial waves, which propagate in the air, directly from the transmitter to the receiver;
- Ground waves, which propagate in the soil. These are recorded only at short distance from the source.

Source signal is described by cylindrical coordinates, but it is better to transform it into Cartesian system. The primary field contains a vertical electric field component (E_z) and a horizontal (tangent) magnetic field component (H_y), both perpendicular to the direction of propagation, x, and a horizontal (radial) electric field component (E_x). Electromagnetic field properties usually are discussed for two areas: the near zone (induction) and far zone (radiation). The far zone refers to the regions in which the distance between transmitter and receiver is many times bigger than the wavelength. The VLF measurements are carried out at a great distance (far zone). Conductors in the earth modify the H_y component and create the vertical component H_z. The vertical magnetic component is measured as ratio H_z/H_y. The primary H_y penetrates into the ground and induces eddy currents forming a secondary horizontal electric component in buried conductive structures. At a great distance, the horizontal electric component is much less than the vertical electric component (Tabbagh *et al.* 1991). A secondary magnetic field is generated which is out of phase with the primary magnetic field. The intensity of the secondary magnetic field depends on the conductivity of the ground (Khalil and Santos 2010). Resultant magnetic field, which is produced by the interference between the primary and the secondary magnetic fields, is elliptically polarized. The parameters of interest are: (a) the orientation of the minor ellipse axis (tilt angle, α), also called the real (in-phase); and (b) the ratio of the minor to the major ellipse axes of the polarization (ellipticity, e), also called the imaginary (quadrature) component (Karous and Hjelt 1977).

As a result of low attenuation and high diffraction, at a small distance from the source, very long waves propagate mainly as spatial waves. However, at a distance of 1000-2000 km from the transmitter, ionospheric wave field strength exceeds the field strength of spatial waves. These bounce off the lowest layers of the ionosphere, which have very low electron density. Propagation range of these waves is very large and extends over several thousand kilometres (McNeill and Labson 1992).

The most powerful transmitters emitting frequencies of interest to us are summarized in Fig. 1. The ellipses mark the ranges of spatial waves. Ionospheric waves are recorded worldwide. Military stations usually do not

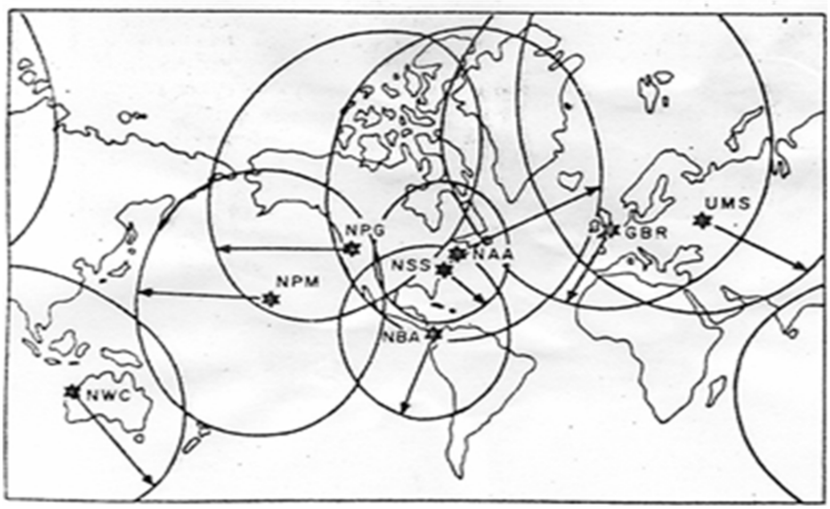


Fig. 1. Location of stations; ellipses mark the maximum range of spatial waves, source: Scintrex (1988).

provide a continuous signal. Station designated as HWU transmits practically without significant breaks. It is the Rosnay station in France, which emits electromagnetic waves with a frequency of 21.75 kHz (HWU France). A list of the transmitters whose signals are received in Poland is presented in Table 1 (Graf *et al.* 2011).

Table 1

List of VLF transmitters whose signals are received in Poland (source: http://www.classaxe.com/dx/ndb/reu/)

ID	Frequency [kHz]	Location
VTX	16.3; 17.0	India
JXN	16.4	Norway
SAQ	17.2	Sweden
RDL	18.1; 21.1	Russia
HWU	18.3; 21.75; 22.6	France
NWC	19.8	Australia
ICV	20.27	Italy
NPM	21.4	USA
GQD	22.1	United Kingdom
DHO38	23.4	Germany
NAA	24.0	USA
NML4	25.4	USA
TBB	26.7	Turkey

3. METHODOLOGY AND DATA ACQUISITION

The method is based on measurement of the secondary fields produced by the underground conductive bodies, which are subject to action of the primary electromagnetic field. Although the source of the waves is artificial, we classify the VLF method as passive and operating in a far field of the source. The reason is that the signal transmission is done completely independently of its registration (serves other purposes), and takes place at a distance of at least several hundred kilometres from the research area. We can therefore consider an electromagnetic wave as a planar rather than spherical wave. In the VLF method, measurements are carried out mostly along profiles parallel to each other. Theoretically, designing profiles perpendicular to the investigated structure should lead to the best results. However, the course of the structures seldom is well understood, which requires previous reconnaissance measurements (Kaufman and Keller 1981).

The sensors do measurement the total field consisting of the primary and secondary field. Like the primary field, secondary field contains also two horizontal components, vibrating perpendicular to the direction of wave propagation. Also, the vertical component is measured. As a result of measurements, we obtain three components: E_x, H_y, and H_z. Horizontal components E_x and H_y are derived from primary and secondary fields, while the component H_z comes from the secondary field only. Due to measured components, the measurement can be divided into magnetic mode (H_z and H_y components are measured) and electric mode (E_x and H_y). In the VLF method, the apparent resisitivity ($\rho_a\,[\Omega\mathrm{m}]$) is obtained by calculating the impedance ($Z = E_x/H_y$) from measured components in electric mode, according to the following formula similar to magnetotelluric measurement (Zdhanov and Keller 1994):

$$\rho_a = \frac{1}{\omega\mu_0} \cdot \left| \frac{E_x}{H_y} \right|^2 , \tag{1}$$

where $\rho_a\,[\Omega\mathrm{m}]$ is the apparent resistivity; ω is the angular frequency, which is equal to $2\pi f$, where f is a frequency of an electromagnetic wave in Hertz; μ_0 is the magnetic permeability of a vacuum, which is equal to $4\pi\cdot 10^{-7}$ [Vs/Am]; and E_y and H_x are horizontal components of the electric and magnetic field, respectively, expressed in millivolts per meter [mV/m] and in microamperes per meter [μA/m].

The unit of apparent resistivity is the same as for the true resistivity [Ωm], but only for a homogeneous medium these parameters are equal in value. In any other case, we obtain a function reflecting the distribution of resistivity in the ground only approximately (Tabbagh *et al.* 1991). At the beginning, the analysis of apparent resistivity along individual profiles is done. Interpolated data from several profiles allow obtaining an estimated apparent resistivity distribution in the study area. The second parameter obtained in the electric mode is the phase shift of impedance, expressed in degrees (Zhdanov and Keller 1994).

$$\varphi = \tan^{-1}\left[\frac{Im\dfrac{E_x}{H_y}}{Re\dfrac{E_x}{H_y}}\right] . \tag{2}$$

For a homogeneous half-space, the phase shift of impedance should be equal to –45 degrees or 45 degrees. The sign depends on the approach used for the solution of Helmholtz equation. The instrument which was used in this measurement calculates it with the negative sign. This parameter behaves inversely to the resistivity changes. The phase is maximum, where the apparent resistivity has its maximal slope. It increases when the resistivity of the soil decreases. Primary field coupled with conductors will be substantially shifted in phase. In the magnetic mode, searching for a buried conductor is performed by analyzing the axis of the elongated conductive structure. In a typical application of the VLF method, the conductive structure would reveal as a change in a sign of the vertical component. This characteristic sign reversal is generally analyzed separately for each profile. However, the placement of the component H_z on the research area map enables to specify the zone of occurrence of anomalous structures approximately. When the anomalous structure is present, an extreme of the curve for the horizontal component (H_y) can be expected. Above the high-conductivity layer the maximum is received. A low-conductivity layer causes a minimum (Karous and Hjelt 1983).

In the ground, the magnetic field is disturbed in phase and direction. It is attenuated, and the depth range is determined by skin effect. The depth of electromagnetic field penetration is the depth at which the wave amplitude decreases by one neper, in relation to the field at the surface of the earth. Depth range can be estimated using the following formula (Spies 1989):

$$\delta = 503.3 \cdot \sqrt{\rho / f} . \tag{3}$$

As is clear from the formula, the depth range depends only on the average resistivity and frequency. Considering that it is based here on a fixed frequency wave, the depth range depends only on the average resistivity in the soil.

4. DATA ANALYSIS AND INTERPRETATION

Research area was located near the Faculty of Management and Social Communication of the Jagiellonian University in Cracow (Lesser Poland, Poland). In the picture below (Fig. 2) there are two maps, which present the location of the survey area. Minor map contains location of Cracow city in Poland.

The major map illustrates the research area with greater accuracy. VLF method was used to detect the sand gutters which cause the uncontrolled water flow and flooding of the lowest level of the building. Eight parallel profiles were made in lengths from 43 to 56 m, spaced to each other by 1.5 m. The distance between the measuring points along the profile was 1 m. Rosnay station in France with a transmitted wave frequency of 21.75 kHz served as a source of electromagnetic waves (Bernard and Valla 1991).

Survey area is located on the right bank of Vistula river within Miocene tectonic graben. The graben is filled by clay deposits mainly containing small inserts of evaporates. In the area of interest, the deposits originate from the middle Miocene–Tortonian (Gradziński 1972, Pitera 2004a, b). Tortonian sediments were covered with a thin layer of Holocene alluvial sediments and locally with the Pleistocene weathered deposits. In the area of our survey, shallowly beneath the surface, impervious clay material is divided with small sand structures that let the water through (Pociask-Karteczka 1994). The clay of marine origin, occurring in the area of study, is characterized by a very

Fig. 2. Location of the survey area. The red polygon stands for study area. Source: Google Earth.

Table 2

Values of resistivity (ρ) and conductivity (σ) of rocks (Kobranova 1989)

Rock	ρ [Wm]	σ [mS/m]
Sand	100-2500	0.4-10
Clays	1-100	10-1000
Loam	5-50	20-200
Marls	3-70	14-300
Sandstone	500-5000	2-20
Limestone	$1\text{-}2\times10^5$	0.5-0.01
Gypsum	$10^5\text{-}10^6$	0.01-0.001
Anhydrite	$10^2\text{-}10^5$	0.001-10
Rock salt	$10^4\text{-}10^8$	$10^{-5}\text{-}0.1$

Table 3

Values of resistivity (ρ) and conductivity (σ) of water (Keller 1966)

Material	ρ [Wm]	σ [mS/m]
Natural waters	1-100	10-1000
Saline waters, 3%	0.15 (average)	6700
Saline waters, 20%	0.05 (average)	20000

Table 4

Values of resistivity (ρ) and conductivity (σ) of sandstones (Guinea *et al.* 2010)

Material	ρ [Wm]	σ [mS/m]
Solid sandstone	> 1000	< 1
Saturated sandstone	< 100	< 10

low resistivity (Table 2). In general, the sands have a high resistivity, but rainwater (Table 3) may reduce its value, even to just several ohmmeters (Table 4). So, a slight resistivity contrast should be visible and it would be possible to separate the layers of sandstone from clay.

The depth range of such layers may vary from about 3 to 10 m. The data has been filtered by a 2×2 window and edited due to disturbances and submitted in the map.

The first analyzed parameter is the distribution of the real magnetic vertical component (H_z in phase) as ratio H_z/H_y (Fig. 3). In the behavior of the vertical component the change of sign is of the greatest interest. In the figure a contour line for the value equal to zero H_z/H_y is marked with a red line. Most of the profiles, between 20 and 30 m, reveal a change of sign in H_z/H_y

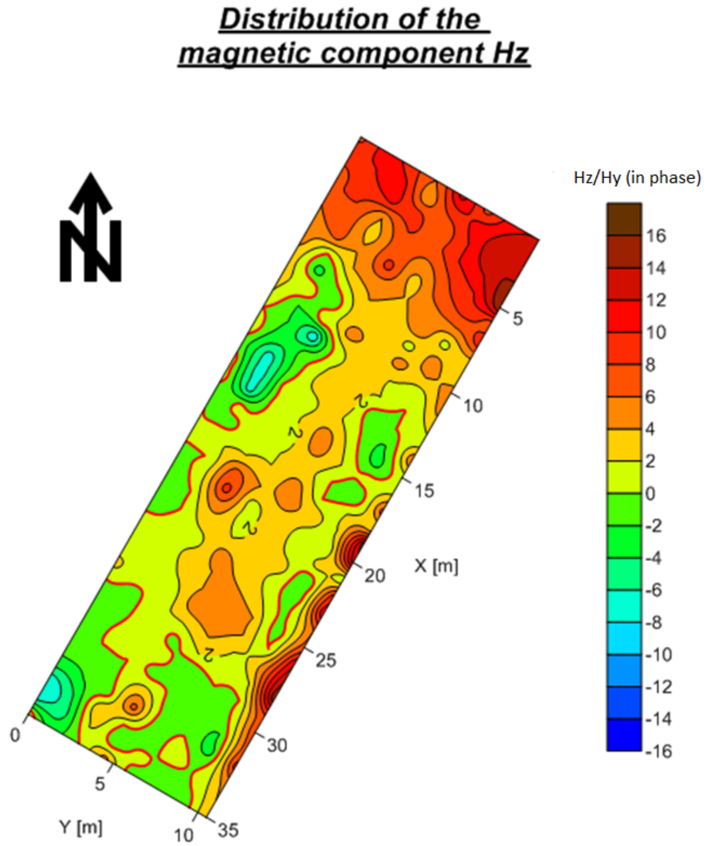


Fig. 3. The distribution of the vertical magnetic field, ratio H_z/H_y (phase).

value. This zone corresponds to the location of the flooded building. In the northern part of the study area, a very strong positive anomaly shows up in the value of H_z/H_y in phase. It coincides with a strong minimum of apparent resistivity in this region (Fig. 6).

The distribution of the values of quadrature component H_z(quad)/H_y is presented in Fig. 4. We can clearly observe an area in the northern part of the map. This zone is characterized by extremely high positive values of the vertical component, while for the rest of the area the values are negative, reaching their minimum in the southern part of the study area. Extremely low values of the magnetic component in the southern part of the study area may be related to rain-water saturated sands.

For a better illustration of results, they distribution of parameter H_z/H_y in phase for Profile IV is additionally shown (Fig. 5). This profile is marked by a red line in Fig. 4. The zero crossing around 27 m of the profile is clearly

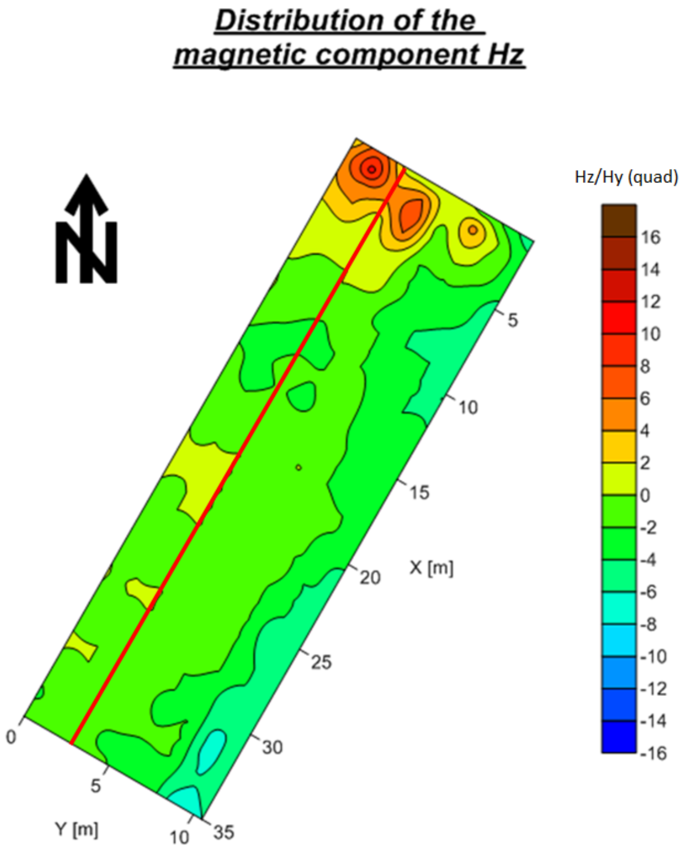


Fig. 4. The distribution of the vertical magnetic field, ratio H_z/H_y (quadrature). The redline marks Profile IV.

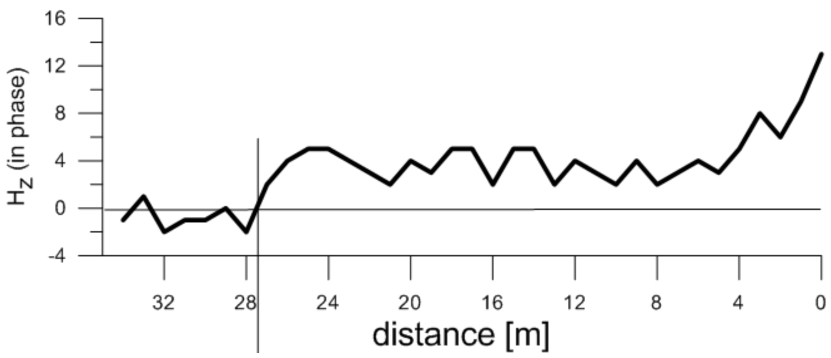


Fig. 5. The distribution of H_z in phase component on Profile IV.

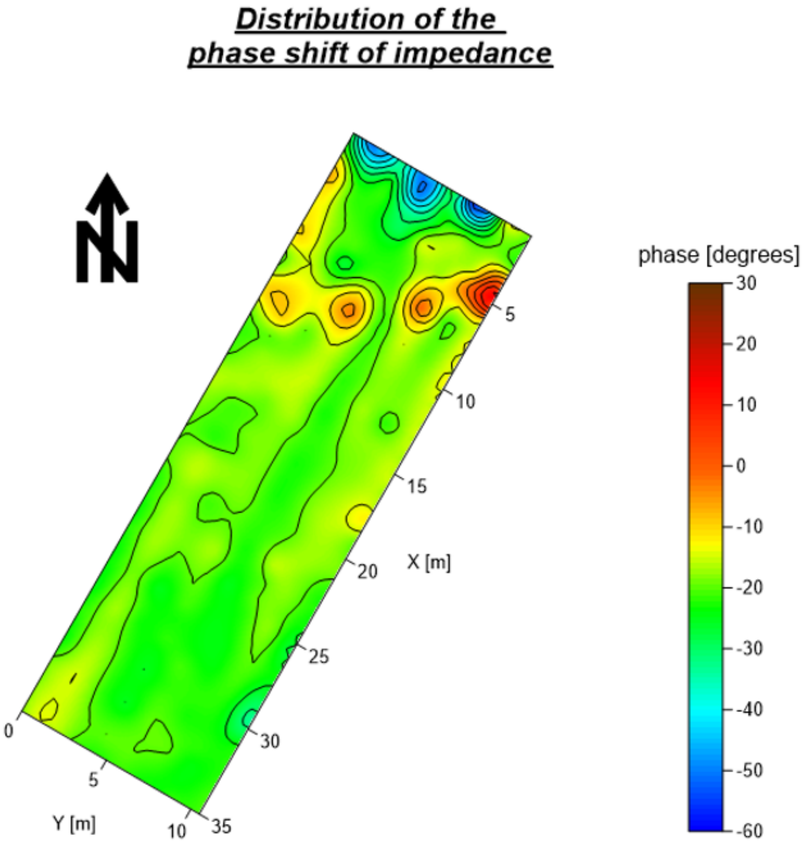


Fig. 6. The distribution of the phase shift of impedance.

visible. This correlates with the location of the gutter. It also coincides with the positive anomaly of the apparent resistivity distribution (Fig. 6).

The next step is to analyze electric parameters as phase shift of impedance and apparent resistivity. Firstly, the distribution of phase shift of impedance is presented (Fig. 6). The instrument used in this measurement calculates this parameter as –45 degrees for homogenous half-space, so it is presented in that way. For the majority of research area the phase shift oscillates about –30 degrees, but in the northern part it grows up to +30 degrees. It is related to accumulation of low resistive clays in there. Phase shift gives better results for good conductors in resistive surroundings, than otherwise. Due to that fact, sandy gutter is not visible in this map.

The last analyzed parameter, which is the most reliable, is the distribution of apparent resistivity (Fig. 7). It is well-known that the vertical component of magnetic field (H_z) well appears in horizontal conductivity

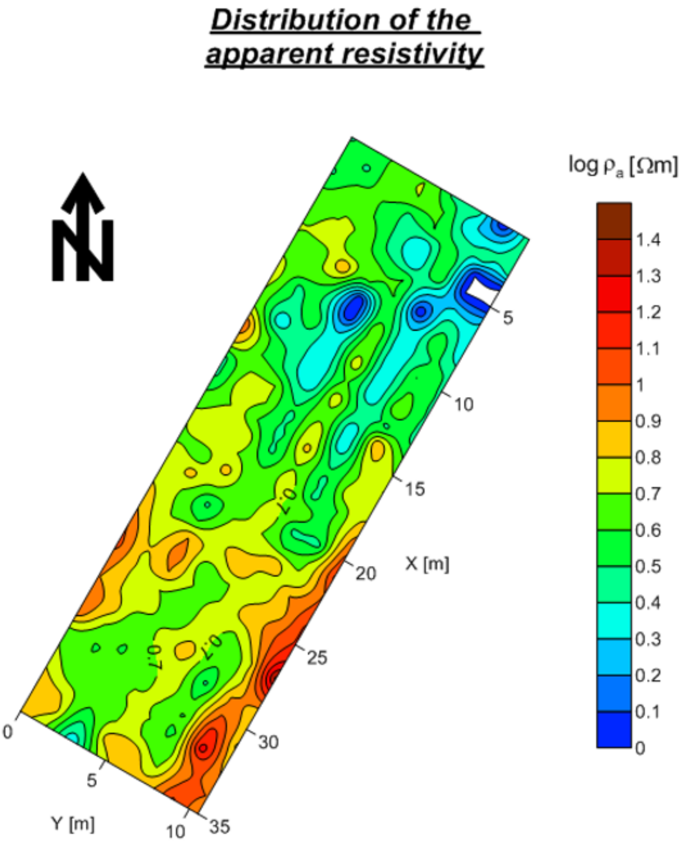


Fig. 7. The distribution of apparent resistivity in a logarithmic scale.

discontinuity, whose strike is not parallel to the direction of propagation of the EM field. In such a situation H_z crosses zero value over conductive body (Saydam 1981). In the discussed area, the sand gutter is low-conductive but the other rock, that is, clay is a good conductor. Clays form a lot of conducting bodies, which probably disturbs the magnetic field (H_z). In this discontinuity (sand-clay), the electric field must be of interest. The electric field (E_x) reacts for local horizontal changes of resistivity because it describes currents flowing across the discontinuity. So the apparent resistivity is probably the most reliable parameter in the area.

The map was made for the logarithm of apparent resistivity due to the high variability of this parameter. The zone of reduced resistivity is clearly visible. It coincides very well with the anomalies of magnetic component H_z/H_y which were described earlier. In the northern part of the map, a region of extremely low resistivity occurs. The reason for that is the presence of the

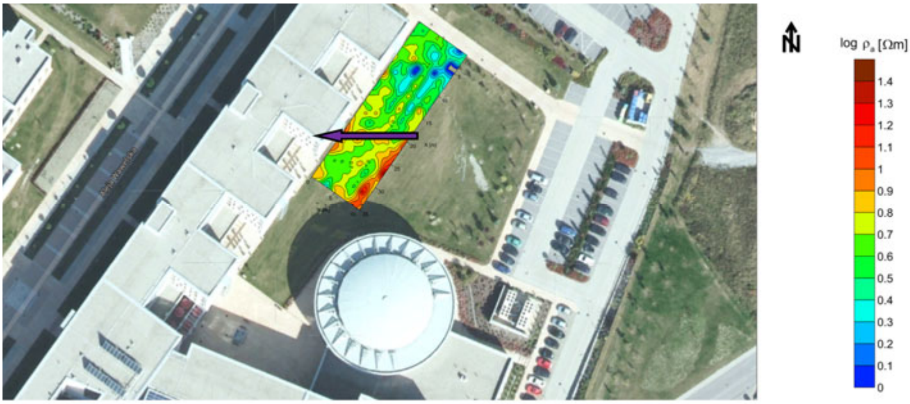


Fig. 8. Distribution of apparent resistivity in a logarithmic scale superimposed on a map of the study area. Purple arrow shows the direction of flooding. Source: www.googlemaps.com.

very low-resistivity clays, which have been described previously. In the central part of the study area, between 20 and 25 m of profile, there is a zone of increased apparent resistivity. It may indicate the presence of a sandy trough, which is responsible for the flooding of the building.

The last figure (Fig. 8) shows the distribution of apparent resistivity superimposed on a map of the study area.With such a result, we can conclude that just in this area a watered sandy trough is located, with an east-west course, responsible for the flooding of the lowest floor of the building.

5. DISCUSSION

These VLF surveys were carried out in order to define the location of a sand gutter. As we can see from the obtained data, our subject has been detected, although with low precision. All analysed parameters revealed an anomaly in the middle part of the study area. It coincides with the expected location of the gutter. The results of the study were affected by an adverse impact of the close proximity to the buildings. It could greatly weaken or distort the signal. This is undoubtedly a disadvantage of VLF method. In these circumstances, the Electrical Resistivity Tomography (ERT) could be very helpful as a complementary method. Both methods seem to be suitable for this particular case from a theoretical point of view. However, taking into account close proximity to the buildings, VLF method is more convenient because it does not require such spacing as the ERT method.

6. CONCLUSIONS

The study shows that although the method is dedicated to the detection of conductors in low-conductive soils, it is also effective for shallow issues detection in low resistivity layers. Small resistivity contrasts between low resistivity materials are very well recorded. Very important matter in this method is to determine the expected depth range. The method is based on a wave with a fixed frequency, so the depth of penetration depends only on the distribution of resistivity in the study area. The higher it is, the greater the range of penetration. The theoretical maximal depth range of the method in this case was calculated as 10 m. For this research the studied structure is located at a depth about 5 m below the ground. We can conclude that registered anomaly comes from the searched structure. In this case, the theoretical depth of penetration is from 3 to 10 m. Therefore, we can conclude that the recorded anomaly derived from the studied structure. VLF method appears to be sensitive to changes in resistivity, even in low-resistivity rocks. It can be successfully used for the determination of water flow zones. This method is suitable for solving problems related to the protection of buildings against flooding.

References

Antoniuk, J., W.J. Mościcki, and K. Janicki (2003), Geoelectric investigations of migration of chemically-polluted waters from the post-flotation settlement reservoir "Żelazny Most", IGSMiE PAN, Kraków, 383-391.

Bernard, J., and P. Valla (1991), Groundwater exploration in fissured media with electrical and VLF method, *Geoexploration* **27**, 1-2, 81-91, DOI: 10.1016/0016-7142(91)90016-6.

Chandra, P.C. (2015), *Groundwater Geophysics in Hard Rock*, CRC Press, 384 pp.

Gradziński, R. (1972), Przewodnik geologiczny po okolicach Krakowa, Wydawnictwa Geologiczne, Warszawa (in Polish).

Graf, K.L., U.S. Inan, and M. Spasojevic (2011), Transmitter-induced modulation of subionospheric VLF signals: Ionospheric heating rather than electron precipitation, *J. Geophys. Res.* **116**, A12, A12313, DOI: 10.1029/2011JA016996.

Guinea, A., E. Playa, L. Riverso, M. Nimi, and R. Bosch (2010), Geoelectrical classification of Gypsum Rocks, *Surv. Geophys.* **31**, 6, 557-580, DOI: 10.1007/s10712-010-9107-x.

Karous, M., and S.E. Hjelt (1977), Determination of apparent current density from VLF measurements. Contribution N. 89, Department of Geophysics, University of Oulu. Finland.

Karous, M., and S.E. Hjelt (1983), Linear filtering of VLF dip-angle measurements, *Geophys. Prospect* **31**, 5, 782-794, DOI: 10.1111/j.1365-2478.1983.tb01085.x.

Kaufman, A.A., and G.V. Keller (1981), *The Magnetotelluric Sounding Method*, Elsevier, Amsterdam.

Keller, G.V. (1966), Electrical properties of rocks and minerale. **In:** S.P. Clark (ed.), *Handbook of Physical Constants*, The Geological Society of America, 587 pp.

Khalil, M.A., and F.M. Santos (2010), Comparative study between filtering and inversion of VLF-EM profile data, *Arab. J. Geosci.* **4**, 1-2, 309-317, DOI: 10.1007/s12517-010-0168-4.

Kleczkowski, A.S., T. Solecki, J. Myszka, and J. Stopa (1994), Krakowskie artezyjskie zdroje wód pitnych z wapieni jury, WGGiOŚ AGH Kraków, 61 pp. (in Polish).

Kobranova, V.N. (1989), *Petrophysics*, Spinger Verlag, Berlin, 375 pp.

McNeill, J.D., and V.F. Labson (1992), Geological mapping using VLF radio fields. **In:** M. Nabighian (ed.), *Electromagnetic Methods in Applied Geophysics*, Vol. 2, Society of Exploration Geophysicists, Tulsa.

Pitera, H. (2004a), Gipsy okolic Krakowa, *Aura* **7**, 11-13 (in Polish).

Pitera, H. (2004b), Gipsy z osiedla Kliny-Zacisze w Krakowie, *Wszechświat* **105**, 7-9, 201-203 (in Polish).

Plewa, M., and S. Plewa (1992), *Pertofizyka*, Wyd. Geol., Kraków (in Polish).

Pociask-Karteczka, J. (1994), Changes of water conditions at the Cracow area, *Zesz. Nauk. UJ, MCXLIV Pr. Geogr.* **96**, 38 pp. (in Polish).

Reynolds, J.M. (2011), *An Introduction to Applied and Environmental Geophysics*, 2nd ed., John Wiley and Sons, Chichester, 796 pp.

Saydam, A.S. (1981), Very low-frequency electromagnetic interpretation using tilt angle and ellipticity measurements, *Geophysics* **46**, 11, 1594-1605, DOI: 10.1190/1.1441166.

Scintrex (1988), User manual Scintrex VLF-3, Geophysical and geochemical instrumentation and services, Scintrex.

Spies, B.R. (1989), Depth of investigation in electromagnetic sounding methods, *Geophysics* **54**, 7, 872-888, DOI: 10.1190/1.1442716.

Tabbagh, A., Y. Bendritter, P. Andrieux, J.P. Decriaud, and R. Guerin (1991), VLF resistivity mapping and verticalization of the electric field, *Geophys. Prospect.* **39**, 8, 1083-1097, DOI: 10.1111/j.1365-2478.1991.tb00360.x.

Zhdanov, M.S., and G.V. Keller (1994), *The Geoelectrical Methods in Geophysical Exploration*, Elsevier Science.

Received 5 October 2015
Received in revised form 13 February 2016
Accepted 5 May 2016

DE GRUYTER OPEN

Acta Geophysica

vol. 64, no. 6, Dec. 2016, pp. 2337-2355
DOI: 10.1515/acgeo-2016-0108

An Estimation Method of Pore Structure and Mineral Moduli Based on Kuster–Toksöz (KT) Model and Biot's Coefficient

Da PENG, Cheng YIN, Hu ZHAO, and Wei LIU

Southwest Petroleum University, Sichuan Province Key Laboratory of Natural Gas Geology, School of Geoscience and Technology,Chengdu, China;
e-mail: pengda2012@163.com

Abstract

Pore structure and mineral matrix elastic moduli are indispensable in rock physics models. We propose an estimation method of pore structure and mineral moduli based on Kuster-Toksöz model and Biot's coefficient. In this technique, pore aspect ratios of five different scales from 100 to 10^{-4} are considered, Biot's coefficient is used to determine bounds of mineral moduli, and an estimation procedure combined with simulated annealing (SA) algorithm to handle real logs or laboratory measurements is developed. The proposed method is applied to parameter estimations on 28 sandstone samples, the properties of which have been measured in lab. The water saturated data are used for estimating pore structure and mineral moduli, and the oil saturated data are used for testing these estimated parameters through fluid substitution in Kuster–Toksöz model. We then compare fluid substitution results with lab measurements and find that relative errors of P-wave and S-wave velocities are all less than 5%, which indicates that the estimation results are accurate.

Key words: rock physics, Kuster–Toksöz model, mineral matrix moduli, pore structure, pore aspect ratios spectrum.

1. INTRODUCTION

Reservoir rocks consist of a variety of components, such as minerals, clays, cements, pores, pore fluids, *etc.* Rock physics models can be used to establish relationships of these physical parameters and dynamic elastic moduli of rocks. Consequently, they can bridge seismic attributes with reservoir parameters. The Gassmann model (Gassmann 1951) is a very useful tool to build these relationships by means of fluid substitution. It considers the effects of mineral components, pore fluids, porosity, and dry rock elastic moduli, and seismic velocities. The Gassmann model assumes well connected pores with no isolated pores. The inclusion model (Ament 1953, Eshelby 1957, Walsh 1965, Wu 1966, Mori and Tanaka 1973, O'Connell and Budiansky 1974, Berryman 1980, Zimmerman 1984) can be used to construct relationships between pore microstructure parameters and dynamic elastic moduli of rocks. The Kuster–Toksöz (KT) model (Kuster and Toksöz 1974), differential effective medium (DEM) model (Cleary *et al.* 1980, Norris *et al.* 1985, Zimmerman 1991), and self-consistent (SC) model (Budiansky 1965, Hill 1965) are often used to analyze effects of pore shapes and percentages of pore aspect ratios on dynamic elastic moduli of rocks and seismic velocities. Rock physics models have many different parameters, such as lithology, clay content, porosity, fluid saturation, mineral matrix moduli, pore aspect ratio spectrum, *etc.* If these parameters are unknown or given incorrectly, rock physics models cannot implement fluid substitution and predict velocities. Therefore, many researchers have investigated estimation methods to estimate these rock physics parameters.

Tran proposed a modified DEM velocity estimation scheme to estimate pore aspect ratio spectra on a suite of Berea sandstones (Tran 2008). Lev Vernik developed empirical relations for estimating pore-shape factors based on the noninteraction approximation with the Mori-Tanaka model (Vernik and Kacanov 2010). Jensen gave a procedure for estimating the mineral elastic moduli of smectite and kaolinite by using Hashin-Shtrikman upper and lower bounds (Jensen *et al.* 2011). Lin *et al.* developed an estimation method for self-adapting mineral matrix bulk moduli based on Gassmann theory (Lin *et al.* 2011). Spikes applied the self-consistent model to estimation of pore aspect ratios for both patchy and uniform fluid saturation in Haynesville Shale. Bakhorji used the DEM model which has a set of pores with low aspect ratios to estimate elastic properties in low porosity sandstones (Bakhorji 2012). Mavko and Saxena developed an embedded-bound method for estimating the change in dynamic bulk moduli upon substitution of solid in the pore space (Mavko and Saxena 2013). Johansen proposed an inverse rock physics modeling (IRPM) strategy for estimation of lithology and other rock properties from seismic data (Johansen *et al.* 2013). Mikhail Markov proposed a novel approach for simulating the elastic properties of porous rocks based on the generalized differential effective medium (GDEM) method (Markov *et al.* 2013).

The above estimation methods, however, cannot estimate mineral matrix elastic moduli and pore structure simultaneously. In this paper, we propose an estimation method to estimate pore structure and mineral moduli simultaneously based on the KT model and Biot's coefficient. The technique is tested on 28 sandstone samples with high and low porosities. Firstly, pore aspect ratios of five different scales are constructed: 10^0, $N_1\times10^{-1}$, $N_2\times10^{-2}$, $N_3\times10^{-3}$, and $N_4\times10^{-4}$, and the initial values of N_1, N_2, N_3, and N_4 are randomly generated. Secondly, Biot's coefficient is used to build relationships among different rock elastic moduli; upper and lower bounds of mineral matrix moduli for all sandstone samples are calculated; and the initial estimated values within the bounds are randomly generated. Finally, simulated annealing (SA), global optimization algorithm (Metropolis *et al.* 1953), is used to optimize the randomly generated rock physics parameters until the minimum of two objective functions are reached and the generated parameters remain unchangeable. This leads to estimation of pore structure and mineral moduli of 28 sandstone samples. Comparison of relative errors of P-wave and S-wave velocities between fluid substitution results with lab measurements shows that the relative errors are all less than 5%. Moreover, when only having log data (porosity, bulk density, saturation, P-wave and S-wave velocities), this new estimation method can be directly used to estimate pore aspect ratios spectrum and mineral matrix elastic moduli accurately.

1.1 KT model and pore structure

The KT model bridges porosity, pore shape, mineral moduli, and fluid inclusion moduli with P-wave and S-wave velocities. This model assumes a low concentration of the inclusions, and also assumes that all pores are ellipsoidal, which can be described with pore aspect ratios. This parameter is defined by the ratio of short and long axes of the inclusion. The KT model can be written as:

$$(K_{sat}-K_{ma})\frac{(K_{ma}+\frac{4}{3}\mu_{ma})}{(K_{sat}+\frac{4}{3}\mu_{ma})}=\sum_{i=1}^{N}c_i(K_i-K_{ma})P^{mi} \tag{1}$$

$$(\mu_{sat}-\mu_{ma})\frac{(\mu_{ma}+\frac{4}{3}\zeta_{ma})}{(\mu_{sat}+\frac{4}{3}\zeta_{ma})}=\sum_{i=1}^{N}c_i(\mu_i-\mu_{ma})Q^{mi} \tag{2}$$

$$\zeta_{ma} = \frac{\mu_{ma}}{6} \frac{9K_{ma} + 8\mu_{ma}}{K_{ma} + 2\mu_{ma}} \tag{3}$$

$$\varphi = \sum_{i=1}^{N} c_i \tag{4}$$

In Eqs. 1-4, K_i and μ_i, respectively, are bulk moduli and shear moduli of i-th fluid inclusion, and K_{sat} and μ_{sat}, respectively, are saturated bulk moduli and saturated shear moduli. c_i is the volume fraction of i-th inclusion, and φ is the porosity. P^{mi} and Q^{mi} are strain concentration coefficients when adding i-th inclusion in mineral matrix (see Appendix). The coefficients are functions of mineral matrix elastic moduli, K_{ma} and μ_{ma}, and aspect ratio. Coefficients P^{mi} and Q^{mi} can be calculated by Wu's arbitrary aspect ratio (Wu 1966) and Berryman's 3D special pore types (Berryman 1995).

The spectrum of pore aspect ratios is the most important parameter in the KT model. Some researchers have studied the spectra of sandstones, limestones, and granites through electron microscope scanning measurements (Timur *et al.* 1971, Sprunt and Brace 1974, Hadley 1975). Table 1 lists pore aspect ratio spectra of some typical sandstone samples by using SEM electron scanning results (Toksöz *et al.* 1979). In Table 1, α is pore aspect ratio, $c(\alpha)$ is the volume fraction of the pore aspect ratio.

Table 1 shows results assuming that there are pore aspect ratios of five different scales for sandstone samples: 10^0, $N_1 \times 10^{-1}$, $N_2 \times 10^{-2}$, $N_3 \times 10^{-3}$ and

Table 1

The spectrum of pore aspect ratios of some typical sandstone samples (Toksöz *et al.* 1976)

Name	The spectrum of pore aspect ratios										Porosity
Boise sandstone	α	1	1E-1	3.5E-3	2.8E-3	2.1E-3	1.7E-3	1.3E-3	9E-4	5E-4	0.25
	$c(\alpha)$	1.877E-1	6.2E-2	3.7E-5	6.1E-5	6.4E-5	1.4E-4	1.1E-4	8.5E-5	6.2E-5	
Berea sandstone	α	1	1E-1	1E-2	1.7E-3	1.4E-3	1E-3	6E-4	3E-4	-	0.163
	$c(\alpha)$	1.399E-1	2.2E-2	3.7E-4	1.4E-4	1.4E-4	1.6E-4	1.3E-4	1.1E-4	-	
Navajo sandstone	α	1	1E-1	1E-2	3.5E-3	2.8E-3	2.1E-3	1.6E-3	1.2E-3	9E-4	0.1614
	$c(\alpha)$	1.416E-1	2.1E-2	2.4E-4	1.4E-4	2E-4	1.5E-4	1.8E-4	1.4E-4	6.5E-5	

$N_4 \times 10^{-4}$ (N_1, N_2, N_3, and N_4 are positive integers which are less than or equal to 10). We use these five pore aspect ratios to describe the pore structure in sandstones. Specifically, $\alpha = 10°$ represents a spherical pore, which is not closed under pressure. $\alpha = N_1 \times 10^{-1}$ is an intergranular pore. $\alpha = N_2 \times 10^{-2}$, $\alpha = N_3 \times 10^{-3}$, and $\alpha = N_4 \times 10^{-4}$ are crack pores, which are easily closed under pressure. For example, at $P_d = 500$ bars, all crack pores with aspect ratios smaller than 0.01 are closed (Toksöz *et al.* 1976).

Table 1 also shows that different pore aspect ratios have different volume fraction values. The crack pores only take a very small part of pore structure (less than 1% in each sandstone). These crack pores are indispensable, because the effect of lithological and porosity variations is minimal compared to the effect of the shape and size of crack pores (Kuster and Toksöz 1974, Anselmetti and Eberli 1999, Saleh and Castagna 2004). A low porosity rock may have more flat pores with low aspect ratios, while a high porosity rock may have more spherical pores with high aspect ratios (Wang 2001).

1.2 Biot's coefficient and mineral moduli

In rock physics models, the mineral matrix elastic moduli are essential. Although they can be calculated by Voigt-Reuss-Hill (VRH) averaging (Hill 1952) or Hashin-Shtrikman (HS) averaging (Hashin and Shtrikman 1963), mineral components and their percentages must be known firstly through rock thin section analysis. However, the rock thin section results have some problems:

1) There are a large number of mineral components in rocks. The thin section analysis can only provide contents of the main components. Mineral components with low concentrations cannot be observed. However, these minerals may have large elastic moduli that have significant effects on the overall responses of seismic signals.

2) The thin section analysis gives the average percentages of minerals in the samples. If the results of some mineral components have errors, the VRH averaging will fail.

3) Mineral matrix elastic moduli can be affected by diagenesis, formation pressure, temperature, and lithology. It's not always appropriate to refer and use elastic modulus of a mineral component from Mavko (Mavko 1998).

4) Sometimes, it is prone to have errors and abnormal results due to relaxed protocols in lab.

Therefore, an effective and accurate method for estimating mineral matrix elastic moduli is needed. In this paper, we use Biot's coefficient which had been discussed by Krief *et al.* (1990) and Nur *et al.* (1998). By building relations between different rock elastic moduli, the upper and lower bounds of mineral matrix bulk modulus and mineral matrix shear modulus can be obtained.

Krief *et al.* (1990) used the sandstone data of Raymer *et al.* (1980) to find an empirical relation for Biot's coefficient versus porosity:

$$K_{dry} = K_{ma}(1-\varphi)^{(3/(1-\varphi))} \tag{5}$$

$$\mu_{dry} = \mu_{ma}(1-\varphi)^{(3/(1-\varphi))} \tag{6}$$

These two equations are then transformed:

$$K_{ma} = K_{dry}(1-\varphi)^{(3/(\varphi-1))} \tag{7}$$

$$\mu_{ma} = \mu_{dry}(1-\varphi)^{(3/(\varphi-1))} \tag{8}$$

In rock physics, there are two general relations between mineral matrix elastic moduli and rock elastic moduli:

$$K_{ma} > K_{sat} > \mu_{dry} \tag{9}$$

$$\mu_{ma} > \mu_{sat} \tag{10}$$

Thus, when knowing porosity, dynamic rock bulk moduli and shear moduli of saturated sandstones, and using Eqs. 7-10, we get the upper and lower bounds for mineral matrix elastic moduli:

$$K_{sat} < K_{ma} = K_{dry}(1-\varphi)^{(3/(\varphi-1))} < K_{sat}(1-\varphi)^{(3/(\varphi-1))} \tag{11}$$

$$\mu_{sat} < \mu_{ma} = \mu_{dry}(1-\varphi)^{(3/(\varphi-1))} < K_{sat}(1-\varphi)^{(3/(\varphi-1))} \tag{12}$$

And these two equations are simplified to:

$$K_{sat} < K_{ma} < K_{sat}(1-\varphi)^{(3/(\varphi-1))} \tag{13}$$

$$\mu_{sat} < \mu_{ma} < K_{sat}(1-\varphi)^{(3/(\varphi-1))} \tag{14}$$

1.3 Objective function in SA

The upper and lower bounds only can limit the variation ranges of mineral matrix elastic moduli, but cannot give specific values. Therefore, in this paper, SA method is used to obtain the specific values within the upper and lower bounds of the mineral moduli. To perform the SA calculations, proper objective functions are required. In this new estimation method, the two objective functions can be defined by:

$$\Delta f_1 = \frac{\left|K_{sat}^{obs} - K_{sat}^{KT}\right|}{K_{sat}^{obs}} \le 1\% \tag{15}$$

$$\Delta f_2 = \frac{\left|\mu_{sat}^{obs} - \mu_{sat}^{KT}\right|}{\mu_{sat}^{obs}} \le 1\% \tag{16}$$

where Δf_1 is the objective function for mineral matrix bulk moduli while Δf_2 is for mineral matrix shear moduli. K_{sat}^{obs} and μ_{sat}^{obs} are bulk moduli and shear moduli, respectively, of the observed data. K_{sat}^{KT} and μ_{sat}^{KT} are bulk moduli and shear moduli, respectively, of the KT model calculation results. The objective functions for both mineral bulk moduli and shear moduli must be less than 1% in order to obtain high precision rock physics parameters in the estimation.

1.4 Estimation procedure

Estimating the bulk and shear moduli and pore aspect ratio distribution from velocity and porosity measurements is an underdetermined problem; we solve this problem by implementing a simulated annealing (SA) procedure. The execution of this estimation method consist of three steps (Fig. 1):

Step 1: Input P-wave velocity v_p^{obs}, S-wave velocity v_s^{obs}, water saturation s_w^{obs}, bulk density ρ, and porosity φ from laboratory data or log data.

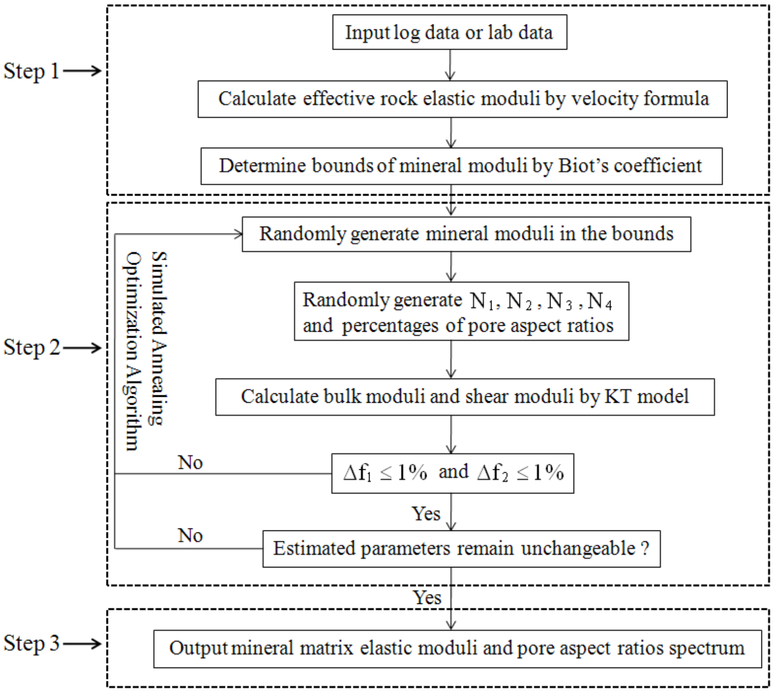


Fig. 1. Estimation procedures of this new estimation method.

Calculate dynamic rock bulk modulus K_{sat}^{obs} and shear modulus μ_{sat}^{obs} of saturated sandstones by velocity formula. And then determine bounds of mineral matrix bulk modulus and shear modulus by using Eqs. 13 and 14.

Step 2: Randomly generate K_{ma} and μ_{ma} within the bounds and randomly generate N_1, N_2, N_3, N_4, and percentages of pore aspect ratios: c_1, c_2, c_3, c_4, c_5. Use these randomly generated values in KT model to calculate bulk and shear moduli of rock samples. Then compare the calculated results of the rock elastic moduli with lab data and apply SA algorithm to minimize the relative errors to less than 1% by optimizing the corresponding two objective functions Δf_1 and Δf_2. And judge whether the estimated parameters (mineral moduli and pore aspect ratios spectrum) remain unchangeable or not in the SA iteration process.

Step 3: Repeatedly implement step 2, until the two objective functions reach the minimums (less than 1%) and the estimated parameters remain unchangeable. Finally, output the estimated parameters: pore aspect ratios spectrum and mineral matrix bulk moduli and shear moduli.

2. EXPERIMENTAL RESULTS

In order to validate this new estimation method, 28 sandstones core samples from a depth of 1446 m to 3496 m are used in simulations. The KT model assumes a low concentration of the inclusions, and our thin section identification results showed that the assumption is met for the experimental sandstones core samples. The experimental measurement procedure of these 28 sandstones core samples includes preparation, washing oil, drying, porosity measurement under pressure at 1.72 MPa, complete water and oil saturation, bulk density measurement (complete water and oil saturated conditions), P-wave velocity and S-wave velocity measurements (complete water and oil saturated conditions). The measurement was processed under the simulated formation temperature and pressure, and the results are shown in Fig. 2.

Figure 2a shows that the porosities of these 28 sandstone samples are from 1.96% to 22.42%. Thus, we can differentiate these sandstone samples between intermediate-low porosity samples (porosity less than 12.36%) and intermediate-high porosity samples (porosity greater than 12.36%). Figure 2b shows that the densities of these 28 sandstone samples are from 2.1 g/cm3 to 2.7 g/cm3, and the water saturated data are larger than the oil saturated data. Figure 2c shows that the P-wave velocities of these 28 sandstone samples are from 3500 m/s to 6000 m/s, and the water saturated data are larger than the oil saturated data. Figure 2d shows that the S-wave velocities of these 28 sandstone samples are from 2000 m/s to 3500 m/s, and the water saturated data are larger than the oil saturated data.

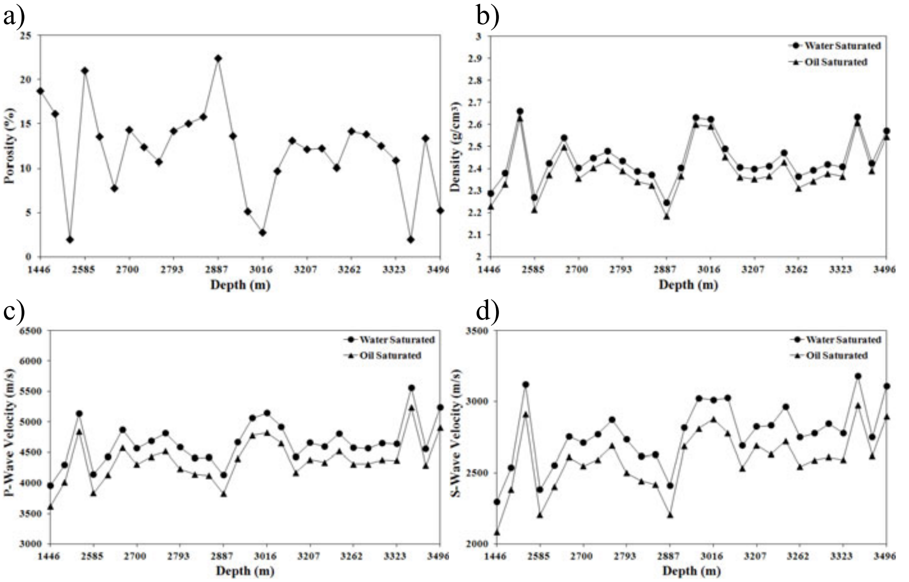


Fig. 2. Measured data of 28 sandstone samples under high pressure (pore pressure at 1.72 MPa and 68.95 MPa for porosity, bulk density and velocity measurements) and complete water and oil saturated conditions: (a) porosity, (b) density, (c) P-wave velocity (frequency at 0.7 MHz), (d) S-wave velocity (frequency at 0.2 MHz). Round line is water saturated data, triangle line is oil saturated data.

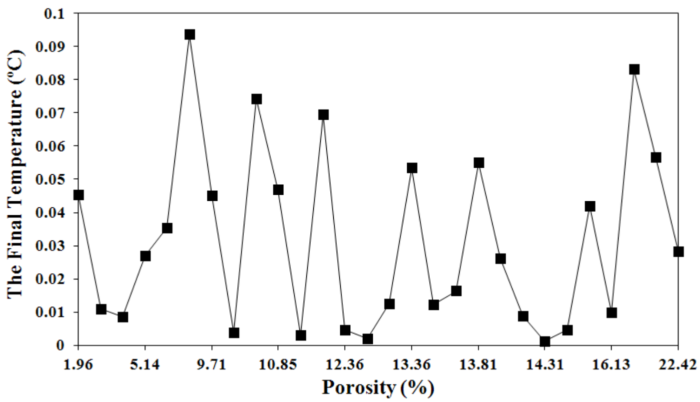


Fig. 3. The final temperature of every iteration process in the estimation procedures of 28 sandstone samples' parameters. In SA optimization algorithm, the initial temperature and the stop temperature of every iteration process respectively are 2000°C and 0.001°C, and the cooling factor is 0.99.

In SA optimization algorithm, the initial temperature and the stop temperature of every iteration process, respectively, are 2000°C and 0.001°C, and the cooling factor is 0.99, which ensure enough iteration time (the maximum iteration time is 1443) in the estimation procedures. In every iteration process, when the two objective functions are both less than 1% and the estimated parameters remain unchangeable, the final temperature in SA is recorded. Figure 3 is the final temperature of every iteration process in the estimation procedures of 28 sandstone samples' parameters. Figure 3 shows that all samples' final temperatures are less than 0.1 (iteration time is 985), and 9 samples' final temperatures are less than 0.01 (iteration time is 1215). This indicates that the final temperatures of every iteration process in the estimation are stable, and SA optimization algorithm is effective.

Figure 4 presents the calculated upper and lower bounds of mineral matrix bulk moduli, mineral matrix shear moduli and the corresponding estimated results. All estimated results are between the upper bounds and lower bounds while a few overlap with the bounds, and the bounds of shear moduli are greater than those of bulk moduli. The mineral elastic moduli are related to the mineralogy of samples, and from Figure 4 we can learn that the mineral constituent of our sandstone samples is various, and there is no clean sandstone sample.

Note that out of 13 samples with porosity less than 12.36%, only for 6 samples the mineral shear moduli are greater than the mineral bulk modulus. In addition, among 15 samples with porosity greater than 12.36%, only 6 samples' mineral shear moduli are greater than mineral bulk moduli. Therefore, for mineral matrix of this sandstone reservoir, bulk moduli are generally greater than shear moduli.

Specifically, for the 13 intermediate-low porosity samples (porosity less than 12.36%), the estimated results of bulk moduli are from 26.5 GPa to 45.9 GPa, and the estimated results of shear moduli are from 27.8 GPa to 40.4GPa For the 15 intermediate-high porosity samples (porosity greater

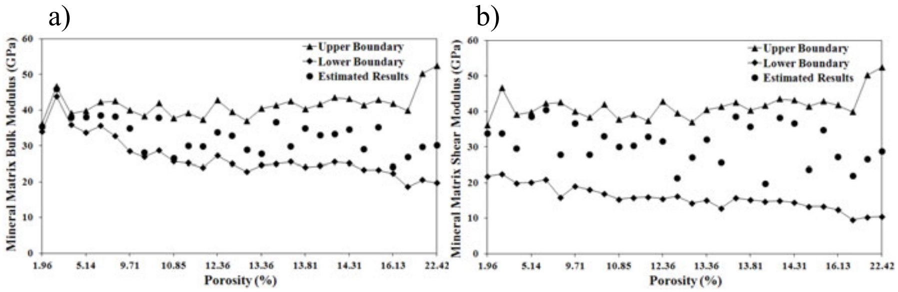
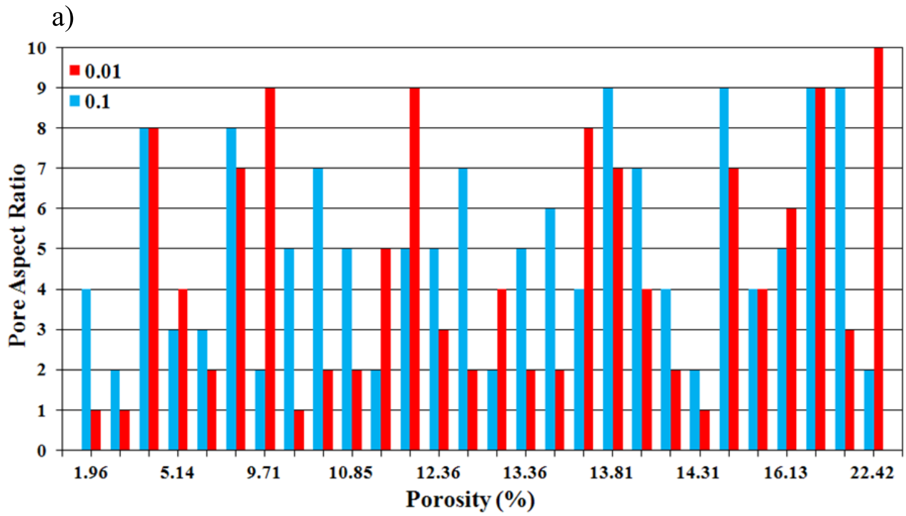


Fig. 4. Calculation results of upper and lower bounds for mineral moduli, and estimated results of these mineral moduli for 28 sandstone samples by this estimation method: (a) mineral matrix bulk modulus, (b) mineral matrix shear modulus. Triangle line is calculated upper boundary, diamond line is calculated lower boundary, and round dots are estimated results.



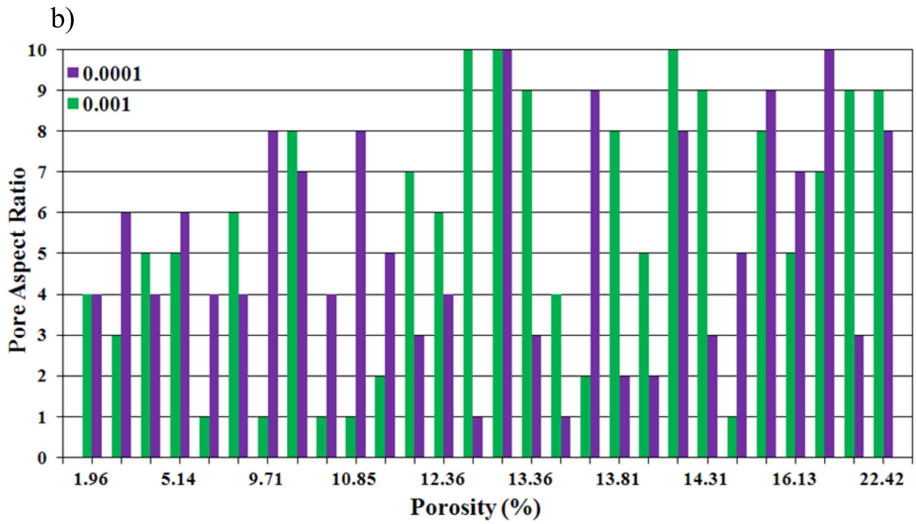


Fig. 5. The estimated results of pore aspect ratios of four different scales for 28 sandstone samples: (a) blue bar is estimated pore aspect ratio for N_1, red bar is estimated pore aspect ratio for N_2; (b) green bar is estimated pore aspect ratio for N_3, purple bar is estimated pore aspect ratio for N_4.

than 12.36%), the estimated results of bulk moduli are from 24.1 GPa to 36.6 GPa, and the estimated results of shear moduli are from 19.8 GPa to 38.6 GPa.

Figure 5 presents the estimated results of pore aspect ratios of four different scales for 28 sandstone samples. In Figure 5a, blue bar is for the estimated pore aspect ratio for N_1, red bar is for the estimated pore aspect ratio for N_2. In Figure 5b, green bar is for the estimated pore aspect ratio for N_3, and purple bar is for the estimated pore aspect ratio for N_4. Although

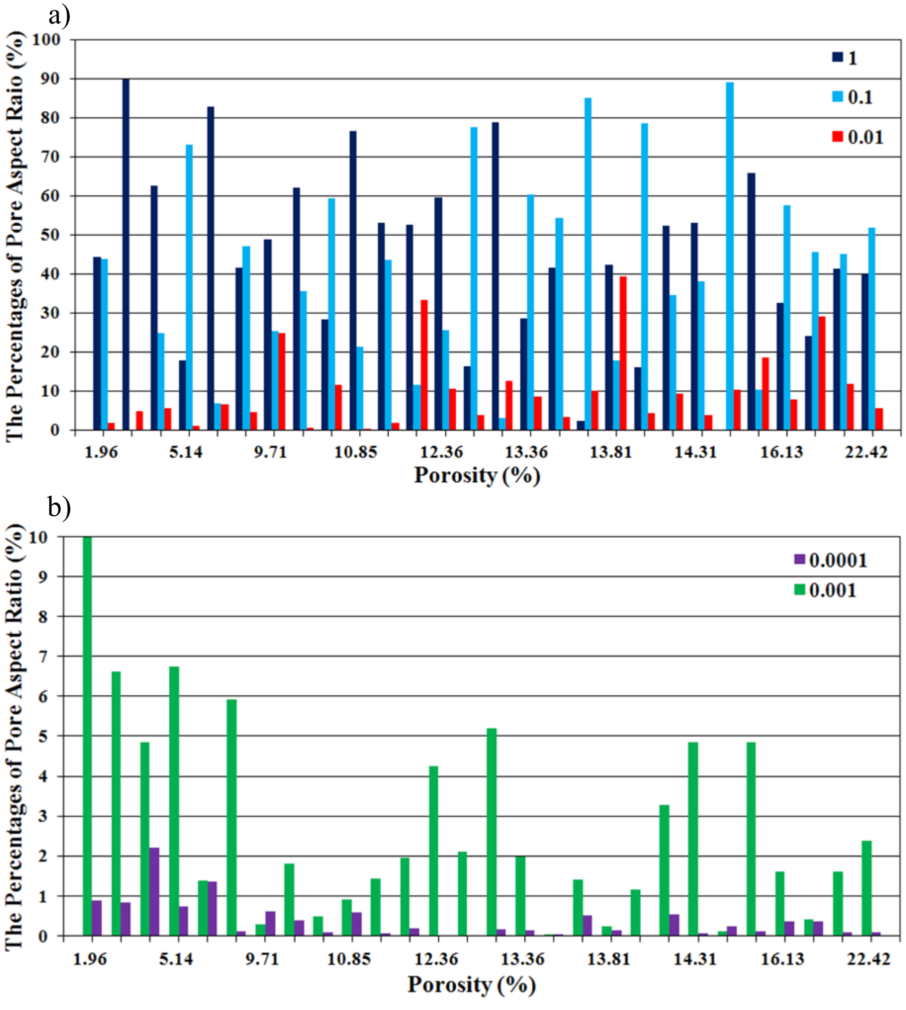


Fig. 6. The estimated results of the percentages of pore aspect ratios of five different scales for 28 sandstone samples: (a) dark blue bar is estimated percentages of pore aspect ratio for α = 10°, blue bar is estimated percentages of pore aspect ratio for $\alpha = N_1\times10^{-1}$, red bar is estimated percentages of pore aspect ratio for $\alpha = N_2\times10^{-2}$; (b) green bar is estimated percentages of pore aspect ratio for $\alpha = N_3\times10^{-3}$, purple bar is estimated percentages of pore aspect ratio for $\alpha = N_4\times10^{-4}$.

there is no significant feature for these estimated pore aspect ratios, it can be carefully seen from Fig. 5 that the estimated pore aspect ratios in intermediate-low porosity samples are a little smaller than those in intermediate-high porosity samples.

Figure 6 presents the estimated results of the percentages of pore aspect ratios of five different scales for 28 sandstone samples. In Fig. 6a, dark blue bar is for the estimated percentage of pore aspect ratio for $\alpha = 10°$, blue bar is for the estimated percentage of pore aspect ratio for $\alpha = N_1 \times 10^{-1}$, red bar is for the estimated percentage of pore aspect ratio for $\alpha = N_2 \times 10^{-2}$. In Fig. 6b, green bar is for the estimated percentage of the pore aspect ratio for $\alpha = N_3 \times 10^{-3}$, and purple bar is for the estimated percentages of pore aspect ratio for $\alpha = N_4 \times 10^{-4}$.

It can be concluded from Fig. 6 that the spherical pores and the intergranular pores appear to be the dominant type in the sandstone samples. Per centages of the spherical pores in 12 samples are between 50% and 90%, while percentages of the intergranular pores in 12 samples are between 50% and 89%. Crack pores have the lowest percentages, especially for the percentages of crack pores with $\alpha = N_3 \times 10^{-3}$ and $\alpha = N_4 \times 10^{-4}$. The former are less than 10% and the latter are less than 2.2%. It also can be learnt from Fig. 6b that there are more crack pores in the 6 low porosity samples (porosity less than 9.71%) than in the 15 intermediate-high porosity samples (porosity greater than 12.36%).

Thus, in this sandstone reservoir, pore structures of sandstone can mainly be characterized by spherical pores and intergranular pores; although the usage of the crack pores has a very low probability, they should not be neglected, especially in the low porosity sandstone samples.

3. FLUID SUBSTITUTION TESTING AND DISCUSSION

In order to test reliability and accuracy of the estimated parameters by this estimation method, we put estimated parameters (mineral moduli and pore aspect ratios spectrum of 28 samples) into KT model, change the fluid from water to oil saturated condition, and calculate P-wave and S-wave velocities of these sandstone samples, then compare calculation results with measured data and also calculate relative errors between the KT fluid substitution results and measured data. In the fluid substitution process, bulk modulus and shear modulus of the fluid inclusion are 1.35 GPa and 0 GPa, respectively.

Figure 7a is a comparison of theoretical data and measured data for oil saturated P-wave velocity. The theoretical data and measured data match very well for the 13 samples with intermediate-low porosity. The triangle dots (measured data) and round dots (theoretical data) are overlapped in 12 samples. And for the 15 samples with intermediate-high porosity, triangle

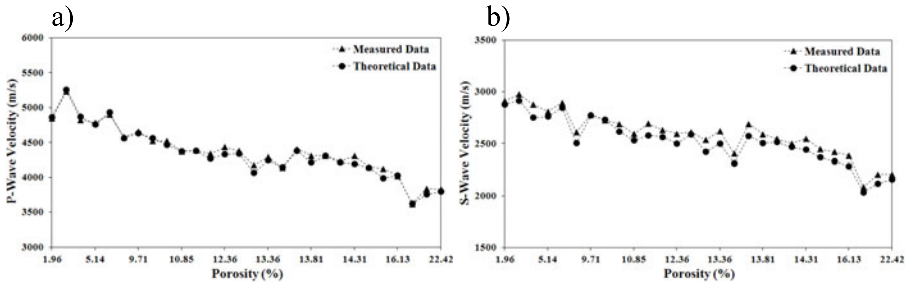


Fig. 7. Comparison between theoretical data (KT fluid substitution results) and measured data for 28 sandstone samples: (a) oil saturated P-wave velocity, (b) oil saturated S-wave velocity. Triangle dash line is measured data, round dash line is theoretical data.

dots and round dots coincide in 11 samples. These excellent data matches demonstrate that the estimation results by this new estimation method are reliable and accurate.

Figure 7b is a comparison of theoretical data and measured data for oil saturated S-wave velocity. Comparing Fig. 7b with 7a shows that the fitting of S-wave velocity is not as good as that of P-wave velocity. For the 13 samples of intermediate-low porosity, triangle dots and round dots are overlapped in 5 samples. For the 15 samples with intermediate-high porosity, triangle dots and round dots have excellent matches in 5 samples. But it doesn't mean the KT fluid substitution results for oil saturated S-wave velocity have poor fitting with the measured data. We also calculate the relative errors between KT fluid substitution results and measured data:

$$R_1 = \frac{\left| K_{oil}^{obs} - K_{oil}^{KT} \right|}{K_{oil}^{obs}} \tag{17}$$

$$R_2 = \frac{\left| \mu_{oil}^{obs} - \mu_{oil}^{KT} \right|}{\mu_{oil}^{obs}} \tag{18}$$

where R_1 is the calculated relative error for mineral matrix bulk moduli while R_2 is the calculated relative error for mineral matrix shear moduli. K_{oil}^{obs} and μ_{oil}^{obs} are bulk moduli and shear moduli of oil saturated data, respectively, of the measured results. K_{oil}^{KT} and μ_{oil}^{KT} are bulk moduli and shear moduli of oil saturated data, respectively, of the KT model fluid substitution calculation results.

Figure 7 shows relative errors between theoretical data (KT fluid substitution results) and measured data. It shows that the relative errors of both P-wave velocities and S-wave velocities of 28 sandstone samples are all less than 5%. It indicates that the estimated mineral moduli and pore aspect ratios spectrum are accurate by this new estimation method. Except for only 2 samples, the relative errors of P-wave velocities are smaller than those of S-wave velocities in most 26 sandstone samples. It demonstrates that the estimated results are more appropriate for predicting P-wave rock velocity.

4. CONCLUSIONS

This paper presents a new method for estimation of pore structure and mineral moduli based on KT model and Biot's coefficient. The new approach chooses pore aspect ratios of five different scales in KT model to construct pore structure of a sandstone reservoir. This avoids the disadvantages in previous studies that consider only one or few pore shapes to simulate pore structure in sandstone. The Biot's coefficient makes full use of relationships between different rock elastic moduli. It can properly determine upper and lower bounds of mineral elastic moduli when only log data are available.

This new estimation method of rock physics parameters can be applied directly to estimation of pore aspect ratios spectrum and mineral matrix moduli, which can help researchers better understand quantitative and nonlinear relations between different rock elastic moduli. The rock parameters estimated by the estimation method not only can perform fluid substitution and velocity prediction, but also can build mathematical physics relations between rock physics parameters (porosity, density, saturation, fluid type, pore shape) and seismic attributes (P-wave velocity, S-wave velocity, wave impedance, amplitude, AVO response).

Acknowledgements. This work was supported by the Key National Projects of Science and Technology of China (2011ZX05019-008-09) and the National Natural Science Foundation of China (41304115).

Appendix

Details of coefficient equations in KT model

We list the coefficient equations of P^{mi} and Q^{mi} in KT model which are described by Kuster and Toksöz (1974)

$$P^{mi} = \frac{1}{3} T_{iijj} \tag{A1}$$

$$Q^{mi} = \frac{1}{5}(T_{ijij} - \frac{1}{3} T_{iijj}) \tag{A2}$$

The scalars T_{ijij} and T_{iijj} which are used in this study are given by

$$T_{iijj} = \frac{3F_1}{F_2} \tag{A3}$$

$$T_{ijij} - \frac{1}{3} T_{iijj} = \frac{2}{F_3} + \frac{1}{F_4} + \frac{F_4 F_5 + F_6 F_7 - F_8 F_9}{F_2 F_4} \tag{A4}$$

where

$$F_1 = 1 + A[\frac{3}{2}(g+\theta) - R(\frac{3}{2}g + \theta - \frac{4}{3})] \tag{A5}$$

$$F_2 = 1 + A[1 + \frac{3}{2}(g+\theta) - \frac{R}{2}(3g + 5\theta)] + B(3-4R)$$
$$+ \frac{A}{2}(A+3B)(3-4R)[g + \theta - R(g - \theta + 2\theta^2)] \tag{A6}$$

$$F_3 = 1 + \frac{A}{2}[R(2-\theta) + \frac{1+\alpha^2}{\alpha^2} g(R-1)] \tag{A7}$$

$$F_4 = 1 + \frac{A}{4}[3\theta + g - R(g-\theta)] \tag{A8}$$

$$F_5 = A[R(g + \theta - \frac{4}{3}) - g] + B\theta(3-4R) \tag{A9}$$

$$F_6 = 1 + A[1 + g - R(g+\theta)] + B(1-\theta)(3-4R) \tag{A10}$$

$$F_7 = 2 + \frac{A}{4}[9\theta + 3g - R(5\theta + 3g)] + B\theta(3-4R) \tag{A11}$$

$$F_8 = A[1 - 2R + \frac{g}{2}(R-1) + \frac{\theta}{2}(5R-3)] + B(1-\theta)(3-4R) \tag{A12}$$

$$F_9 = A[g(R-1) - R\theta] + B\theta(3-4R) \tag{A13}$$

Then

$$A = (\mu_i / \mu_{ma} - 1) \tag{A14}$$

$$B = [(K_i/K_{ma}) - (\mu_i/\mu_{ma})]/3 \tag{A15}$$

$$R = 3\mu_{ma}/(3K_{ma} + 4\mu_{ma}) \tag{A16}$$

$$g = \alpha^2 (3\theta - 2)/(1 - \alpha^2) \tag{A17}$$

$$\theta = \frac{\alpha^2}{(1-\alpha^2)^{3/2}}[\cos^{-1}\alpha - \alpha(1-\alpha^2)^{1/2}] \tag{A18}$$

References

Ament, W. (1953), Sound propagation in gross mixtures, *J. Acoustic Soc. Am.* **25**, 638-641, DOI: 10.1121/1.1907156.

Anselmetti, F., and G.P. Eberli (1999), The velocity-deviation log: A tool to predict pore type and permeability trends in carbonate drill holes from sonic and porosity or density logs, *AAPG Bull.* **83**, 3, 450-466.

Bakhorji, A., H. Mustafa, S. Aramco, and P. Avseth (2012), Rock physics modeling and analysis of elastic signatures for intermediate to low porosity. **In:** *82nd Annual International Meeting, SEG*, Technical program expanded abstracts, 1-5.

Berryman, J. (1980), Long-wavelength propagation in composite elastic media, *J. Acoust. Soc. Am.* **68**, 1809-1831.

Berryman, J. (1995), Mixture theories for rock properties. **In:** T.J. Ahrens (ed.), *Rock Physics and Phase Relations: A Handbook of Physical Constants*, American Geophysical Union, Washington, D.C., 205-228, DOI: 10.1029/RF003p0205.

Budiansky, B. (1965), On the elastic moduli of some heterogeneous materials, *Mech. Phys. Solids* **13**, 4, 223-227, DOI: 10.1016/0022-5096(65)90011-6.

Cleary, C., G. Coates, and J. Dumanoir (1984), Theoretical and experimental bases for the dual-water model for interpertation of shaley sands, *Soc. Petrol. Eng. J.* **24**, 153-168, DOI: 10.2118/6859-PA.

Eshelby, J. (1957), The determination of the elastic field of an ellipsoidal inclusion, and related problems, *Proc. Roy. Soc. London* **A241**, 1226, 376-396, DOI: 10.1098/rspa.1957.0133.

Gassmann, F. (1951), Elastic waves through a packing of spheres, *Geophysics* **16**, 4, 673-682, DOI: 10.1190/1.1437718.

Hadley, K. (1975), Comparison of calculated and observed crack densities and seismic velocities in westerly granite, *J. Geophys. Res.* **81**, 20, 3484-3494, DOI: 10.1029/JB081i020p03484.

Hashin, Z., and S. Shtrikman (1963), A variational approach to the elastic behavior of multiphase materials, *Mech. Phys. Solids* **11**, 2, 127-140, DOI: 10.1016/0022-5096(63)90060-7.

Hill, R. (1952), The elastic behavior of crystalline aggregate, *Proc Phys. Soc*. **65**, 389, 349-354, DOI: 10.1088/0370-1298/65/5/307.

Hill, R. (1965), A self-consistent mechanics of composite materials, *Mech. Phys. Solids* **13**, 4, 213-222, DOI: 10.1016/0022-5096(65)90010-4.

Jensen, E.H., C.F. Andersen, and T.A. Johansen (2011), Estimation of elastic moduli of mixed porous clay composites, *Geophysics* **76**, 1, 9-20, DOI: 10.1190/1.3511351.

Johansen, T.A., E.H. Jensen, G. Mavko, and J. Dvorkin (2013), Estimate rock physics modeling for reservoir quality prediction, *Geophysics* **78**, 2, 1-18, DOI: 10.1190/geo2012-0215.1.

Krief, M., J. Garat, J. Stellingwerff, and J. Ventre (1990), A petrophysical interpretation using the velocities of P and S waves (full-waveform sonic), *The Log Analyst* **31**, 6, 355-369.

Kuster, G., and M. Toksöz (1974), Velocity and attenuation of seismic waves in two-phase media: Part 1 – Theoretical formulations, *Geophysics* **39**, 5, 587-606, DOI: 10.1190/1.1440450.

Lin, K., X.J. Xiong, X. Yang, Z.H. He, J.X. Cao, Z.X. Zhang, and P. Wang (2011), Self-adapting extraction of matrix mineral bulk moduli and verification of fluid substitution, *Appl. Geophys.* **8**, 2, 110-116, DOI: 10.1007/s11770-011-0278-0.

Markov, M., E. Kazatchenko, A. Mousatov, and E. Pervago (2013), Novel approach for simulating the elastic properties of porous rocks including the critical porosity phenomena, *Geophysics* **78**, 4, L37-L44, DOI: 10.1190/geo2012-0260.1.

Mavko, G., and N. Saxena (2013), Embedded-bound method for estimating the change in bulk moduli under either fluid or solid substitution, *Geophysics* **78**, 5, L87-L99, DOI: 10.1190/geo2013-0074.1.

Mavko, G., T. Mukerji, and J. Dvorkin (1998) *The Rock Physics Handbook*, Cambridge University Press, Cambridge.

Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller (1953), Equation of state calculations by fast computing machines, *J. Chem. Phys.* **21**, 1087-1092, DOI: 10.1063/1.1699114.

Mori, T., and K. Tanaka (1973), Average stress in matrix and average elastic energy of materials with misfitting inclusions, *Acta Metall.* **21**, 5, 571-574, DOI: 10.1016/0001-6160(73)90064-3.

Norris, A., P. Sheng, and A. Callegari (1985), Effective-medium theories for two-phase dielectric media, *Appl Phys*. **57**, 6, 1990-1996, DOI: 10.1063/1.334384.

Nur, A., G. Mavko, and J. Dvorkin (1998), Critical porosity: The key to relating physical properties to porosity in rocks, *The Leading Edge* **17**, 357-362, DOI: 10.1190/1.1887540.

O'Connell, R., and B. Budiansky (1974), Seismic velocities in dry and saturated cracked solids, *J. Geophys. Res.* **79**, 35, 5412-5426, DOI: 10.1029/JB079i035p05412.

Raymer, L.L., E.R. Hunt, and J.S. Gardner (1980), An improved sonic transit time-to-porosity transform. **In:** *21st Ann. Logg. Symp.*, SPWLA.

Saleh, A., and J.P. Castagna (2004), Revisiting the Wyllie time average equation in the case of near spherical pores, *Geophysics* **69**, 1, 45-55, DOI: 10.1190/1.1649374.

Sprunt, F., and W. Brace (1974), Direct observations of microcavities in crystalline rocks, *Int. J. Rock Mech. Min. Sci. Geomech. Abstr*. **11**, 4, 139-150, DOI: 10.1016/0148-9062(74)92874-5.

Timur, A., W. Hempkins, and R. Weinbrandt (1971), Scanning electron microscope study of pore systems in rocks, *Geophysics* **76**, 20, 4932-4948, DOI: 10.1029/JB076i020p04932.

Toksöz, M., C.H. Cheng, and A. Timur (1976), Velocities of seismic waves in porous rocks, *Geophysics* **41**, 4, 621-645, DOI: 10.1190/1.1440639.

Tran, D.T., and S. Chandra (2008), Changes in crack aspect-ratio concentration from heat treatment: A comparison between velocity estimation and experimental data, *Geophysics* **73**, 123-132.

Vernik, L., and M. Kachanov (2010), Modeling elastic properties of siliciclastic rocks, *Geophysics* **75**, 6, E171-E182, DOI: 10.1190/1.3494031.

Walsh, J. (1965), The effect of cracks on the compressibility of rock, *J. Geophys. Res.* **70**, 2, 381-389, DOI: 10.1029/JZ070i002p00381.

Wang, Z.J. (2001), Fundamentals of seismic rock physics, *Geophysics* **66**, 2, 398-412, DOI: 10.1190/1.1444931.

Wu, T. (1966), The effect of inclusion shape on the elastic moduli of a two phase material, *Int. J. Solids Struct.* **2**, 1, 1-8, DOI: 10.1016/0020-7683(66)90002-3.

Zimmerman, R. (1991), *Compressibility of Sandstones*, Elsevier, New York.

Zimmerman, R. (1984), The elastic moduli of a solid with spherical pores: New self-consistent method, *Int. J. Rock Mech. Min. Sci. Geomech. Abstr.* **21**, 6, 339-343, DOI: 10.1016/0148-9062(84)90366-8.

Received 19 November 2015
Received in revised form 17 April 2016
Accepted 15 June 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2356-2381
DOI: 10.1515/acgeo-2016-0116

Acoustic Full Waveforms as a Bridge between Seismic Data and Laboratory Results in Petrophysical Interpretation

Kamila WAWRZYNIAK-GUZ

AGH University of Science and Technology, Faculty of Geology, Geophysics and Environmental Protection, Department of Geophysics, Kraków, Poland; e-mail: wawrzyni@agh.edu.pl

Abstract

Mutual relationships between geological and geophysical data obtained by using methods of different scale are presented for the Miocene sandy-shaly thin-bedded formation and for the Zechstein carbonate formation. The common basis of laboratory results, well logging and seismic data was a recognition of elastic and reservoir properties of rocks. The common basis of laboratory results, well logging and seismic data were elastic and reservoir properties of rocks. Seismic attributes calculated from acoustic full waveforms were a link between the considered data. Seismic attributes strongly depend on small changes observed in rock formation related to lithology variations, facies changes, structural events and petrophysical properties variability. The observed trends and relationships of high correlation coefficients in the analysed data proved the assumption made at the beginning of research that common physical basis is a platform for data scaling. Proper scaling enables expanding the relationships determined from laboratory and well logging of petrophysical parameters to a seismic scale.

Key words: seismic attributes, acoustic full waveforms, elastic properties of rocks, upscaling.

---

1. INTRODUCTION

Successful combination of seismic data and well logging information that includes the results of laboratory experiments is an important way of scaling petrophysical parameters. Since the acoustic log is run in boreholes, seismic sections are correctly converted from time to depth scale with the use of velocity models based on well logging measurements. Similarly, seismic inversion that provides acoustic pseudoimpedance sections belongs to the methods which combine well logs and seismic data (Veeken and Da Silva 2004, Huuse and Feary 2005). Results of laboratory measurements, in particular velocity and bulk density, can also be included in seismic calibration procedures.

Laboratory data are the results of direct measurements of rock properties that are performed on the core or plug samples, sparsely taken from the borehole. They are precise and accurate but related to very small portion of rocks, and therefore they are considered as point-type data. Well logs are recorded continuously along the borehole and can be regarded as a 1D measurement, while seismic surveys represent 2D or 3D data sets. Different vertical resolution of these methods makes proper joint interpretation difficult. Scale dependence of geophysical data is an important and still current challenge for geophysicists and petrophysicists in precise determination of rock formation properties (Pechnig *et al.* 1997, Zoback 2010, Bui *et al.* 2010, Wenlong *et al.* 2012, Marzec *et al.* 2014, Krakowska and Puskarczyk 2015).

Measurements of the physical properties of rocks in the boreholes are recorded by many different logs, which is a characteristic feature of well logging. There are a lot of well established processing and interpretation procedures that are implemented in the commercial software and applications. However, vast and diverse information on geological, petrophysical and reservoir properties of rocks encourage scientists to looking for a new method or improving the existing processing and interpretation methods. Well logging data are suitable for statistical approach − there are many papers that propose application of various statistical methods for enhanced interpretation (*e.g.*, Szabó 2011, ElGhonimy and Sonneneberg 2015, Puskarczyk *et al.* 2015, Wawrzyniak-Guz *et al.* 2016). The other way of updating the interpretation process, especially in today's more complex formations, is adapting techniques known from other methods, as presented in this paper. Technique of seismic attributes computation was adapted for raw acoustic waveform data processing. The new acoustic logging algorithm was the key part of the research.

This paper proposes combination of data from laboratory experiments, well logging and seismic surveys by means of full waveform acoustic log.

Elastic properties and elastic wave propagation phenomena considered at each type of data were the common platform for coupling the information acquired at various scales. Acoustic full waveforms (AFW) were processed in commercial software just like seismic traces, and then seismic-origin attributes were calculated. Seismic attributes reflect instantaneous characteristics related to small changes of rock properties or small-scale tectonic events (Chopra and Marfurt 2005). Similarly, instantaneous characteristics of AFW are the response to small differences in rock formation (Knize 1989, Wawrzyniak-Guz and Jarzyna 2012, Wawrzyniak-Guz and Gruszczyk 2013). Due to distinct differences in seismic and sonic wave frequency and the other scale of inhomogeneity recognized by both methods, AFW attributes are not exact equivalents of seismic attributes. Nonetheless, attributes applied to AFW enable the interpreters to take advantage of the same physical background of sonic logs and seismics and, at the same time, to get closer to the scale of laboratory results than if they work with the primarily recorded signals only.

2. GEOLOGICAL FORMATIONS SELECTED FOR METHODOLOGY TESTS

Formations from two different geological units in Poland were selected for investigation: the Miocene (Sarmatian) formations in the Carpathian Foredeep and the Main Dolomite in the Fore-Sudetic Monocline. Three major conditions were taken into consideration to accept the region for the planned works: (i) good quality 3D seismic survey; (ii) full wave sonic logs (*i.e.*, acoustic full waveforms) recorded in the wells within the area of seismic survey; and (iii) great variety of laboratory results in the cored intervals of wells within the area of seismic survey, in particular P- and S-wave velocities.

The first region of research is located in the area of the Carpathian Foredeep (Fig. 1). The TCZ 3D seismic survey was done in the investigated area where C-2, C-3, C-5K, and M-1 wells were available (Fig. 2a). Research was focused on the Miocene part of the deposit in C-2, C-3, and C-5k wells and additional working material was included from M-1 well where acoustic full waveforms were recorded only in the carbonates of the Carboniferous basement. Intergranular porosity and high shale volume are typical features of the Sarmatian sandy-shaly thin-bedded formation (Bała *et al.* 2012, Jarzyna *et al.* 2013). A comprehensive data interpretation coming from various sources (seismics, well logging and laboratory experiments) is very important in thinly-bedded formations like heterolithes in molasse basins in the mountain foredeeps. The Miocene sediments in the Polish part of the Carpathian Foredeep are a good example of such formations (Marzec and Pietsch 2012, Pietsch *et al.* 2007). Sandy-shaly thin-bedded Miocene succes-

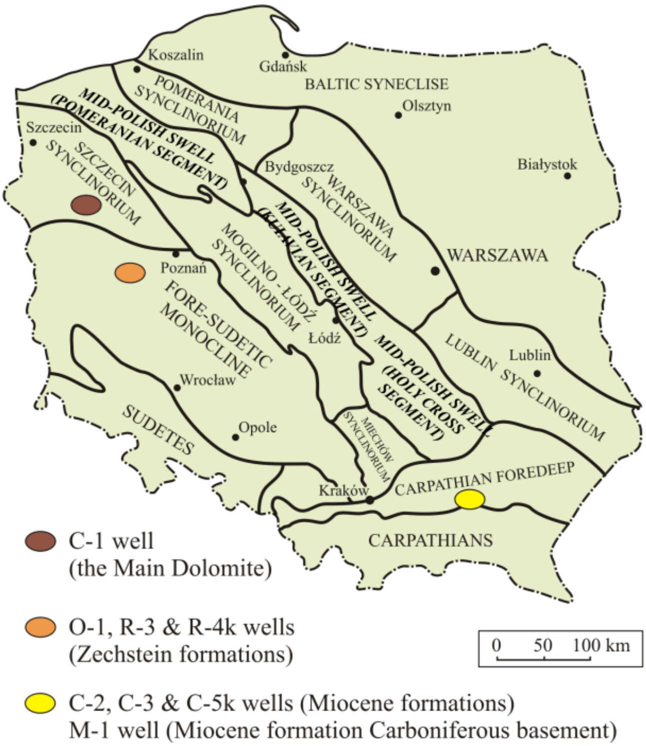


Fig. 1. Study areas on the background of main geological units in Poland.

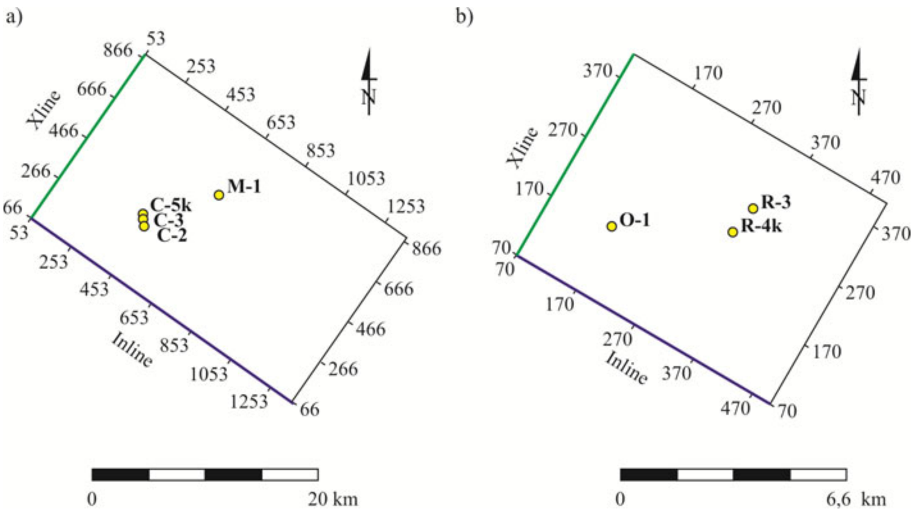


Fig. 2. Study areas: wells location on the background of selected Xline and Inline of TCZ 3D (a) and OR 3D (b) seismic projects.

sions, where pore-space is occupied by gas in sandstone layers (treated as reservoirs) and in mudstone layers (considered as source rocks), are very difficult to resolve because the thickness of both types of layers is very small; it ranges from several to dozens of centimetres. Small gas deposits identified by seismic surveys confirmed by well logs are very frequent in that region. Unfortunately, as often, there are boreholes without gas inflow drilled after seismic interpretation, despite similar stratigraphic and sedimentary positions. Unconventional shale gas deposits are similar to the discussed thin-bedded formations.

The second subject of investigation was the OR 3D seismic project where three wells: O-1, R-3, and R-4k were located in the study area (Figs. 1 and 2b). Full wave sonic logs covered the sequence of the Zechstein evaporates and carbonate sediments, including the Main Dolomite horizon. Additional C-1 well from the Polish Lowland (Fig. 1) was also included in research where the AFW were available in the Main Dolomite section. The Main Dolomite (Ca2) is a very important geological formation in hydrocarbon prospecting in the Polish Lowland and Fore-Sudetic Monocline. Within this dolomitic formation, formed on the carbonate-anhydrite platform, the biggest hydrocarbon deposits in the Polish Lowland were discovered. The Main Dolomite horizon is about 40-90 m thick and in many wells it may be divided into three different parts with regard to porosity and shale content. The upper part of the Main Dolomite horizon is always the most porous, the middle one is hard and of low porosity and low shaliness. The lowest part covers relatively shaly section. The dolomitic horizon reveals intergranular porosity as well as fractures and fissures which are frequently observed.

3. ACOUSTIC FULL WAVEFORM (AFW) ATTRIBUTES

Acoustic full waveforms are elastic wavefield recordings that are generated in the borehole by the transmitter in a sonic tool. Monopole source develops both body and surface waves, *i.e.*, P (compressional), S (shear) and Stoneley (surface) waves. Processing of AFW can include some seismic procedures, such as seismic attributes calculation, since sonic logs and seismics are based on the same physical phenomenon − elastic wavefield.

Seismic attributes were previously incorporated into AFW interpretation, however they were mainly considered as a visual enhancement of qualitative approach (Bała and Jarzyna 1996). In the presented paper, numerical values of attributes were determined for quantitative interpretation. The results were applied to combine acoustic wavefield (represented by AFW recorded in borehole) with seismic wavefield (recorded on the surface). Seismic attributes were treated as a link between these two representations of elastic wavefields. AFW were also the connection between well logging and labora-

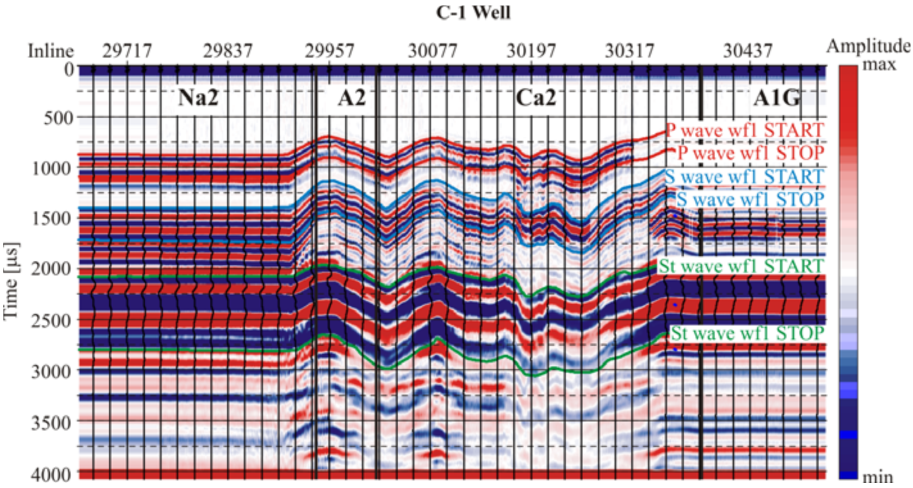


Fig. 3. AFW from C-1 well loaded to seismic software. Horizontal axis represents logging depth (MD) in meters (multiplied by factor 100), vertical axis represents recording time of AFW in μs. Colour lines indicate arrival (_START) and end (_STOP) time of P, S, and Stoneley (St) wave packets. Na2 – Older Halite, A2 – Basal Anhydrite, Ca2 – Main Dolomite, A1G – Upper Anhydrite.

tory results. AFW attributes were related to small changes of physical properties of rock and reflected instantaneous characteristics of formation.

Seismic attributes for P, S, and Stoneley waves were calculated with the use of standard seismic software: Hampson-Russell Suite, version 9. An example of AFW recorded in the Zechstein formations: salts (Na2), anhydrites (A2 and A1G), and carbonates (Ca2) in the Polish Lowland is presented in Fig. 3. Data were recorded by Full Wave Sonic tool (Halliburton Co.).

4. METHODOLOGY AND EXAMPLES

Calculation of attributes from AFW required special preprocessing of full waveform sonic logs that enabled loading AFW signals into the seismic software just like any other seismic section. It included, inter alia, file format conversion and editing the file headers. Once the preprocessing was completed, the attributes could be computed with the use of algorithms available in the seismic program, such as instantaneous or windowed attributes. Scheme of the methodology is presented on the basis of C-1 well in the Polish Lowland; more details can be found in Wawrzyniak-Guz and Jarzyna (2012).

Recording time of AWF was presented on the vertical axis and the logging depth was on the horizontal axis (Figs. 3 and 4). Next, using the “picking horizon” tool in the seismic software, the arrival time (_START) and the

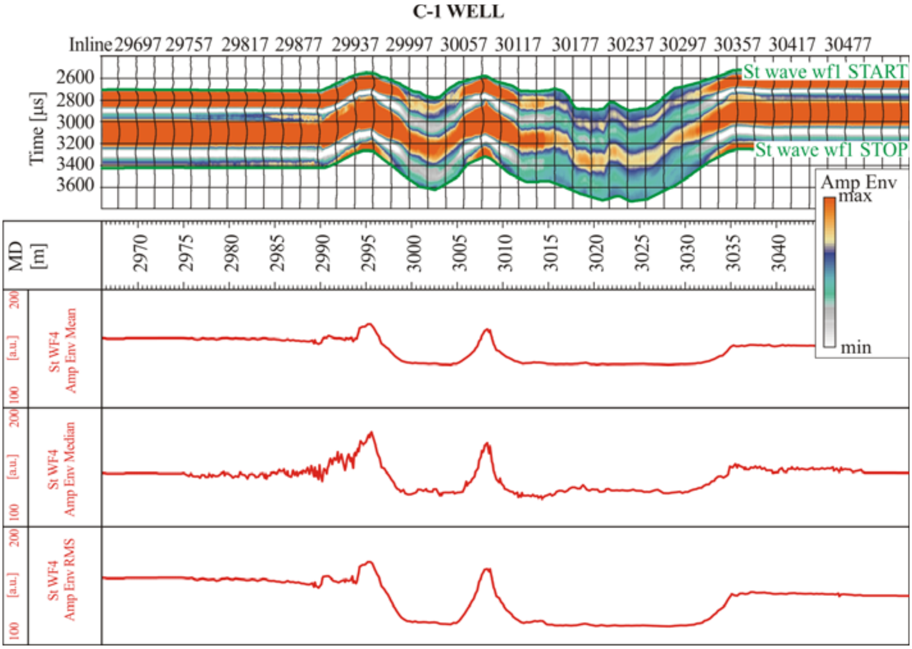


Fig. 4. Method of seismic attributes for AFW calculation. All attributes, calculated between the arrival and the end time of P, S, and Stoneley (St) waves, were averaged with arithmetic mean, median and root mean square (RMS). Here, examples of the amplitude of the envelope (Amp Env) for Stoneley (St) from C-1 well are presented.

end time (_STOP) of P, S, and Stoneley waves were pointed at every waveform (Fig. 4). The end time was chosen arbitrarily; however, the idea was to include the main oscillations of the wave packets. The same phase of the signal was very carefully picked throughout the sections, similar to the phase correlation in picking horizons in seismics. Five seismic attributes for P, S, and Stoneley waves were calculated: Isochron (the time interval between _START and _STOP times) and the complex trace attributes (Taner *et al.* 1979): Amplitude Envelope, Instantaneous Frequency, Instantaneous Phase, and Cosine of Instantaneous Phase. Attributes were calculated only between _START and _STOP times for individual waves (*i.e.*, between "horizons"), and then averaged using arithmetic mean, median, and root mean square (RMS). As a result, five seismic attributes for each P, S, and Stoneley wave were available for further analyses. Additional processing allowed to present these attributes as the standard well logs, *i.e.*, they showed how the attributes changed with depth (Figs. 5-7). Results for C-1 well are presented in Fig. 5. Significant changes of the attributes *versus* depth related to fluctuations in mineral composition (admixture of anhydrite is one of the most substantial in

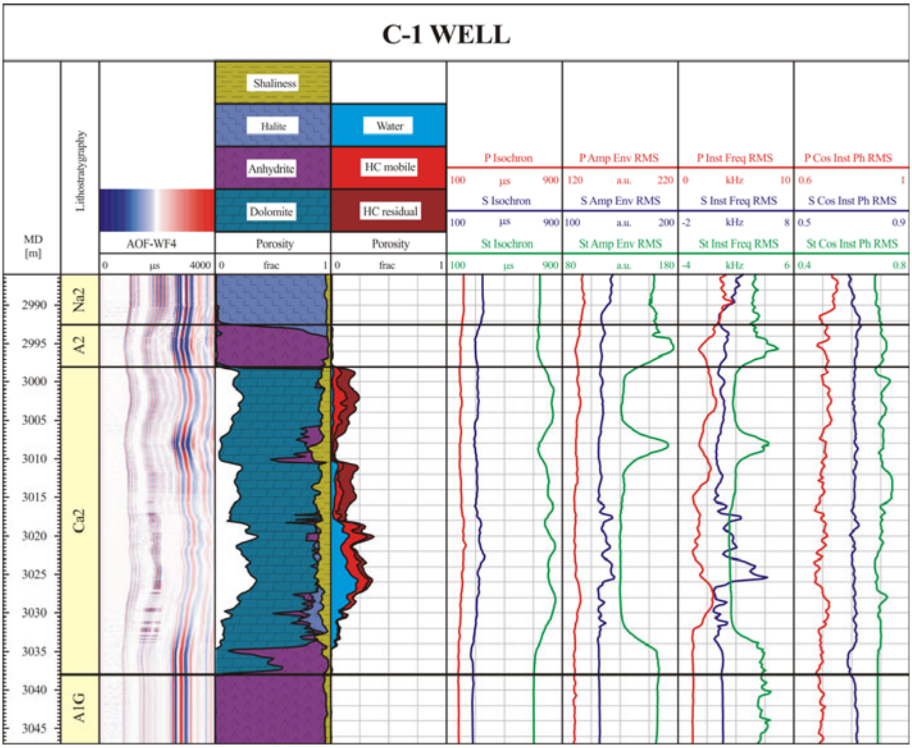


Fig. 5. Seismic attributes computed for AFW in C-1 well, Polish Lowland. The scales of the attributes are adjusted for better visualisation and curve separation. Here, only results of RMS averaging of the Amplitude Envelope (Amp Env), Instantaneous Frequency (Inst Freq), and Cosine of Instantaneous Phase (Cos Inst Ph) are presented.

this case), porosity, and water/hydrocarbon saturation are distinctly visible. For example, anhydrite contribution significantly amplifies amplitudes of Stoneley waves, whereas an increase of porosity lengthens the time span of all waves expressed by the Isochron attribute. Calculated attributes are vulnerable to unwanted logging effects and errors. For instance, artificial amplification of S wave between 3014-3027 m had a harmful effect on the Amplitude Envelope and the Instantaneous Frequency (S Amp Env and S Inst Freq, respectively) (Fig. 5).

AFW attributes, with the use of the developed methodology, were computed for the sandy-shaly thin bedded Miocene formation in C-2 well (Fig. 6) and for carbonates of the Carboniferous basement in M-1 well (Fig. 7). The amplitude Envelope of P, S, and Stoneley waves represent wave amplitudes, and Instantaneous Frequency informs about the frequency of the particular waves.

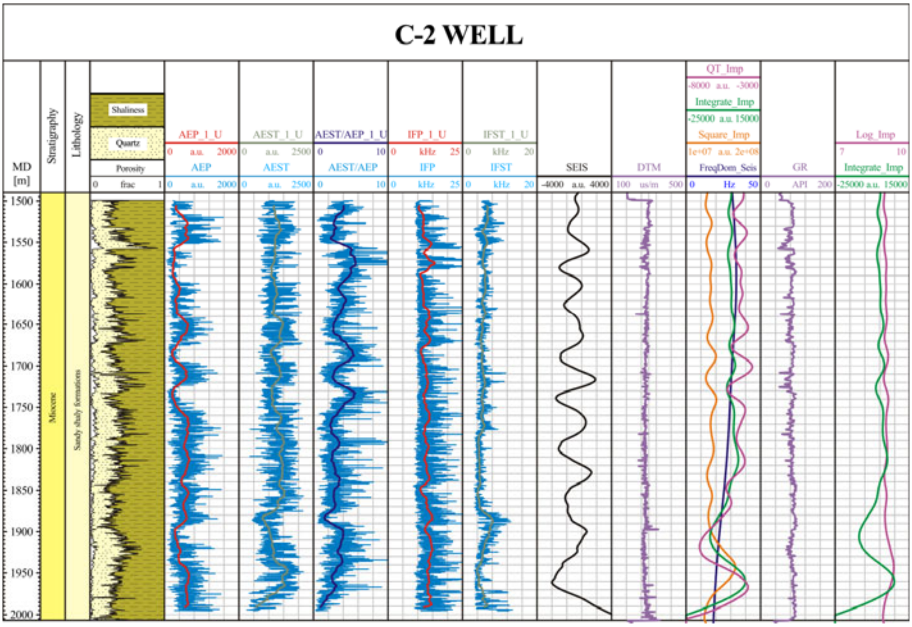


Fig. 6. Attributes calculated from acoustic waveforms and from seismic data in C-2 well, Carpathian Foredeep. AE – Amplitude Envelope; IF – Instantaneous Frequency; P, S, St – indicate compressional, shear and Stoneley waves, respectively; _1_U suffix means upscaled data; SEIS – seismic trace, QT_Imp, Integrate_Imp, Square_Imp, FreqDom_Seis, Log_Imp are seismic attributes (explanations are in the text); DTM and GR are sonic and gamma ray logs.

Well log vertical resolution is much greater than seismic data, so it was necessary to upscale the attributes calculated from AFW to the seismic wavelength. Running average was applied to each attribute. The averaging window length (L) used in this study was determined on the basis of Backus theory and the method proposed in the paper of Liner and Fei (2006): $L = V_{SV}^{(\min)}/N{\cdot}F_{\mathrm{dom}}$, where $V_{SV}^{(\min)}$ is the minimum S-wave velocity derived from sonic log (from the standard interpretation of full wave sonic log), F_{dom} is the dominant frequency of seismic data, here taken from the seismic attribute FreqDom_Seis (Fig. 6), and N is a positive integer to be chosen; $N = 3$ gives a reasonable level of accuracy that preserves essential part of the seismic wavefield (Liner and Fei 2007).

Applying the Backus theory was justified here because the Miocene formations in the Carpathian Foredeep exhibit thin layering. Similarly, limestones in Carboniferous basement reveal layer-induced seismic anisotropy (VTI anisotropy). Hence, the attributes were upscaled to seismic wavelength with 13-m and 28.5-m averaging window in C-2 and in M-1 well, respectively (Figs. 6 and 7).

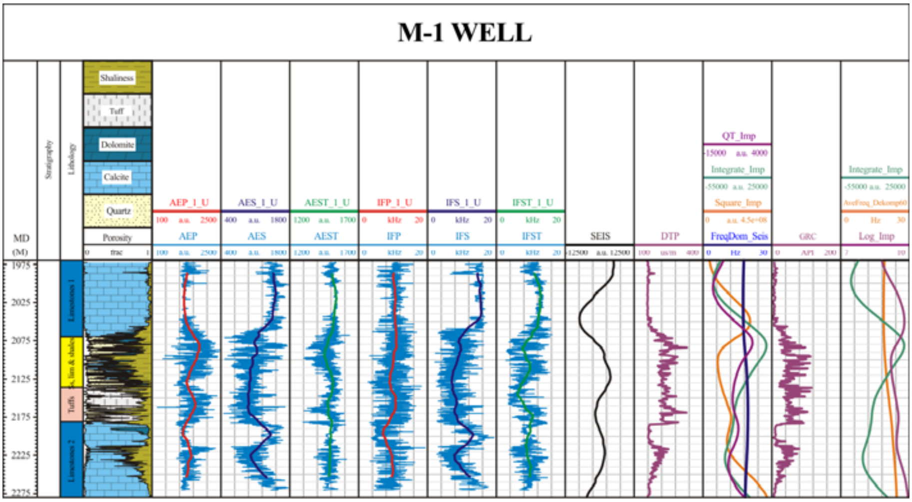


Fig. 7. Attributes calculated from acoustic waveforms and from seismic data in M-1 well, Carpathian Foredeep. Explanations are the same as in Fig. 6.

5. SEISMIC ATTRIBUTES

There are plenty of attributes that can be calculated from seismic signals (*e.g.*, Taner *et al.* 1979, Chopra and Marfurt 2005). Seismic attributes are usually applied to qualitative seismic interpretation, aimed to recognize geological structures, sedimentation environment and water/hydrocarbon saturation. Selected sets of seismic attributes can be also used for well log data prediction (Schultz *et al.* 1994a, b; Ronen *et al.* 1994). Similar approach was proposed in this research. Two logs − sonic and gamma − ray were chosen as the representatives of elastic properties (DT) and lithology features (GR) of the investigated formations. Several seismic attributes were calculated at well locations on the basis of standard seismic traces, traces after spectral decompositions, and acoustic impedance traces (Wawrzyniak-Guz and Gruszczyk 2013). The set of best-fitted attributes were chosen with the use of multi-linear regression. Calculations were done for wells located in the area of TCZ 3D seismic survey. Though the prediction was made in the sandy-shaly Miocene formation (Fig. 8), the attributes were calculated over the whole time interval of seismic data. The attributes were later converted to depth domain and presented in well logging data manner along with AFW attributes. The complete set of attributes for DT prediction was as follows: Quadrature Trace (from impedance traces), QT_Imp; Dominant Frequency (from seismic traces), FreqDom_Seis; Integrate (from impedance traces), Integrate_Imp; and Square (from impedance traces), Square_Imp. For GR prediction, the following groups of attributes were chosen: Average Frequency

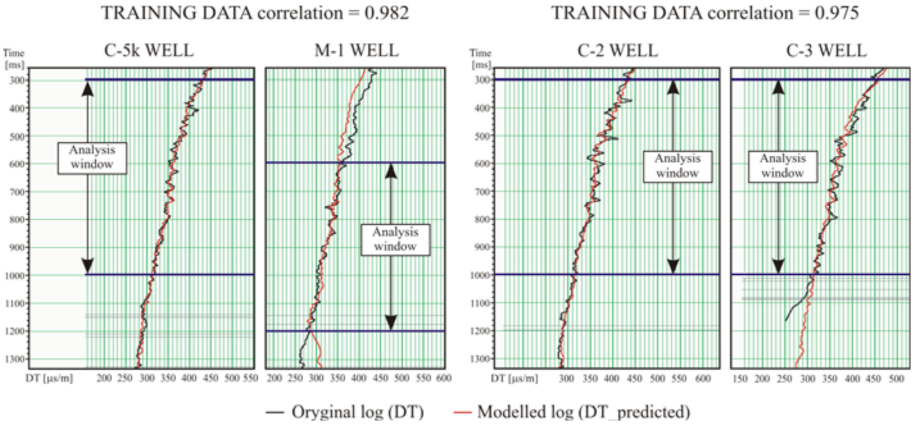


Fig. 8. Results of acoustic log prediction on the basis of seismic attributes in the Miocene formations, Carpathian Foredeep.

(from spectral decomposition 60Hz traces), AveFreq_Dekomp60; Logarithm (from impedance traces), Log_Imp; and Integrate (from impedance traces). Results are presented in Figs. 6 and 7. These attributes were later combined with attributes from AFW. The same procedures were applied to data from OR 3D seismic project.

6. RELATIONSHIPS BETWEEN SEISMIC DATA, WELL LOGS AND LABORATORY INFORMATION

Dozens of relationships between seismic data, well logging results and laboratory parameters were tested to illustrate the mutual dependence of petrophysical properties determined from different methods. The most important ones were these including reservoir properties and elastic parameters since these parameters were considered as the common platform for comparisons. Results of laboratory experiments and well logging data (measurements and interpretation) were inspected by cross-plots and correlations. Similarly, relationships between well logs and variety of seismic parameters were sought. Also, the selected seismic attributes were joined with laboratory results. From a great number of relationships that were examined, several distinct examples were selected for illustration of the results.

7. RESULTS OF SEISMIC INVERSION INCLUDING LABORATORY P-WAVE VELOCITY

Fragments of TCZ 3D seismic survey (Carpathian Foredeep) were used for seismic inversion based on geological model. The inversion was computed in the vicinity of M-1, C-2, C-3, and C-5k wells. Input data, which were used for geological model construction, were composed of the P-wave impedance

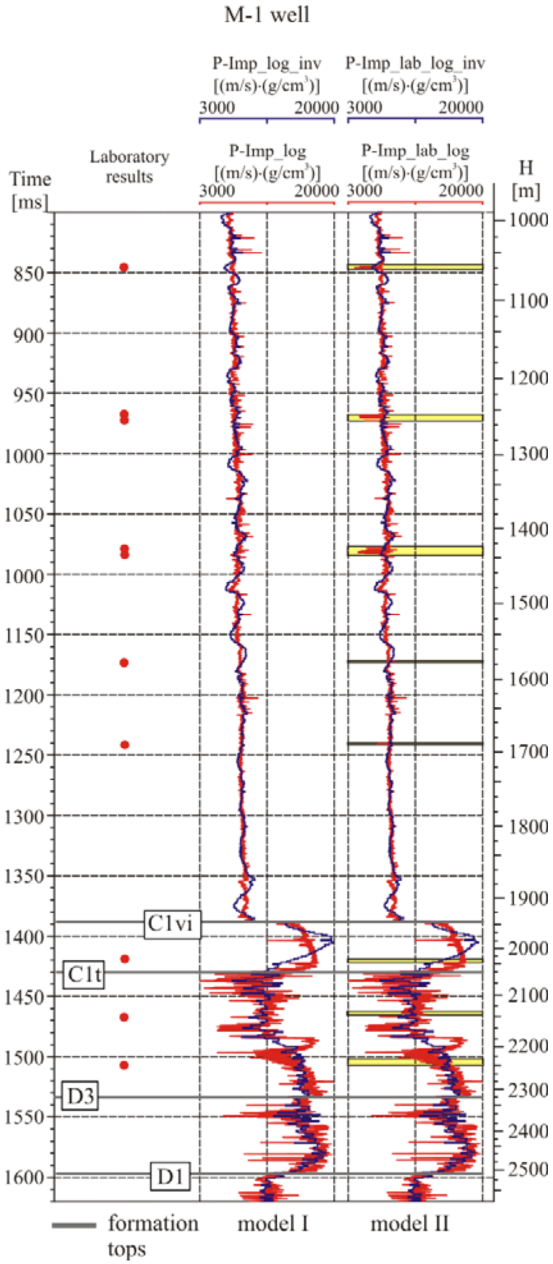


Fig. 9. Velocity models prepared for seismic calibration from time to depth on the basis of well log data only (model I – Track 1) and including laboratory results into well logs (model II – Track 2). Red dots present depth of core samples for laboratory measurements. Explanation of formation tops: C1vi − top of Visean; C1t − top of Tournaisian; D3 − top of Late Devonian; D1 − top of Early Devonian.

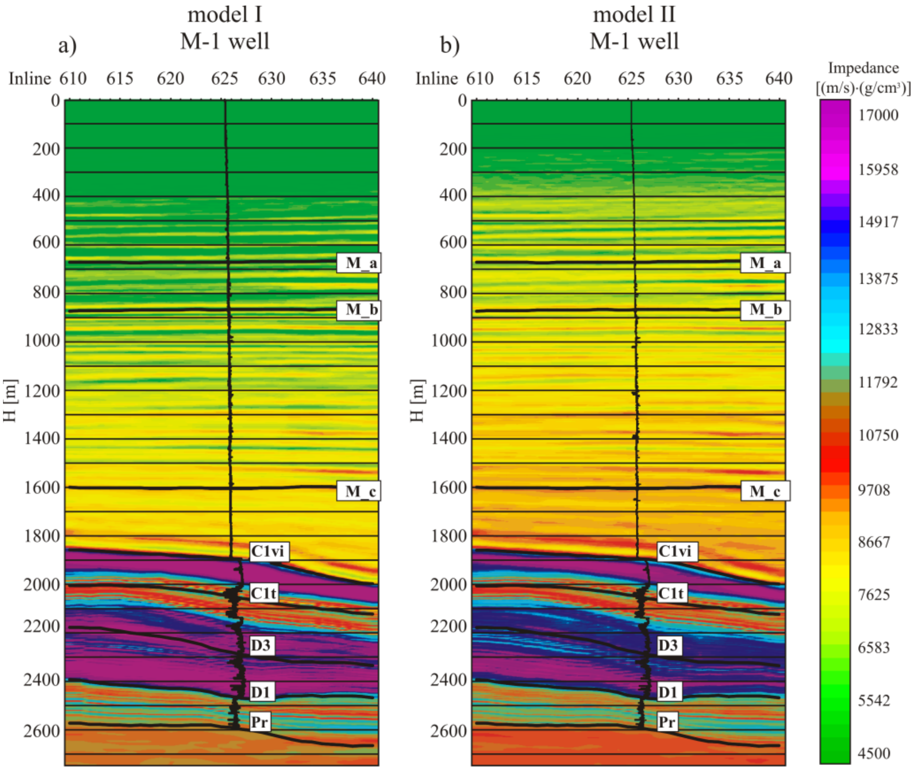


Fig. 10. Result of seismic inversion in the pseudoimpedance version using velocity models I and II. Curves inserted into well location are P-Imp_log and P-Imp_lab_log, respectively. Explanation of formation tops: M_a, M_b, M_c – lithostratigraphic horizons within the Miocene formation; C1vi – top of Visean; C1t – top of Tournaisian; D3 – top of Late Devonian; D1 – top of Early Devonian; Pr – top of Precambrian basement.

curve calculated from sonic and density logs, and previously interpreted seismic horizons. For the M-1 well, a set of laboratory measurements of P-wave velocity and bulk density was available. Including laboratory data in the inversion was advisable for proper combining data acquired at different scales.

In the M-1 well, primary velocity model (model I) was built with the use of logs only. Then, it was modified to model II by including P-wave velocity values from laboratory experiments (Fig. 9). Pseudoimpedance sections presented in Fig. 10 were the results of seismic inversion based on models I and II calculated around M-1 well. In the Sarmatian sediments, the inversion based on model II revealed some layers of higher values of pseudo-impedance in comparison to results obtained with model I. More yellow coloured beds are visible in the upper fragment of the profile, where shaly-sandy sediments are water saturated. In the lower sections of the geological profile (below 1941.5 m), pseudoimpedance from model II is lower than from model I. This interval is composed of Carboniferous and Devonian rocks, including sandstones and carbonates. More blue coloured beds are visible in that section of profile, which means lower impedance. Acoustic impedance from seismic inversion is not simply related only to bed velocities but also depends on reflection coefficients at the seismic boundaries. Thus, the interpretation of results presented in Fig. 10a and b is rather qualitative than quantitative.

Results of inversion based on model II shows slightly lower values in the depth intervals where lab data were included in comparison to P-impedance calculated from logs only (Fig. 11).

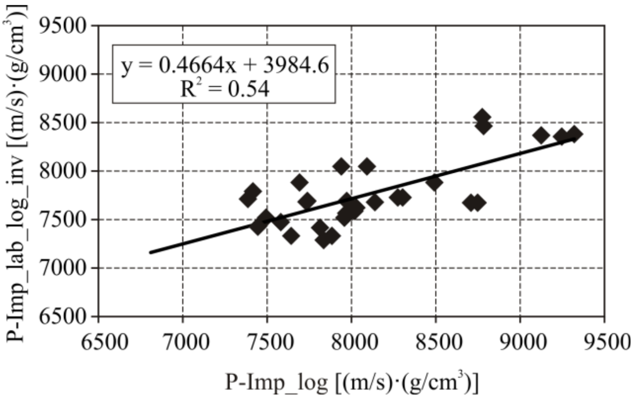


Fig. 11. P-impedance being the result of seismic inversion based on velocity model from logs and laboratory results (model II) *versus* P-impedance from well logs only.

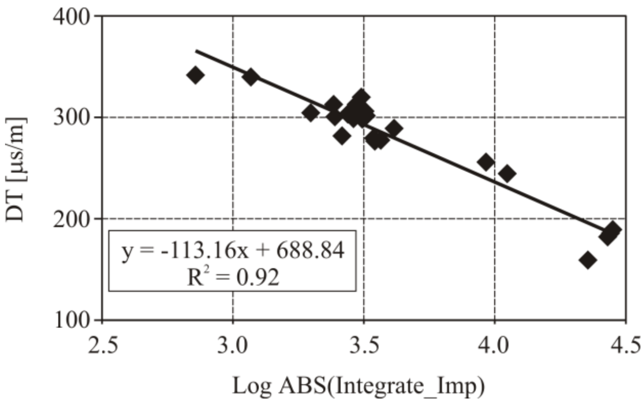


Fig. 12. Interval transit time *versus* logarithm of absolute value of integral calculated from acoustic impedance (seismic attribute).

The logarithm of the absolute value of the integral of acoustic impedance from seismic inversion (Log ABS (Integrate_Imp)) revealed a clear correlation with transit interval time (DT) from acoustic log (Fig. 12). Such strong relationship proves applying seismic attributes to determination P-wave velocity in the area covered by seismic surveys and can enhance the ability to obtain total porosity from seismic inversion.

8. RELATIONSHIPS BETWEEN DIFFERENT SCALE DATA

Crossplots and relationships between parameters acquired at different scales: lab-log, log-seis, log-log, and lab-seis, were calculated in this study and several examples are presented in Figs. 13-22. Found correlations proved the assumption made at the beginning that the common physical background of

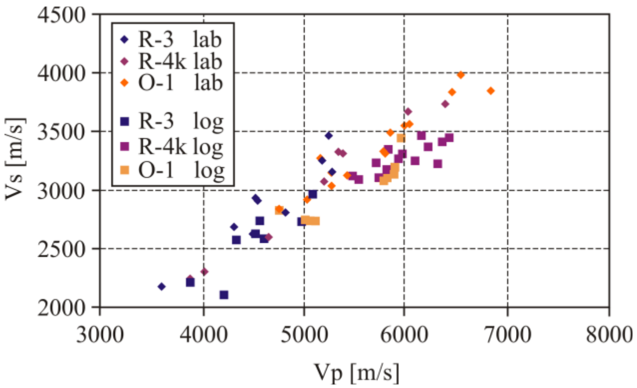


Fig. 13. S wave velocity *versus* P wave velocity – comparison of lab and well log (WL) data in R-3, R-4k, and O-1 wells for the Main Dolomite.

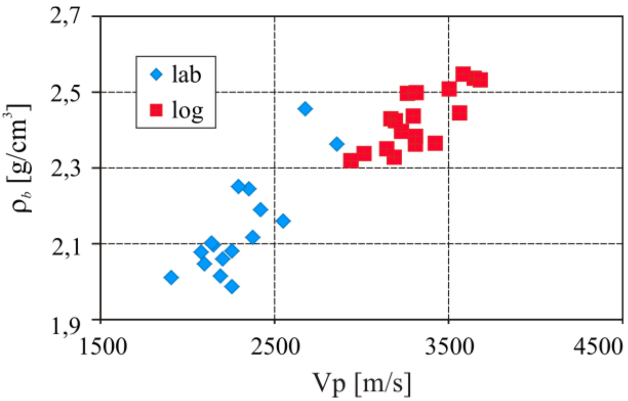


Fig. 14. Relationship between bulk density and P-wave velocity in M-1well for Miocene formation.

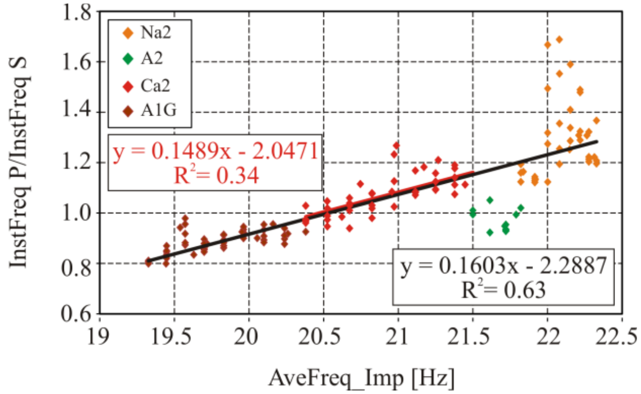


Fig. 15. Amplitude Envelope of P and S wave ratio (from AFW) *versus* average frequency of impedance trace (from seismic inversion) in R-4k well. Data represent the Zechstein formation: the Main Dolomite (Ca2), anhydrites (A1G, A2), and salts (Na2).

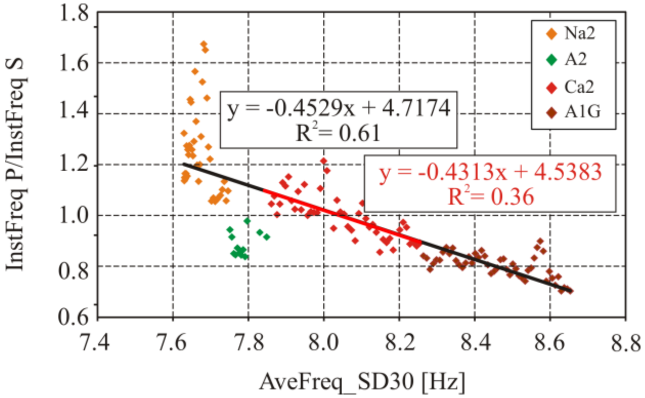


Fig. 16. Instantaneous frequency of P and S waves ratio (from AFW) *versus* average frequency of spectral decomposition 30 Hz trace in R-4k well. Data represent the Zechstein formation: the Main Dolomite (Ca2), anhydrites (A1G, A2), and salts (Na2).

wave velocity measured in laboratory, by acoustic logs, and in seismics, *i.e.*, elastic properties of rocks and elastic wavefield, could be successfully applied to join and combine petrophysical information acquired at different scales.

Firstly, P and S wave velocities measured in the laboratory on rock samples were compared with the velocities from acoustic full waveform logs (Fig. 13). Presented log data were chosen from the same depths as core and rock samples. The comparison reveals that the velocities cover the same

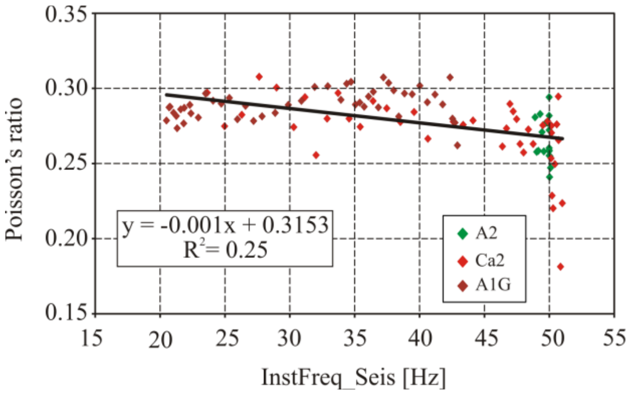


Fig. 17. Poisson's ratio (from well logging) *versus* instantaneous frequency of seismic trace in R-4k well. Data represent the Zechstein formation: the Main Dolomite (Ca2), anhydrites (A1G, A2).

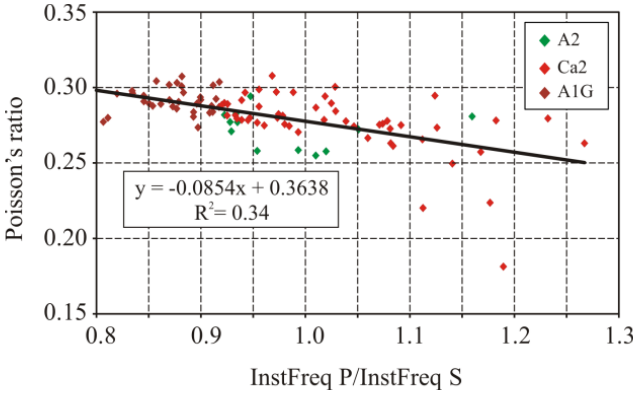


Fig. 18. Poisson's ratio (from well logging) *versus* instantaneous frequency of P and S wave ratio (from AFW) in R-4k well. Data represent the Zechstein formation: the Main Dolomite (Ca2), anhydrites (A1G, A2).

range of values and lie along the same trend; however, the lab velocities are a bit higher than those from the logs. This is the result of higher frequency of elastic waves induced by transducers in laboratory tests (order of MHz) than the wave frequency used in well logging (order of kHz). Correlation proves the consistency of elastic properties of rocks at different scales and justifies further analyses.

Relationship between bulk density and P-wave velocity in M-1 well (Fig. 14) shows two separate groups of data. Laboratory data have lower values than well logging results, but the general trend is similar. This is caused by both geological and technical reasons. The geological reason of

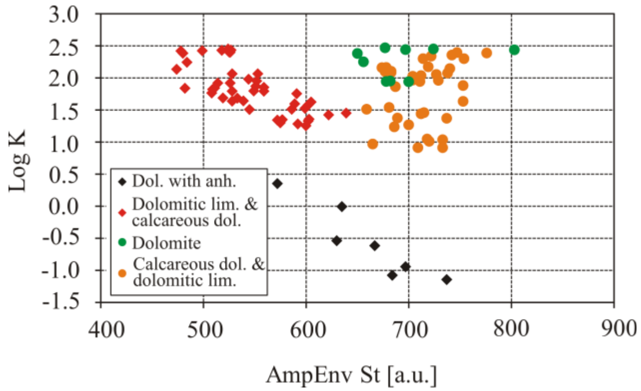


Fig. 19. Permeability of the Main Dolomite calculated from well logs with Zawisza formula (Zawisza 1993) *versus* amplitude envelope of Stoneley wave from AFW in R-3 well.

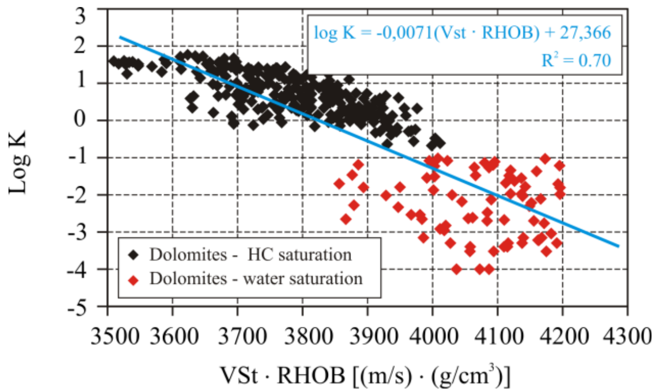


Fig. 20. Permeability of the Main Dolomite calculated from well logs with Zawisza formula (Zawisza 1993) *versus* product of Stoneley wave velocity and bulk density (from well logs) in R-4k well.

such results is related to thinly-bedded formation, consisted of sandstones characterized by higher values of velocity and bulk density, and shales that have lower values of these parameters. The technical reason is related to the limited vertical resolution of logs in comparison to point laboratory results.

Relationships between well log parameters and the results of laboratory measurements confirmed that tested formations were non-homogeneous (Figs. 13 and 14). Poor depth matching due to the defined vertical resolution of well logging devices is only partially responsible for the observed scattering the data. The main role is played by complicated geological structure of the investigated formations which influenced all the presented results and

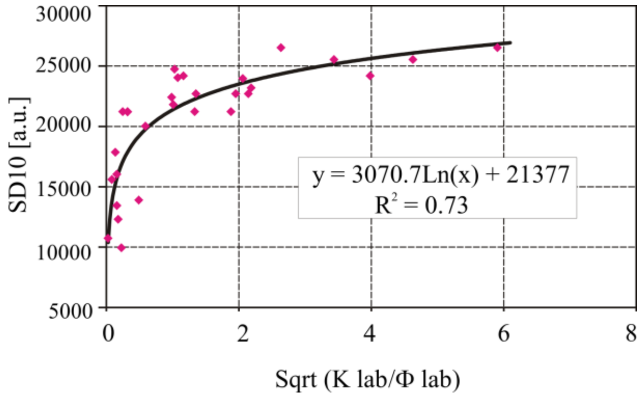


Fig. 21. Spectral decomposition 10 Hz component *versus* ratio of permeability and total porosity from laboratory for the Miocene formation, M-1 well.

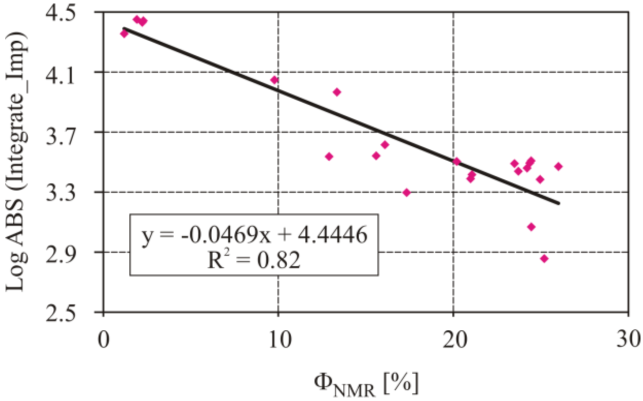


Fig. 22. Logarithm of absolute value of integrated impedance (seismic attribute) *versus* total porosity from NMR laboratory experiment for the Miocene formation, M-1 well.

significantly lowered the correlation coefficients. Despite the geological reason, relationships between determined parameters were established and can be used in further interpretation.

Relationships between log data represented by AFW attributes and seismic attributes are presented in Figs. 15 and 16. Satisfactory trends are observed between the instantaneous frequency from AFW (InstFreq P/InstFreq S) and seismic frequency represented by attributes such as an average frequency of impedance (AveFreq_Imp) and average frequency of SD30 component (AveFreq_SD30), where 30 Hz component was taken from the spectral decomposition of the seismic trace. Though coefficients of determination (R^2) are not high due to the scattered points (different rock types and non-homogeneity of Ca2 itself), the correlations are unquestionable. Similar situations can be observed in Figs. 17 and 18 where the Poisson's ratio is correlated with the instantaneous frequency obtained directly from seismic traces (InstFreq_Seis), and ratio of instantaneous frequencies of P and S waves from AFW (InstFreq P/InstFreq S). However, more detailed analyses show that the relation changes with facies. Figure 19 shows the diversity of the Main Dolomite Ca2 facies.

Stoneley wave sensitivity to permeability (K) is illustrated in Figs. 19 and 20. Increase of K generally reduces the amplitude of the Stoneley wave (here expressed as an Amplitude Envelope – one of the instantaneous attributes from AFW). An interesting relationship was obtained for velocity of Stoneley wave multiplied by the bulk density of the formation. R^2 coefficient is high (0.7); thus, the relation is promising for permeability prediction. Stoneley wave identified from AFW in combination with the logarithm of permeability (Fig. 20) showed the ability to discriminate between water saturated and hydrocarbon saturated parts of the Main Dolomite. Relationships between the parameter related to hydraulic properties of the Sarmatian sandy-shaly reservoir (M-1 well), here expressed as a square root of permeability to porosity ratio, Sqrt (K/Φ), and component of spectral decomposition SD10 of seismic trace (SD10) were a very interesting example of the combination between seismic attribute and laboratory results (Fig. 21). A strong relationship was also observed between the logarithm of the absolute value of integrated impedance as seismic attribute (Log ABS (Integrate_Imp)) and total porosity from an NMR laboratory experiment (Fig. 22). The goal of these considerations was to check whether it was possible to determine the permeability from the seismic data.

The presented relationships between petrophysical parameters from laboratory measurements, attributes from acoustic full waveforms and seismic attributes are a step further in dealing with scaling problem. Results of various geophysical measurements are different, even when based on the same rock parameters, such as elastic and reservoir properties. The found relationships show connection between parameters of a different scale: from lab to log scale, from log to seismic scale, and even from lab to seismic scale. Relationships indicate that it is possible to extrapolate laboratory data, very detailed but measured on small samples cut from cores, to a larger amount of rock formation considered by well logs and seismics. Including acoustic full waveforms in such a research, particularly AFW attributes, substantially improves combining the parameters acquired at different scales. On the one hand, the AWF attributes, which represent log scale, can be related to lab and other log data; on the other hand, they are a natural link between log and seismic data thanks to the same physical background (*i.e.*, elastic wave field). Additionally, seismic attributes that were incorporated instead of raw seismic data increased sensitivity for local and tiny changes of rock formation, which are usually averaged in seismic surveys. These results proved the key role of well logs in combination of different types of data; however, the links are usually not obvious. Applying seismic and AFW attributes along with ratio of different rock properties can be helpful (Wawrzyniak-Guz and Jarzyna 2014).

9. THREE DIMENSIONAL RESULTS VISUALIZATION

3D images of seismic sections were constructed to the show position of wells, depth of wells and horizons important in the lithostratigraphic correlation. The variability of seismic attributes along the borehole axis is well visible on the background of seismic sections. Two examples illustrate relationships between seismic results, well logs and laboratory experiments outcomes (Figs. 23 and 24).

The presented 3D visualizations enable to show similarity of parameters determined from various methods. Such presentation distinctly shows scale differences. Visualization of the petrophysical results on the background of seismic section, reflecting the variability of lithology and stratigraphy, delivers additional global scale information.

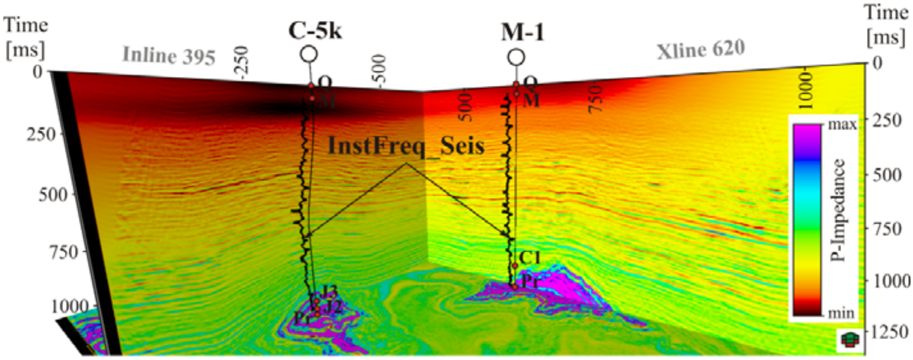


Fig. 23. 3D visualization of instantaneous frequency calculated from the standard seismic processing (InstFreq_Seis) on the background of acoustic impedance seismic sections, TCZ 3D seismic project.

10. CONCLUSION

The main concept of establishing relationships between various parameters obtained in the processing and interpretation of seismic data and well logs, including also the results of laboratory experiments, was successfully realized. Pseudoimpedance sections from seismic inversion along with various

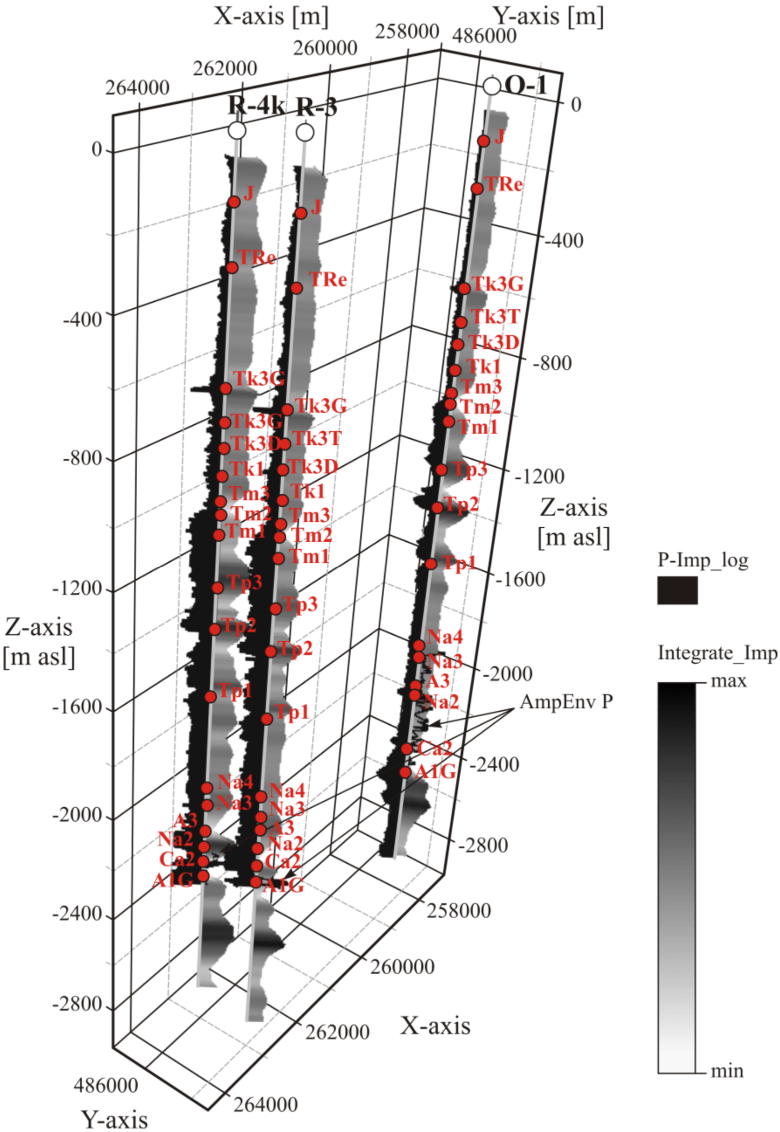


Fig. 24. 3D visualization of P wave instantaneous amplitude (AmpEnv P, right track), integral of acoustic impedance from seismic inversion (Integrate_Imp, right track, gradient of colour, from white to black) and P wave acoustic impedance from well logs (P-Imp_log, left track, black colour), OR 3D seismic project.

seismic attributes and those of calculated from acoustic full waveforms were used in many combinations with velocity and bulk density from well logs or velocity, porosity and permeability from laboratory measurements.

In two different geological formations (sandy-shaly thin-bedded Sarmatian formation and the Main Dolomite carbonate horizon), relationships of high correlation coefficients between seismic attributes, well logs, and laboratory origin values of porosity and permeability were determined. Seismic attributes and AFW attributes as parameters depending on small characteristic features of rock formation revealed good correlation with laboratory results. The combination of these parameters was considered as a type of scaling data that were acquired from methods of different vertical resolution.

Non-homogeneity of the formations was pointed out as a factor lowering the presented relationships.

Acknowledgments. Research were done in the frame of the scientific project No. NN 307 294439 (2010-2013) financed by Ministry of Science and Higher Education, Poland, and was also financially supported by statutory funds No. 11.11.140.769 at AGH University of Science and Technology, Faculty of Geology, Geophysics and Environmental Protection, Kraków, Poland. Authors thank for the data to PO&GC, Warsaw, Poland. Seismic survey and interpretation, and well logs measurements and interpretation were done by Geofizyka Kraków, Poland. Analyses were done in Techlog software (Schlumberger) and Hampson-Russell software (CGGVeritas) thanks to Academic Grants funded by these companies to the AGH University of Science and Technology, Kraków, Poland. Thanks are also directed to Magdalena Niepsuj, M.Sc. Eng., Ph.D. student at the AGH University of Science and Technology, Kraków, Poland, for the seismic inversion results, to Michał Gruszczyk, M.Sc. Eng., from Geofizyka Kraków, Poland for seismic attributes calculations, and to Teresa Staszowska for help in preparation of the figures.

References

Bała, M., and J. Jarzyna (1996), Application of acoustic full wavetrains for the determination of lithology, reservoir and mechanical parameters of formation, *Geophys. Prospect.* **44**, 5, 761-787, DOI: 10.1111/j.1365-2478.1996.tb00173.x.

Bała, M., J. Jarzyna, and Z. Mortimer (2012), Statistical analysis of petrophysical parameters of Middle Miocene rocks from the Polish Carpathian Foredeep, *Geol. Q.* **56**, 4, 665-680, DOI: 10.7306/gq.1048.

Bui, H., P. Ng, D. Becker, J. Durrani, and M. Smith (2010), Well-seismic Tie in the Green Canyon, deep-water area in the Gulf of Mexico – A valuable indicator of anisotropy. **In:** *Ext. Abstr. 72nd EAGE Conf. & Exhib. Incorporating SPE EUROPEC, 14-17 June 2010, Barcelona, Spain*, P251, DOI: 10.3997/2214-4609.201401154.

Chopra, S., and K.J. Marfurt (2005), Seismic attributes – A historical perspective, *Geophysics* **70**, 5, 3SO-28SO, DOI: 10.1190/1.2098670.

ElGhonimy, R.S., and S. Sonnenberg (2015), Statistical methods of predicting source rock organic richness from open-hole logs, Niobrara Formation, Denver Basin, CO. **In:** *Unconventional Resources Technology Conf., 20-22 July 2015, San Antonio, Texas, USA*, SPE-178487-MS, DOI: 10.2118/178487-MS.

Huuse, M., and D.A. Feary (2005), Seismic inversion for acoustic impedance and porosity of Cenozoic cool-water carbonates on the upper continental slope of the Great Australian Bight, *Mar. Geol.* **215**, 3-4, 123-134, DOI: 10.1016/j.margeo.2004.12.005.

Jarzyna, J., M. Bała, Z. Mortimer, and E. Puskarczyk (2013), Reservoir parameter classification of a Miocene formation using a fractal approach to well logging, porosimetry and nuclear magnetic resonance, *Geophys. Prospect.* **61**, 5, 1006-1021, DOI: 10.1111/j.1365-2478.2012.01102.x.

Knize, S. (1989), Evaluation of full waveform sonic data by analysis of instantaneous characteristics and colograms. **In:** *Trans. 12th Int. Logging Symposium of SAID, 24-27 October, Paris, France*, SAID-004.

Krakowska, I.P., and E. Puskarczyk (2015), Tight reservoir properties derived by nuclear magnetic resonance, mercury porosimetry and computed microtomography laboratory techniques. Case study of palaeozoic clastic rocks, *Acta Geophys.* **63**, 3, 789-814, DOI: 10.1515/acgeo-2015-0013.

Liner, C.L., and T.W. Fei (2006), Layer-induced seismic anisotropy from full-wave acoustic sonic logs: Theory, application, and validation, *Geophysics* **71**, 6, D183-D190, DOI: 10.1190/1.2356997.

Liner, C.L., and T.W. Fei (2007), The Backus number, *The Leading Edge* **26**, 4, 420-426, DOI: 10.1190/1.2723204.

Marzec, P., and K. Pietsch (2012), Thin-bedded strata and tuning effect as causes of seismic data anomalies in the top part of the Cenomanian sandstone in the Grobla-Rajsko-Rylowa area (Carpathian foreland, Poland), *Geol. Q.* **56**, 4, 691-710, DOI: 10.7306/gq.1050.

Marzec, P., M. Niepsuj, M. Bała, and K. Pietsch (2014), The application of well logging and seismic modelling to assess the degree of gas saturation in miocene strata (Carpathian Foredeep, Poland), *Acta Geophys.* **62**, 1, 83-115, DOI: 10.2478/s11600-013-0177-2.

Pechnig, R., S. Haverkamp, J. Wohlenberg, G. Zimmermann, and H. Burkhardt (1997), Integrated log interpretation of the German Continental Deep Drilling Program: Lithology, porosity and fracture zones, *J. Geophys. Res.* **102**, B8, 18363-18390, DOI: 10.1029/96JB03802.

Pietsch, K., P. Marzec, M. Kobylarski, T. Danek, A. Leśniak, A. Tatarata, and E. Gruszczyk (2007), Identification of seismic anomalies caused by gas saturation on the basis of theoretical P and PS wavefield in the Carpathian Foredeep, SE Poland, *Acta Geophys.* **55**, 2, 191-208, DOI: 10.2478/s11600-007-0002-x.

Puskarczyk, E., J. Jarzyna, and S. J. Porębski (2015), Application of multivariate statistical methods for characterizing heterolithic reservoirs based on wireline logs – example from the Carpathian Foredeep Basin (Middle Miocene, SE Poland), *Geol. Q.* **59**, 1, 157-168, DOI: 10.7306/gq.1202.

Ronen, S., P.S. Schultz, M. Hattori, and C. Corbett (1994), Seismic-guided estimation of log properties (Part 2. Using artificial neural networks for nonlinear attribute calibration), *The Leading Edge* **13**, 6, 674-678, DOI: 10.1190/1.1437027.

Schultz, P.S., S. Ronen, M. Hattori, and C. Corbett (1994a), Seismic-guided estimation of log properties (Part 1. A data-driven interpretation methodology), *The Leading Edge* **13**, 5, 305-310, DOI: 10.1190/1.1437020.

Schultz, P.S., S. Ronen, M. Hattori, M. Mantran, and C. Corbett (1994b), Seismic-guided estimation of log properties (Part 3. A controlled study), *The Leading Edge* **13**, 7, 770-776, DOI: 10.1190/1.1437036.

Szabó, N.P. (2011), Shale volume estimation based on the factor analysis of well-logging data, *Acta Geophys*, **59**, 5, 935-953, DOI: 10.2478/s11600-011-0034-0.

Taner, M.T., F. Koehler, and R.E. Sheriff (1979), Complex seismic trace analysis, *Geophysics* **44**, 6, 1041-1063, DOI: 10.1190/1.1440994.

Veeken, P.C.H., and M. Da Silva (2004), Seismic inversion methods and some of their constraints, *First Break* **22**, 6, 47-70, DOI: 10.3997/1365-2397.2004011.

Wawrzyniak-Guz, K., and M. Gruszczyk (2013), Combination of seismic and well log data with the use of attributes and acoustic full waveforms. **In:** *Ext. Abstr. 75th EAGE Conf. & Exhib. Incorporating SPE EUROPEC 2013, 10-13 June 2013, London, United Kingdom*, We P15 03, DOI: 10.3997/2214-4609.20131047.

Wawrzyniak-Guz, K., and J. Jarzyna (2012), Seismic attributes for acoustic full waveforms. **In:** *Ext. Abstr. 74th EAGE Conf. & Exhib. Incorporating EUROPEC, 4-7 June 2012, Copenhagen, Denmark,* P072, DOI: 10.3997/2214-4609.20148475.

Wawrzyniak-Guz, K., and J. Jarzyna (2014), Combination of rock formation properties derived from lab, log and seismic measurements scale. **In:** *Ext. Abstr. 76th EAGE Conf. & Exhib., 16-19 June 2014, Amsterdam, Netherlands*, We G104 02, DOI: 10.3997/2214-4609.20141168.

Wawrzyniak-Guz, K., E. Puskarczyk, P.I. Krakowska, and J.A. Jarzyna (2016), Classification of Polish shale gas formations from Baltic Basin, Poland based on well logging data by statistical methods. **In:** *Conf. Proc. SGEM*

2016 16th Int. Multidisciplinary Scientific Geoconf.: Science and Technologies in Geology, Exploration and Mining, 30 June – 6 July 2016, Albena, Bulgaria, Vol. 3, 761-768.

Wenlong, D., L. Chao, L. Chunyan, X. Chungchun, J. Kai, Z. Weite, and W. Liming (2012), Fracture development in shale and its relationship to gas accumulation, *Geosci. Front.* **3**, 1, 97-105, DOI: 10.1016/j.gsf.2011.10.001.

Zawisza, L. (1993), Simplified method of absolute permeability estimation of porous beds, *Arch. Min. Sci.* **38**, 4, 343-352.

Zoback, M.D. (2010), *Reservoir Geomechanics*, Cambridge University Press, Cambridge, DOI: 10.1017/CBO9780511586477.

Received 30 July 2016
Received in revised form 14 October 2016
Accepted 3 November 2016

DE GRUYTER
OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2382-2409
DOI: 10.1515/acgeo-2016-0111

The New Algorithm for Fast Probabilistic Hypocenter Locations

Wojciech DĘBSKI and Piotr KLEJMENT

Institute of Geophysics, Polish Academy of Sciences, Warsaw, Poland

Abstract

The spatial location of sources of seismic waves is one of the first tasks when transient waves from natural (uncontrolled) sources are analysed in many branches of physics, including seismology, oceanology, to name a few. It is well recognised that there is no single universal location algorithm which performs equally well in all situations. Source activity and its spatial variability in time, the geometry of recording network, the complexity and heterogeneity of wave velocity distribution are all factors influencing the performance of location algorithms. In this paper we propose a new location algorithm which exploits the reciprocity and time-inverse invariance property of the wave equation. Basing on these symmetries and using a modern finite-difference-type eikonal solver, we have developed a new very fast algorithm performing the full probabilistic (Bayesian) source location. We illustrate an efficiency of the algorithm performing an advanced error analysis for 1647 seismic events from the Rudna copper mine operating in southwestern Poland.

Key words: hypocenter location, probabilistic inverse theory, error analysis, time reversal mirroring, numerical methods.

1. INTRODUCTION

Determining the spatial location (hypocenter) and origin time of the source of seismic waves is one of the first tasks undertaken when the waves from uncontrolled sources are analyzed (Aki and Richards 1985, Gibowicz and Kijko 1994). Depending on the type of event, different location techniques are used. If recorded waveforms for the event at hand exhibit well defined and sharp waves onsets, the event is usually located by seeking a point in space for which numerically predicted wave onset times fit the observed ones (Aki and Richards 1985, Bulland 1976, Thurber and Rabinowitz 2000). This can be accomplished by any optimization technique (Aki and Richards 1985, Thurber and Rabinowitz 2000). The price for simplicity and numerical efficiency of such approach is, however, a lack of reliable estimation of the location errors. Within this optimization-based approach, the most popular method of error evaluation is a linearization of the forward modelling procedure around the optimum location found and calculation of the covariance matrix (Husen and Hardebeck 2010, Menke 1989, Pavlis 1992). The diagonal elements are then interpreted as squares of the location errors for each coordinate, respectively, and the "horizontal part" of the covariance matrix gives a nice looking "error ellipse" confidence region - a region around the solution found where the epicenter is supposed to be located within obtained errors. However, as it has been pointed out, for example, by Bai *et al.* (2006), Husen *et al.* (1999), Lomax *et al.* (2001), Wiejacz and Debski (2001), in realistic situations this estimation of location errors is too simplified and often unrealistic. The reason is that the cornerstone of the method linearization of the forward modelling relation leads to a quadratic misfit function which is a good choice only when all uncertainties pertaining location procedure (data errors, modelling errors) are Gaussian. This is often not the case (Husen and Hardebeck 2010, Lomax *et al.* 2009, Rudzinski and Debski 2012).

To evaluate the location errors in a more systematic way, a probabilistic, also called Bayesian, inversion technique must be used (Tarantola 2005). The approach relies on exploring the space of all possible source locations and assigning to each point in this space (*i.e.,* each possible location) a probability of being the true hypocenter location (Debski 2010, Sambridge and Mosegaard 2002, Tarantola 2005). This *a posteriori* probability can then be used for any analysis of location errors using standard statistical methods (Debski 2010, Lehmann and Casella 1998).

For completness it is worth mentioning the third method of error estimation which is based on the direct Monte Carlo simulation (Giovambattista and Barba 1997, Husen and Hardebeck 2010). The method relies on adding the Monte Carlo generated "noise" to the synthetic data calculated for the formerly found optimum location and relocating the event using these "noisy" data. The size of a "cloud" of the obtained solutions is used as an estimator of the location errors. However, as pointed out by Debski (2004), this approach reduces to the above-mentioned probabilistic approach if the "noise" is generated according to the probability distribution of joint modeling and observational errors.

The probabilistic approach to hypocenter location provides the most complete information about the solution found (see, *e.g.*, Lomax *et al.* 2009, Tarantola 2005). However, the method is computationally demanding as it requires an exhaustive exploration of the 3 or 4-dimensional model space. Even the use of modern, efficient Markov Chain Monte Carlo based methods (Chib and Greenberg 1995, Debski 2010, Gilks *et al.* 1995, Lomax *et al.* 2009) does not allow to employ this approach in applications requiring prompt results. The new perspective for using probabilistic approach are opened by using the time reversal and reciprocity principles of the wave equation together with finite-difference type forward modelling solvers. In this case, as we shall discuss latter on, one can avoid an explicit sampling of the *a posteriori* distribution – the most time consuming part of the probabilistic inversion requiring tens of thousends forward modellings by performing an “implicit” sampling through an examination of the *a posteriori* distribution over the finite grid used by the forward modelling solver. This reduces a number of required forward modelings from thousands to only a few – a number of used seismic stations. In this paper we descibe such an algorithm. We call it TRMLOC.

The paper is structured as follow. Firstly, after a short description of the source location task the proposed algorithm is described. Next, two basic elements of the algorithm, namely probabilistic inversion technique and the FSM algorithm are discussed. Next, the algorithm is compared to other popular methods and its performance is illustrated by analyzing 1647 mining induced seismic events from Rudna (Poland) deep copper mine. The conclusions and general discussion ends up the paper.

2. LOCATION ALGORITHMS

Let us assume that to locate a given source the arrival times $\mathbf{t}_i^{obs}$ $i = 1 \cdots N_s$ are read from waveforms recorded by N_s sensors (geophones, seismometers, piezoceramic transducers, *etc.*). Let $\mathbf{t}_i^{th}(\mathbf{m})$ denote the theoretically predicted arrival time of waves originating at the point $\mathbf{m}$ and recorded by i^{th} sensor where the location parameters $\mathbf{m} = (\vec{r}, t_o)$ include three spatial coordinates ($\vec{r} = (x, y, z)$) and the origin time of the event t_o , so

$$\mathbf{t}_i^{th}(\mathbf{m}) = t_o + \Delta_i(\vec{r}) \tag{1}$$

where $\Delta_i(\vec{r})$ is the propagation time from the source to a given sensor.

Finding the hypocenter location can now be formulated as the optimization task (Aki and Richards 1985): searching for the model $\mathbf{m}^{ml}$ which minimizes the differences between observed ($\mathbf{t}_i^{obs}$) and predicted $(\mathbf{t}_i^{th}(\mathbf{m}))$ travel times. The solution can be obtained by means of any convenient numerical optimization algorithm.

To complete the location task, an error analysis should be performed in order to evaluate the reliability of the solution. The most popular approach to this task is based on the linearizion of the optimized function $S(\mathbf{m})$ around the solution $\mathbf{m}^{ml}$ and calculation of the covariance matrix (Gibowicz and Kijko 1994). However, this simple method is often not reliable. It fails if observational and/or modeling errors have nontrivial statistics, the recording network geometry is complicated or the velocity model has a complex structure (Lomax 2005, Lomax *et al.* 2009, Rudzinski and Debski 2012).

Another, probabilistic approach to the source location task relies on assigning to each model $\mathbf{m}$ (point in space and time) the *a posteriori* probability of $\mathbf{m}$ being the true source location (Debski 2010, Lomax *et al.* 2009, 2000). The advantage of this approach is the possibility of full and exhaustive error and resolution analysis (Debski 2010). In the simplest case, the *a posteriori* probability density $\sigma(\mathbf{m})$ assigned to model $\mathbf{m}$ reads (Debski 2010, Mosegaard and Tarantola 2002)

$$\sigma(\mathbf{m}) = \frac{1}{Z} f(\mathbf{m})\, L(\mathbf{m}) \tag{2}$$

where Z is the normalization factor called *evidence*, $f(\mathbf{m})$ is the probability density function describing the *a priori* estimation of the source location. The second term, traditionally called the likelihood function, is defined as follows

$$L(\mathbf{m}) = \exp\left(-S(\mathbf{m})\right), \tag{3}$$

where

$$S(\mathbf{m}) = ||\mathbf{t}_i^{th}(\mathbf{m}) - \mathbf{t}_i^{obs}|| \tag{4}$$

is the so-called misfit function and $||\cdot||$ is a norm in the data space. The choice of a given norm (l_1, l_2, Cauchy, *etc.*) reflects our expectations about errors statistics, existence of outliers, systematic bias, etc (Debski 2010).

Various numerical estimators, like the maximum likelihood model ($\mathbf{m}^{ml}$) which maximizes $\sigma(\mathbf{m})$, the average model ($\mathbf{m}^{avr}$), the covariance matrix, *etc.* can easily be calculated from $\sigma(\mathbf{m})$. The technique is very general but it requires exhaustive sampling of the model space to determine the characteristics of $\sigma(\mathbf{m})$. Consequently, the approach is computationally demanding even if the very efficient Markov Chain Monte Carlo sampling technique (Debski 2010, Gilks *et al.* 1995, Lomax *et al.* 2009) is used.

The new possibilities of the full probabilistic seismic data inversion for hypocenter coordinates open when the time reversal mirroring technique is employed. This technique, carefully analyzed in laboratory experiments (see, *e.g.*, Fink 1997, Ulrich *et al.* 2008), by numerical simulations (see, *e.g.*, Kremers *et al.* 2011, Scalerandi *et al.* 2009, Steiner and Saenger 2012), and theoretical investigations (see, *e.g.*, Masson *et al.* 2014, Tromp *et al.* 2005, Ulrich *et al.* 2009) has already found application in seismic prospecting (see, *e.g.*, Gajewski and Tessmer 2010, Witten and Artman 2011) and location of seismic tremors (see, *e.g.*, Artman *et al.* 2010, Larmat *et al.* 2008, O'Brien *et al.* 2011). Combaining this technique with a modern eikonal solver has lead us to the proposition of the new TRMLOC location algorithm which enables very fast Bayesian inversion of travel time onset data for hypocenter location.

3. TRMLOC ALGORITHM

The wave equation (Aki and Richards 1985) describing the propagation of seismic waves exhibits two very important features. First of all, being a second-order partial differential equation with even time derivatives, the equation is invariant under time reversal. Thus, if only the boundary conditions do not depend on time, the solutions for forward and back propagation of waves in time are identical. Secondly, the equation exhibits spatial reciprocity invariance, which means that wave propagation between two arbitrary points is invariant with respect to the exchange of these points: the seismogram from the source located at point A and recorded at point B is the same as the seismogram from the source set at B and recorded at A. Combining both properties of the wave equation has enabled construction of a very simple and efficient numerical algorithm used for an analysis presented in this paper. It relies on putting the virtual sources at the receiver locations (reciprocity principle) and simulating propagation of seismic waves from such virtual sources adopting the recorded real signals with reversed time as their temporal signature: the last arriving signal is "re-sent" as the first one. Due to complicated interference, the re-sent signals not only focus in the point where the real source ruptured and at the origin time of the rupture but also provide very important information on the kinematics and dynamics of the rupture process (Fink *et al.* 2000, Fink and Tanter 2010, Kremers *et al.* 2011).

The existing application of the above-described time reversal technique for hypocenter location consists of two steps, namely back propagation ("re-sending") of recorded signals and scanning of the model (location) space in order to find the optimum location where all the back-propagated signals positively interfere (O'Brien *et al.* 2011). The algorithm used in this paper extends the above idea in two aspects. First of all, instead of searching for a point at which the largest positive interference occurs, we propose to construct the *a posteriori* probability density based on the differences of the back-propagated observational data. Secondly, for the location of sources with well determined time onsets on a given set of sensors, we do not need to perform a full waveform back propagation. Instead, we can consider only the wavefronts whose propagation in time is described by the much simpler to solve eikonal equation (Aki and Richards 1985) for which modern fast algorithms like the Fast Sweeping Method (Zhao 2005), the Fast Marching Method (Sethian 1999) or the more traditional finite difference approaches (Podvin and Lacomte 1991, Vidale 1990) can be used.

We propose to construct the *a posteriori* probability density using the differential misfit function $\bar{S}(\mathbf{m})$

$$\bar{S}(\mathbf{m}) = \frac{1}{2N_s} \sum_{i,k;i\neq k} ||\mathbf{t}_m^i - \mathbf{t}_m^k|| \tag{5}$$

based on the difference of the back propagated wavefronts from all considered sensors. In this equation, $\mathbf{t}_m^i$ and $\mathbf{t}_m^k$ stand for waveform onsets recorded by *i*-th and *k*-th receivers, respectively, and back-propagated to the point $\mathbf{m}$. N_s is the number of the receivers used (number of available observational data) and 1/2 takes into account the symmetry of the sum. Following this assumption, we postulate the *a posteriori* probability distribution as

$$\sigma(\mathbf{m}) = \mathrm{const.} f(\mathbf{m}) \exp\left(-\bar{S}(\mathbf{m})\right) \tag{6}$$

The physical intuition behind the above definition of $S(\mathbf{m})$ is quite clear. In an ideal case (no noise, exact forward modelling) all back-propagated travel times should be equal to the source origin time (t_o) at the true hypocenter location point. Thus, the condition for the hypocenter location is the equality of all back-propagated arrival times. Due to the presences of observational and modeling errors, this condition cannot in general be directly fulfilled and thus a reasonable solution is to look for the point in space where $S(\mathbf{m})$ gets minimum. Let us note that, as follows from Eq. 1, the origin time t_o does not enter $S(\mathbf{m})$ and thus an original 4D inversion is reduced to 3D problem: a search for the hypocenter's spatial coordinates only.

The idea of using the differential-time form of the misfit function is by no means new and can be traced back in time to Zhou (1994) and Matsu'ura (1984). In various forms it has already been implemented in different optimization-based location algorithms under various names, among which the equal difference time (EDT) is the most popular (see, *e.g.*, Font *et al.* 2004, Lomax *et al.* 2009, Zhou 1994). The EDT formulation relies on searching the point $\mathbf{m}$ for which hyperbolic surfaces defined by the condition

$$\Delta_i(\mathbf{m}) - \Delta_j(\mathbf{m}) = \mathbf{t}_i^{obs} - \mathbf{t}_j^{obs} \tag{7}$$

intersect for all pairs of stations (i, j). This condition can be rewritten as $\mathbf{t}_i^{obs} - \Delta_i(\mathbf{m}) = \mathbf{t}_j^{obs} - \Delta_j(\mathbf{m})$ for all (i, j), which is actually the condition of equality of all back-propagated observational time onsets at the hypocenter location. The advantage of using the EDT-type differential misfit function relies in removing of origin time from inversion procedure (Matsu'ura 1984) and also lower sensitivity of location results to velocity model (Font *et al.* 2004, Rudzinski and

Debski 2011, Waldhauser and Ellsworth 2000, Zhou 1994). Additionally, in the developed algorithm, it has also allowed to perform an implicit sampling of the *a posteriori* distribution, as discussed latter on. The EDT type misfit function is the cornerstone of the modern relative location methods, namely the double differences and extended double differences techniques (Rudzinski and Debski 2012, Waldhauser and Ellsworth 2000).

Finding the minimum of $S(\mathbf{m})$ will provide the hypocenter location, so $S(\mathbf{m})$ can serve as the cost function for any optimization-based location algorithm. Much less obvious is whether this misfit function can also be used within the probabilistic inversion framework for generating the likelihood function according to Eq. 6. The problem is that the "true" likelihood function $L(\mathbf{m})$ defined by the probabilistic inverse theory is actually a convolution of probability distributions of observational and modelling errors (Debski 2010, Tarantola 2005). Thus, from the statistical point of view, it describes the statistic of sum of errors. Apparently the function $L(\mathbf{m})$ defined by $\bar{S}(\mathbf{m})$ is not such a statistic. It is rather the statistic of sum of differential errors so the question is if the errors estimated by using this proxy of the likelihood function are not systematically biased. Although this point has not been clarified yet, the differential misfit function has already been implemented in some probabilistic location algorithms (Lomax *et al.* 2000, Rudzinski and Debski 2011) and we use it also in the TRMLOC algorithm.

Having defined the *a posteriori* distribution $\sigma(\mathbf{m})$ we have to explore the space of model parameters in order to obtain various characteristics of $\sigma(\mathbf{m})$ including the position of the global maximum, checking an existence of secondary maxima, *etc*. This is the most demanding numerical part of any probabilistic inversion. However, in case of the location task the model space which has to be sampled is exactly the same space (3D configuration space) as that over which the forward modelling operator acts. This opens a possibility of performing an implicit sampling of the *a posteriori* distribution simultaneously with solving the forward problem. The idea is as follows. Assume that the forward modeling method used to calculate $\Delta_i(\mathbf{m})$ is able to provide the back-in-time propagated observed time onsets for a set (for example, regular grided) of spatial points. Then, according to Eq. 5, the *a posteriori* distribution $\sigma(\mathbf{m})$ can be immediately calculated with minimum numerical computations for all grid points. This way we have sampled $\sigma(\mathbf{m})$ at all these points. If the points form a dense enough, regular set, we end up with the well sampled $\sigma(\mathbf{m})$ so we do not need any additional sampling indispensable in the classical probabilistic inversion. We call this mechanism the implicit sampling. The forward modeling techniques fulfilling the above requirement are the all wave equation or eikonal solvers based on the finite difference, finite element, spectral elements, or similar numerical methods (Sethian 1999, Virieux *et al.* 2009). Thus, summarizing the

- Discretize space $\mathbf{m} = (X_i, Y_j, Z_k), \quad i, j, k = 1, 2, \ldots$
- Set the *a priori* density function $f(\mathbf{m})$
- Repeat for each receiver (in parallel)
  - back propagate observed time onsets $\mathbf{t}_i^{obs}$ using eikonal (FSM) solver on defined spatial grid
- Calculate $S(\mathbf{m}) = \frac{1}{2N_s} \sum\limits_{i,k;i \neq k} (\mathbf{t}_m^i - \mathbf{t}_m^k)^2 / C_p^2$
- Determine $\sigma(\mathbf{m}) = \mathrm{const.} f(\mathbf{m}) \exp(-S(\mathbf{m}))$
- Calculate statistical estimators $\mathbf{m}^{ml}$, $\mathbf{m}^{avr}$, $\Delta\mathbf{m}$, evidence, entropy, etc.
- If needed, perform inspection of the full $\sigma(\mathbf{m})$ or marginal *a posteriori* distributions

Fig. 1. The basic steps of the TRMLOC algorithm.

above consideration we propose the algorithm whose flowchart is shown in Fig. 1.

One very important feature of the algorithm is its high speed, as will be demonstrated later on, which follows from:

- reducing inverse problem from 4D to 3D by eliminating event's origin time from inversion,
- employing the modern finite-difference very fast eikonal solver,
- avoiding explicit sampling of the model space: $\sigma(\mathbf{m})$ is evaluated at each grid nodes simultaneously with forward modelings,
- parallelization of the algorithm.

Let us also note that the backward propagation of the observed time onsets through the back-in-time forward modeling has to be performed only N_s times - as many as the number of sensors is used. This is a direct advantage of using the time-reversal and reciprocity invariance principle.

Actually, the TRMLOC algorithm is very similar to the algorithm NLloc developed by Lomax *et al.* (2000). Both approaches use the probabilistic inverse approach, eikonal solvers for forward modelling and similar EDT-based likelihood function. The main differences arise from using different eikonal solvers (NLloc uses the method of Podvin and Lacomte (1991) while TRMLOC the Fast Sweeping Method) and from different implementation of the *a posteriori* pdf sampler.

The very important element of the TRMLOC algorithm is the eikonal solver which enables very efficient calculation of the wavefront positions in the entire 3D domain for a general velocity model. Constructing the TRMLOC algorithm, two finite-difference type eikonal solvers were considered, namely the Fast Marching Method (FMA) developed by Sethian (1999) and the Fast Sweeping Method (FSM) developed by Zhao (2005). The FMA algorithm exhibits numerical complexity of the order of $N\log(N)$, where N is the number of all grid nodes and is optimal for complex velocity models (Sethian 1999). The FSM method is faster for smooth velocity models with numerical complexity proportional to N but it is over-performed by FMA in cases of complex velocity models. TRMLOC has been designed for a local/regional analysis when velocity models are relatively smooth so the FSM technique has been selected. Since this algorithm is practically unknown to seismological community we give here its detailed description restricting ourselves to 2D case for the sake of compactness.

3.1 Eikonal solver: Fast Sweeping Method

Under the high frequency approximation the full wave equation can be split into the eikonal equation describing spatial propagation of wavefronts and transport equation describing changes of wave amplitudes. As we are interested here in travel times modeling, we consider only the eikonal equation which together with the boundary condition at source location Γ reads

$$\begin{aligned} &\nabla T \cdot \nabla T = \frac{1}{v^2} \\ &T|_\Gamma = 0 \end{aligned} \tag{8}$$

where T describes wavefront position in space originating from the source at Γ and v denotes velocity. This is a special case of the Hamiltonian-Jacobi, hyperbolic type nonlinear equation for which the term on right-hand side is always positive. For numerical purpose, such equation can be discretized by using the first-order Godunov upwind type discretization (Sethian 1999, Zhao 2005), For internal grid points this finite difference scheme leads to the following discrete approximation of Eq. 8

$$\left[(T_{i,j} - T_{xmin})^+\right]^2 + \left[(T_{i,j} - T_{ymin})^+\right]^2 = h^2 s_{ij}^2 \tag{9}$$

where $i,\ j$ are indexes of the grid point $\mathbf{x}_{i,j} = (x_i, y_j)$, h is the grid size (for simplicity, the quadratic grid is assumed), s_{ij} is the value of slowness at grid point $(\mathbf{x}_{i,j})$ ($s_{ij} = 1/v_{ij}$) , and the following shorthand notation is used:

$$T_{xmin} = \min(T_{i-1,j}, T_{i+1,j}), \qquad T_{ymin} = \min(T_{i,j-1}, T_{i,j+1}) \tag{10}$$

and

$$(x)^+ = \begin{cases} x,\, x > 0 \\ 0,\, x \le 0 \end{cases} \tag{11}$$

The Fast Sweeping Algorithm proposed by Zhao (2005) is using the above discretization and solves the resulting system of nonlinear equation iteratively as follows:

- Initialization: a large positive value is assigned to all $T_{i,j}$. Then, for all grid points $(\mathbf{x}^s_{i,j})$ within the source of waves (it can be a single grid node for a point-like seismic source model or an extended area Γ for the finite source model) the boundary condition $T(\mathbf{x}^s) = 0$ is set.
- Iterations with alternating sweeping: the following procedure is repeated until the conversion to stable solution is reached.
  - At each grid point $(\mathbf{x}_{i,j})$ not set during the initializations the solution $\tilde{T}$ is computed using current values of T at neighborhood points and then $T_{i,j}$ is updated us follows

$$T^{new}_{i,j} = \min\left(T^{cur}_{i,j}, \tilde{T}\right) \tag{12}$$

where the updating solution $\tilde{T}$ is the solution of Eq. 9 and reads

$$\tilde{T}_{i,j} = \begin{cases} \min(T_{xmin}, T_{ymin}) + s_{ij}h & |T_{xmin} - T_{ymin}| \geq s_{ij}h, \\ \frac{T_{xmin}+T_{ymin}+\sqrt{2s_{ij}^2h^2-(T_{xmin}-T_{ymin})^2}}{2} & |T_{xmin} - T_{ymin}| < s_{ij}h, \end{cases} \tag{13}$$

  - During one iteration the value of $\tilde{T}$ is recalculated four times with different alternating orderings of grid sweeping:

$$\begin{array}{l} \text{a) } i = 1 : N_x,\ \ j = 1 : N_y \\ \text{b) } i = 1 : N_x,\ \ j = N_y : 1 \\ \text{c) } i = N_x : 1,\ \ j = 1 : N_y \\ \text{d) } i = N_x : 1,\ \ j = N_y : 1 \end{array} . \tag{14}$$

As follows from the above description, the proposed scheme shows the numerical complexity of order $O(kN)$ where N is the total number of the grid nodes and k is a constant depending on the number of iterations.

The number of iterations to be performed depends on the complexity of the velocity model. In many cases, if the velocity model is reasonably smooth and without large velocity contrasts only a few (usually 2-3) iterations are sufficient for convergence of the algorithm. The reason is that each sweep (Eq. 14) provides the exact solution in one iteration for one spatial quarter, provided the characteristics of the eikonal equation do not intersect (Zhao 2005). This is the case of smooth velocity models. Moreover, the upwind Godunov difference scheme enforces the causality of the solution (Sethian 1999), because the solution at a given grid point is determined by only those neighborhood points for which T is smaller. This is exactly what happens (Heughen's principle) during an advancing of the wavefront. In consequence, the iteration procedure converges very quickly and the solution is optimally accurate. However, we have to keep in mind that the used upwind scheme is based on the first order difference stencil. This implies the first order accuracy of the method.

In Fig. 2 an example of a simulation of wavefront positions for the velocity model of Rudna mine (vertical section) is shown. The ability of the algorithm to model the complex wavefront structure, including reflection and refraction effects, is clearly visible. The computation time for this simulation (2D grid with $N = 4 \cdot 10^5$ grid points) on 4 cores 2.4 GHz clocked Intel processor was about 90 milliseconds.

Being based on general concepts of the modern probabilistic inverse theory, the TRMLOC algorithm exhibits the same level of generality as any other, more traditional Bayesian location algorithms, However, unlike the classical probabilistic approaches, it performs implicit sampling simultaneously with the forward modeling due to the use of the finite difference based eikonal solver and EDT type of the misfit function.

The algorithm has also some limitations. First of all, the eikonal solver provides solutions only for the first arriving seismic phases (or the first arriving *P* or *S* waves in case of elastic waves). Including other phases within TRMLOC is possible, but it requires using the full waveform modeling algorithms, or multi-phases extensions to eikonal solver (Hauser *et al.* 2008, Rawlinson and Sambridge 2004). In both cases, however, the numerical efficiency of the algorithm is lost.

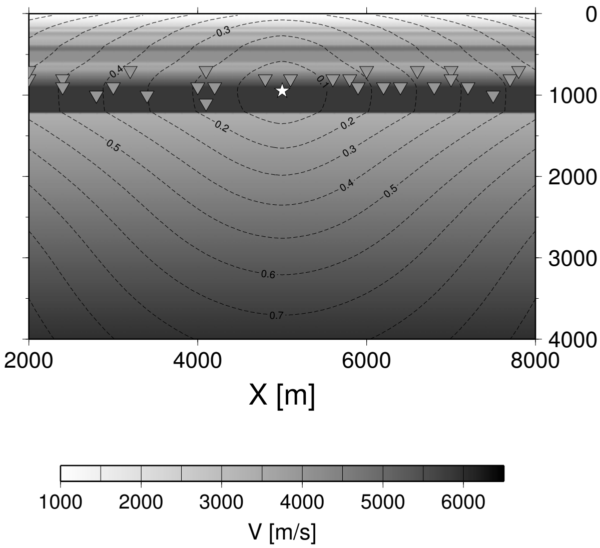


Fig. 2. The vertical section of the synthetic velocity model and wavefront positions simulated by FSM technique from the hypothetic rockburst (star). Open triangles denote seismometers of the underground seismic network operating in the mine.

The next limitation of the algorithm is related to the fact that the spatial resolution achieved by the algorithm is limited by the grid size used by the forward modeling algorithm. Achieving higher resolution requires a finer spatial grid but this increases the computation time linearly. Another problem connected to the spatial grid is that the eikonal solver used by TRMLOC is the first-order differential solver requiring quite fine grid for high numerical accuracy. Using higher-order solvers or more advanced front propagation techniques (Zhang *et al.* 2005) may thus be advisable.

An accumulated experience gathered when using the TRMLOC algorithm shows that the most time-critical part of the algorithm is calculation of integrated statistical characteristics of the *a posteriori* distribution like evidence, entropy, average model, *etc*. For a large grid, with the number of nodes of order 10^8 it takes about 70-80% of the whole calculation time. A remedy to this bottle-necked part of the algorithm is its redesigning using GPGPU technology which is extremely efficient in this type calculations (Kloc and Danek 2012). Further efficiency improvement is expected by porting the algorithm, especially the forward modeling part, to the parallel distributed computational platform, using, for example, MPI paralelization schemata (Quinn 2008).

3.2 Algorithms comparison

While getting the solution of the location task is conceptually rather simple, at least when the problem is formulated as the optimization task, estimating the reliability of the obtained solution is already more problematic and actually determines an efficiency of all location algorithms. Here we present a simple comparison of some popular algorithms based on counting the number of forward modelings (N_f) necessary for obtaining the solution and accompanying error estimators. Although such a comparison is not quite fair because different algorithms use different forward modeling techniques (ray tracing, eikonal solver, waveform modeling, *etc.*) it well illustrates the general feature, namely overall calculation times for different algorithms. In Table 1 we summarize such information for some most popular location softwares.

The first raw in Table 1 lists the most popular location softwares which solve the location tasks through the optimization procedure. The underlying algorithms are very fast, especially due to the fact that they use fast ray-tracers for point-to-point forward modelling. However, the reliability of the estimated location errors can be questioned, especially in complex velocity settings. The second listed group represents algorithms performing probabilistic location. All of them provide possibilities of an exhaustive error analysis, but in general are slower than those from the previous groups.

Table 1

Comparison of a number of forward modelings N_F performed by selected location algorithms

Method	Implementation	Modeling	Sampling	N_F
Linearized	Hypo71$^{(A)}$	ray tracer	—	$K \times N_s$
iterative	HYPOELLIPSE$^{(B)}$	ray tracer	—	$K \times N_s$
inversion	HypoInverse$^{(C)}$	ray tracer	—	$K \times N_s$
Probabilistic	Rloc $^{(E)}$	ray tracer	Metropolis	$\sim 10^6$
inversion	NLloc$^{(D)}$	eikonal solver	Gibbs, Oct-Tree	N_s
	TRMLOC	eikonal solver	implicit grid sampling	N_s

Explanations: A – Lee and Lahr (1975), B – Lahr (1989), C – Klein (2002), D – Lomax *et al.* (2000), E – Rudzinski and Debski (2012). N_s stands for number of stations (data) used for hypocenter location and K is a constant which depends on a number of iterations, velocity model, station distributions, *etc.*, and typically ranges between 10 and 50.

The most interesting is a comparison of NLloc and TRMLOC software as both implement in a different way almost identical location algorithms. The NLloc code provides a large flexibility (two different types of the misfit function and two different sampling methods), which makes it very versatile. On the other hand, TRMLOC is more specialized (mining or local events location) and uses only the EDT type misfit function (with l_1, l_2, and Cauchy norms). The most important difference between NLloc and TRMLOC is that NLloc performs classical, explicit sampling of the location space (3D or 4D) using the modern Oct-Tree or Gibbs samplers. It requires generation and evaluation of $\sim 10^4$ or more samples (Lomax *et al.* 2000) and the *a posteriori* pdf is then retrieved from the samples distribution. On the other hand, TRMLOC performs implicit exhaustive grid sampling evaluating *a posteriori* pdf for all grid nodes (typically $\sim 10^7$) but it is done simultaneously with forward modelling. Computational costs of this implicit sampling are comparable to an additional forward modelling. The other difference is that NLloc saves the forward modelling results on disk and then retrieves it when necessary, while TRMLOC keeps them in the memory. Finally, the TRMLOC algorithm uses the shared memory model of parallel computation via the OMP standard (Dagum and Menon 2002, Quinn 2008) and can easily be ported to MPI and GPGPU parallel computational models, while NLloc (version 6) does not support parallel computations. In consequence of the above implementation differences, the TRMLOC will typically run significantly faster than NLloc on modern computers with multi-core processors and large memories but will be over-performed by NLloc on simpler computers.

4. EXAMPLE: RUDNA COPPER MINE

The TRMLOC algorithm has been applied to locate a set of mining tremors induced by mining activity in the Rudna copper mine. This mine, situated in south-western Poland, runs a digital seismic network composed of 32 vertical seismometers located underground at exploitation depths ranging from 550 to 1150 m. The sampling period is $dt = 2$ ms. We have analysed 1647 events which occurred in two sections of the mine, between 2012 and 2014. The magnitudes M_L of the selected events is ranging from about 1.0 up to 3.5. The histogram of the magnitude distribution of the analysed events is shown in Fig. 3. We have assumed the Gaussian form of the *a priori* function

$$f(\mathbf{m}) = \exp\left(-\frac{(x-x_a)^2+(y-y_a)^2}{C_e^2}+\frac{(z-z_a)^2}{C_z^2}\right) \tag{15}$$

with $C_z = 500$ and $C_e = 2000$ and the *a priori* solutions (x_a, y_a, z_a) were provided by the mine. Choosing such values of C_z and C_e guarantees a very weak *a priori* constraining of the final solution. The parameter C_p defining the likelihood function was taken as $C_p = 0.01$ s. and its setting is discussed in Debski (2015). Let us note at this point that parameters C_z and C_e quantify the *a priori* expected location uncertainties with respect to the *a priori* location (x_a, y_a, z_a) (Debski 2010). Finally, following the standard mining practice we have used in this preliminary study the constant velocity model assuming P-wave velocity $V = 5900$ m/s.

Using the mining data we have firstly verified the scaling property of the TRMLOC algorithm with respect to the number of grid nodes used for calculations. For this analysis we have chosen another event of magnitude 3.6 which occurred on 26 June 2010, and due to its energy was recorded by all 30 running seismometers. Different grid sizes h, ranging from 10 up to 200 m, were used

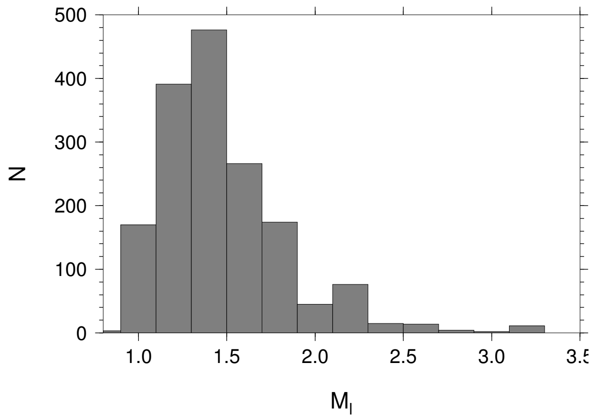


Fig. 3. Magnitude distribution of the analysed 1647 events.

for data processing, keeping the search region (X $\times$ Y $\times$ Z) 10 km $\times$ 10 km $\times$ 2 km the same for all grid settings. The computational efficiency of the algorithm is determined by a few factors, namely the efficiency of the forward modelling procedure, complexity of calculation of the *a posteriori* distribution and finally, calculations of various characteristics of the *a posteriori* distribution. Since numerical efficiency of the FSM solver is $O(N)$ and calculations of the *a posteriori* distribution and its characteristic is also of order $O(N)$, the expected overall efficiency is also proportional to N. The calculations were performed on 8 cores shared-memory computer with 2.4 GHz clock. The calculations were performed 50 times for each grid size to avoid a bias introduced by other tasks performed by operating system and the minimum computational times were recorded. Their dependence on the total number of grid nodes N and the grid size h is shown in Fig. 4. The theoretically predicted linear re-

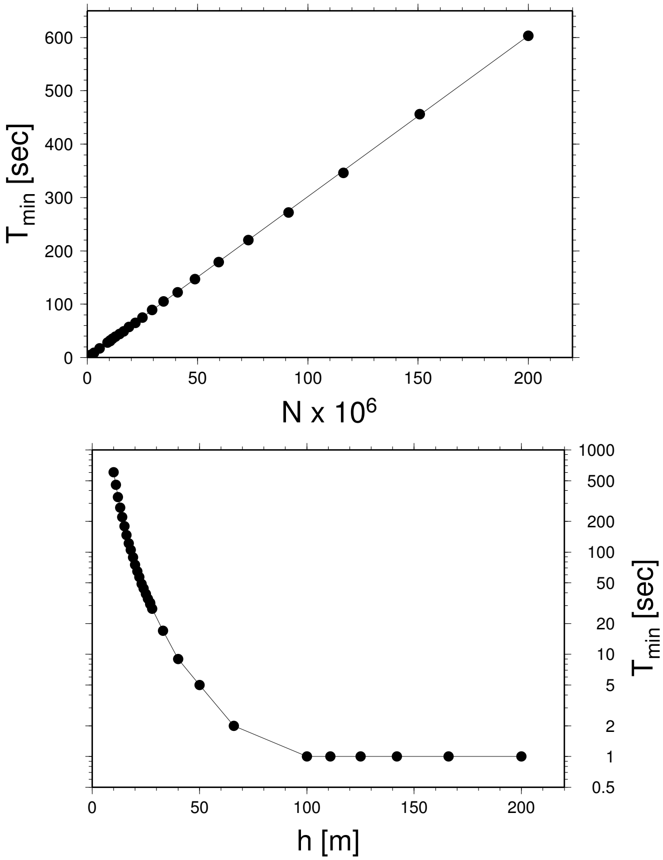


Fig. 4. The TRMLOC minimum computational time taken from an ensemble of 50 runs as a function of number of grid nodes N (up) and grid spacing h (down). Theoretically predicted linear dependence of computational time with N is clearly visible.

lation between computational time and number of grid nodes is clearly seen. Moreover, for all reasonable grid choices for which grid spacing h is smaller than the desired location accuracy (20-50 m) the calculation times are about 1 min. This is a fully acceptable computational efficiency which allows to use the algorithm in time critical application like a routine seismic data processing in mines.

In the next step we have analysed the location accuracy estimated by the diagonal elements of the *a posteriori* covariance matrix. The results are shown in Fig. 5.

It is clearly visible from Fig. 5 that epicentral (x and y) coordinates are much better resolved than the vertical one. This well known fact is a consequence of an almost planar geometry of the seismic network operating by mine. The horizontal location accuracy ranging in most cases between 10 and 30 meters is quite satisfactory. In case of the depth component the location accuracy is ranging typically between 100 and 300 m. As we discuss latter, on the largest

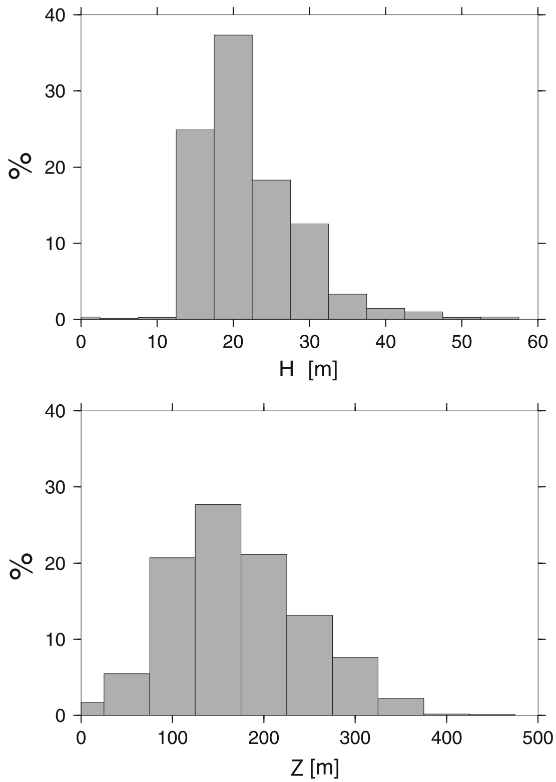


Fig. 5. The histogram of epicentral (up) and vertical (down) location errors for the analysed events.

vertical uncertainties are partially due to multi-modality of the *a posteriori* distribution and thus are overestimated by the used estimator (covariance matrix). The dependence of the location errors upon a number of stations contributing to the hypocenter location is shown in Fig. 6 While the horizontal errors are almost independent of Ns (only a minor increase for small Ns can be observed) the vertical location errors significantly decreases for large $Ns.$ This is not surprising, because even for a planar seismic network a larger horizontal span of the contributing stations is efficiently enhancing the "vertical information" in data (Debski 1996, Debski *et al.* 1997) by differentiating the ratio of vertical-to-horizontal hypocenter-station distances.

Next, we have analysed the correlation between the hypocenter coordinates. They are shown in Fig. 7 where the off-diagonal elements of the *a posteriori* covariance matrix are shown.

The distinct feature visible in this figure is an existence of an overall small correlations (at the level of ± 0.2) for most of events. Besides, we can see that for a number of events a large negative $R_{xy} \sim -0.4$ correlation between x and y coordinates has been reported. We interpret this as an effect of a particular (linear-like) station geometry for a given subset of events. Much more interesting is an existence of the secondary maximum for the R_{xz} coefficients around the value $R_{xz} \sim -0.8.$ This multi-modality of the R_{xz} distribution suggests an existence of two classes of the solution. To verify this hypothesis we have analysed all 1D *a posteriori* marginal probability density distributions and we have found out that all solutions split generally into two classes. To the first class belong solutions for which the $P(z)$ distributions for the depth coordi-

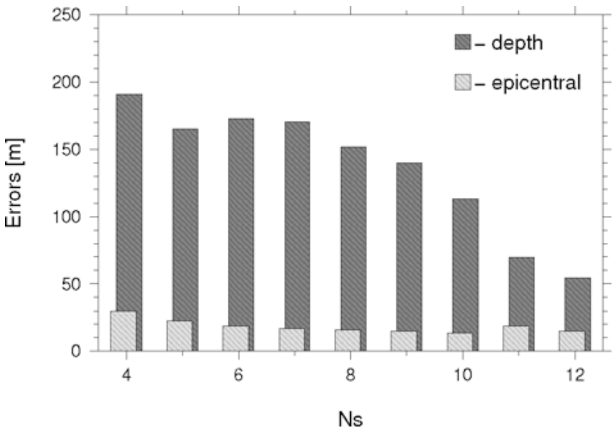


Fig. 6. The histogram of the epicentral and the vertical location errors upon the number of contributing seismic stations $Ns.$ While the horizontal errors are almost independent of Ns (only a minor increase for small Ns can be observed) the vertical location errors significantly decrease for large $Ns.$

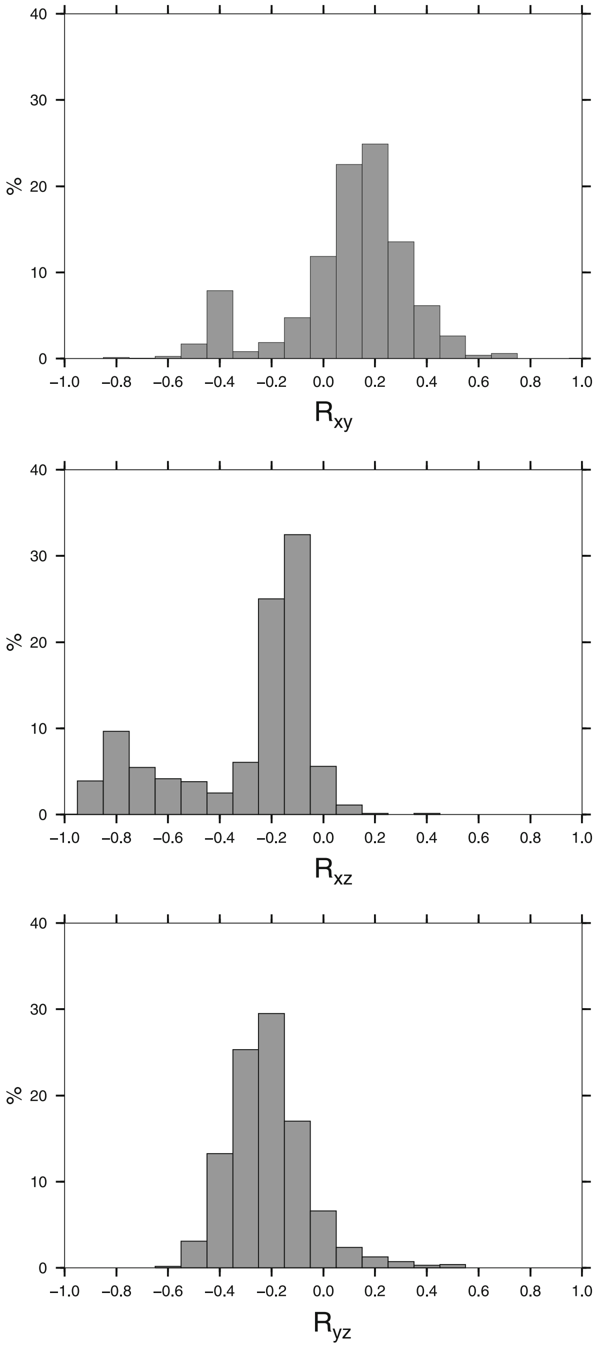


Fig. 7. The histogram of the correlation coefficients between hypocenter coordinates. In case of coefficients R_{xy} and R_{xz}, the existence of two peaks is clearly visible.

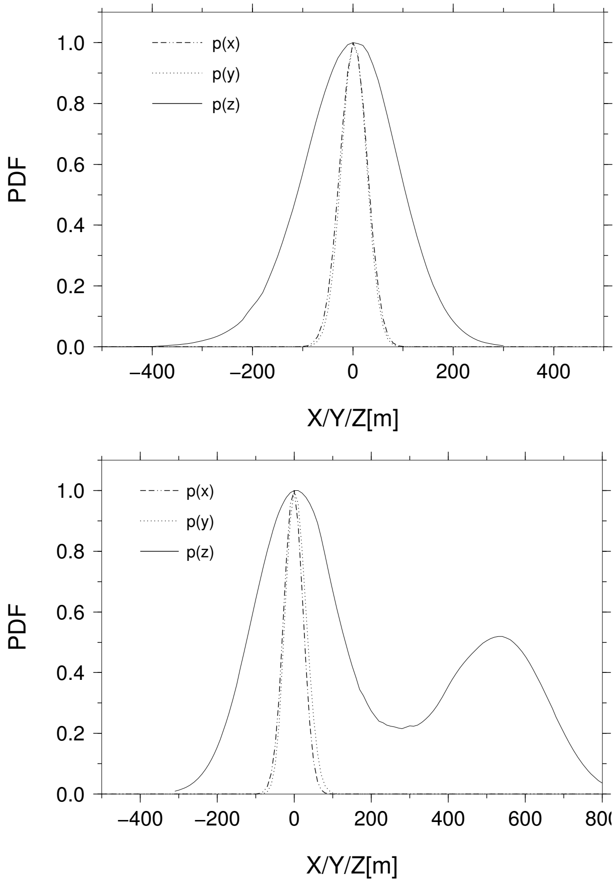


Fig. 8. The examples of two types of solutions with unimodal marginal depth distributions $P(z)$ (up) and two-modal distributions (down). While $P(x)$ and $P(y)$ distributions exibits typically a Gaussian-type shape for all events, the $P(z)$ splits into two class: unimodal (up) and two-modal (down). For the visualization purpose, all distributions were independently normalized and shifted towards unified maximum positions at zero.

nate are unimodal. The second class is formed by two-modal $P(z)$ distribution solutions. The distributions $P(z)$ typical for both classes are shown in Fig. 8

Similar 2D marginal distributions for exemplary events are shown in Fig. 9

The reason of the existence of two-modal depth solutions is obviously the planarity of the seismic underground network. For such a network geometry one can expect an existence of solution above and below the network plane. A possible remedy to this non-uniqueness problem (besides a rather costly hardware enhancing of the network) is using a more realistic, at least 1D velocity model differentiating between rock masses below and above the exploitation

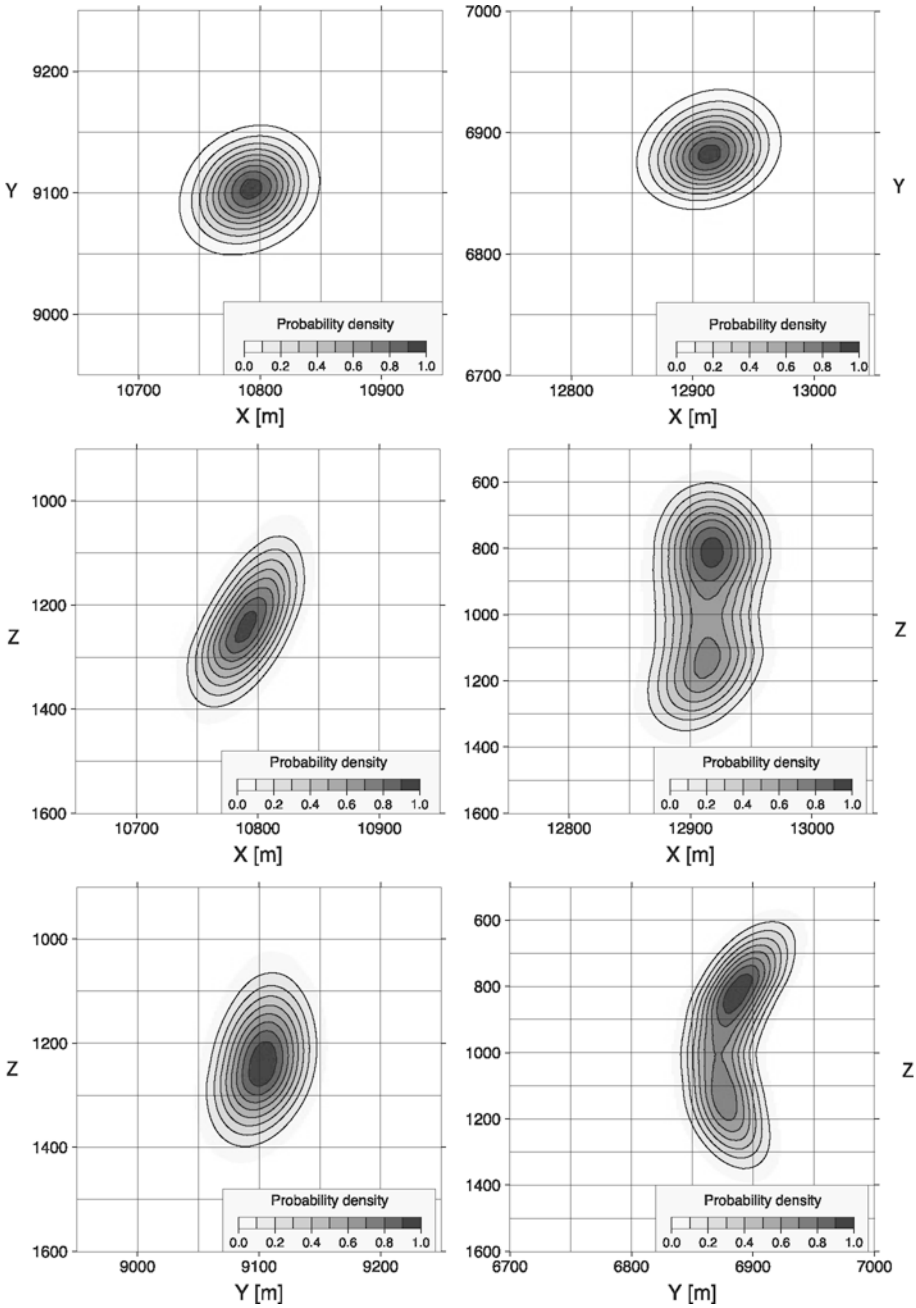


Fig. 9. The examples of two dimensional marginal distributions $P(x, y)$ (upper row), $P(x, z)$ (middle row), $P(y, z)$ (bottom row) for two classes of solutions: unimodal (left column) and two-modal (right column). Local coordinate system is used.

level. The next conclusion following the existence of the multi-modal solutions is that for such a solution an error estimation by the covariance matrix is not justified (Debski 2010). In this case an error estimator "should measure" a width of the main pick only, otherwise, we get an overestimation of the location errors.

Finally, the question arises whether the secondary maxima appear systematically below or above the primary depth solutions. To answer this question we have calculated the skewness γ coefficient for 1D *a posteriori* marginal distributions with respect to the position of the main pick. The results are shown in Fig. 10.

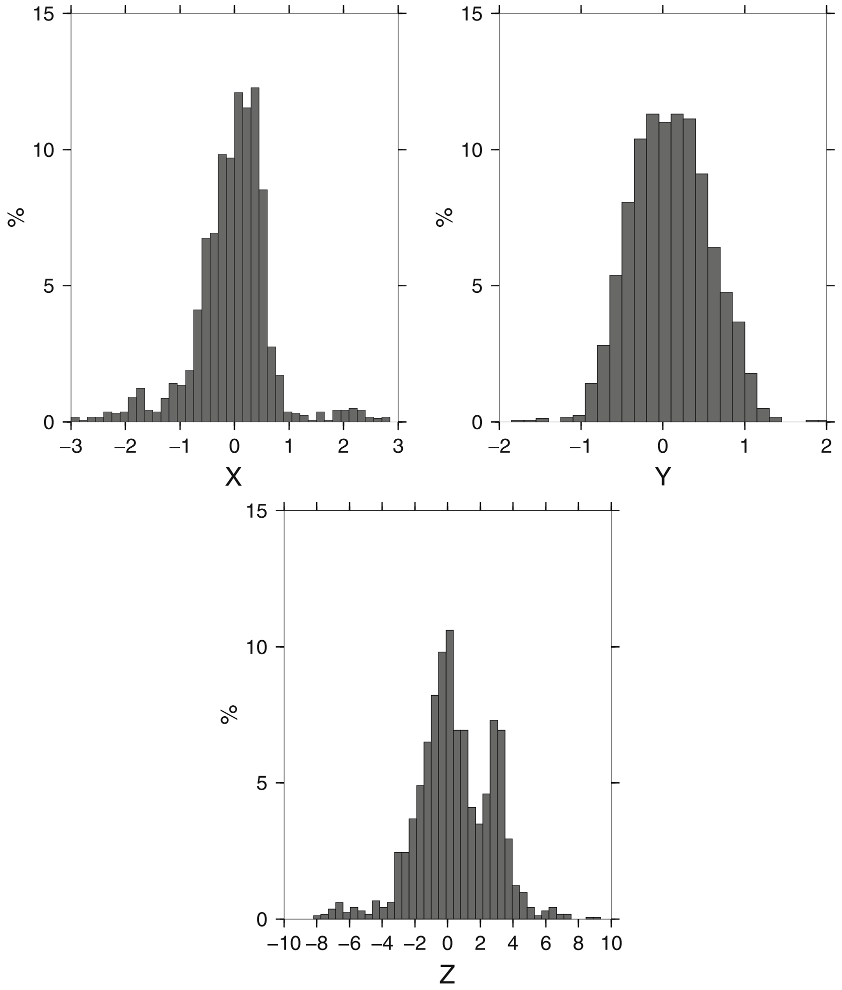


Fig. 10. The distribution of the skewness coefficients for the 1D *a posteriori* marginal distributions.

The visible secondary pick in the skewness distribution of $P(z)$ occurs at the positive value of γ. It means that the secondary pick in two-modal solutions is more frequently localised at larger depths than the main solution.

5. DISCUSSION AND CONCLUSIONS

Evaluation of the quality of inversion results is an important element while solving any inverse task in hand. With the classical approach, when the solution is the optimum-fitting model, our ability of evaluation of inversion quality is very limited and relies in practice on linearized approach. Many statistical tests and methods have been developed along this line of reasoning (see, *e.g.*, Brandt 1999). However, such an approach is more or less dependent on the particular solution found, so the results of the error analysis are "model dependent". The new possibilities are opened by the probabilistic approach which theoretically brings together all available uncertainties occurring during the inversion process and provides the statistics of the *a posteriori* errors. However, the approach is computationally demanding.

Exhaustive computations needed by the probabilistic inversion technique are due to two factors: a need of the sampling of the *a posteriori* distribution and the so-called "curse of dimensionality" effect (Curtis and Lomax 2001). The first factor is connected with the fact that obtaining some characteristics of the *a posteriori* distribution (average model, dispersion, *etc.*) requires inspection and evaluation of many models and thus many additional forward modelings. The second element, namely the fact that the models contributing to nonzero part of the *a posteriori* distribution forms a subset with a very small volume with respect to the whole model space makes the efficient sampling quite difficult (Mosegaard and Sambridge 2002). While building the TRMLOC algorithm we have explored the fact that for the source location task the model space (configuration space) is isomorphic with the space over which the forward operator (eikonal solver) is defined. This enabled us, by using the time reversal principle and the choice of the EDT-type likelihood function to make a calculation of the misfit function at all grid points having solved the forward problem for all receivers. In consequence there was no need of an additional sampling of the *a posteriori* pdf typical for classical Bayesian inversion. Obviously, a similar approach can be used in other situations when the model space coincides with the domain over which the forward operator solution is defined. For example, in case of the source location based on the full waveform inversion the misfit function can be defined by a difference between cross-correlations or more advanced measures (Kennett and Fichtner 2012) of the back propagated waveforms. The method can further be generalized by using different fields and different measures (Larmat *et al.* 2009, Ulrich *et al.* 2009) for constructing the misfit/likelihood function. The important question arises, however, at this point whether a similar method of "implicit sampling" can be developed for other time-reversal inverse problems when analyzed parameters are, for example source moment tensor, source time function, *etc*. An analysis of this issue will be presented elsewhere.

The important point of the developed algorithm is using the EDT type misfit function within the probabilistic inversion. The concept of the EDT misfit function is well established in seismology and its use within the classical, optimization-based inversion technique has proved to be very successful. However, using the EDT misfit to define the likelihood function in probabilistic inversion is by no means obvious. The reason is that the "true" likelihood function is actually the probability density function describing the sum of modelling and observational errors. Apparently, the likelihood-type function defined through the EDT misfit function does not have this property. The open question is thus if this difference is operationally important (eventually under which condition) or not. Apparently the maximum likelihood solution which is equivalent to the minimum of the misfit function is not influenced by this difference. However, this in not the case of other moments, like for example variance of the *a posteriori* distribution and thus some systematic bias can occur. This issue apparently must be clarified.

Finally, let us note that the seismic source location task is a very special type of inverse problem in which the model space is the same as the space over which the forward operator acts. We have explored this identity in the TRM-LOC algorithm to perform the implicit sampling of the *a posteriori* distribution and thus ensuring its very high numerical efficiency. We think that this approach, when combined with time-reversal principle, can further be extend to analyze different characteristics of seismic sources. The crucial point is the observation that the seismic signals back-propagated to the hypocenter location contain all information about the source as the originally recorded waveforms. Thus, to perform the full waveform inversion for seismic source parameters we do not need to use the originally recorded waveforms. Instead we can use the back-in-time propagated signals at the hypocenter location. In consequence, the most time consuming part of any full waveform inversion, namely the propagation of seismic waves from hypocenter to recording stations performed many times during inversion can be replaced by a single back-in-time seismic waves propagation. Apparently, additional uncertainties are introduced with such an approach and a lot of further analysis is necessary.

The developed algorithm has allowed us to make a nontrivial error analysis for a set of 1647 events from Rudna copper mine. The most interesting result of this analysis is demonstrating of an existence of a sub-class of two-modal solutions. We have found out that in most cases the secondary maximum in the *a posteriori* distributions occur at larger depths (often below the exploitation level). Possibilities of such deeper, below the exploitation level, seismic events were discussed, for example, by Gibowicz and Kijko (1994), Gibowicz and Lasocki (2001). In the light of the obtained results, a hypothesis of generation of seismic events in the footwall in Rudna copper mine cannot be ruled out.

Acknowledgment. This paper was partially support by grant No. 2011/01/B/ST10/07305 from the National Science Center, Poland. G. O'Brien and the anonymous reviewers are acknowledged for their help in improving the manuscript.

References

Aki, K., and P. Richards (1985), *Quantitative Seismology*, Freeman and Co., San Francisco

Artman, B., I. Podladtchikov, and B. Witten (2010), Source location using time-revers imaging, *Geophys. Prosp.* **58**, 5, 861-873, DOI:10.1111/j.1365-2478.2010.00911.x

Bai, L., Z. Wu, T. Zhang, and I. Kawasaki (2006), The effect of distribution of stations upon location error: Statistical tests based on the double-difference earthquake location algorithm and the bootstrap method. *Earth Planets Space* **58**, 2, e9-e12, DOI:10.1186/BF03353364.

Brandt, S. (1999), *Data Analysis. Statistical and Computational Methods for Scientists*, Springer-Verlag.

Bulland, R. (1976), The Mechanics of locating earthquakes, *Bull. Seismol. Soc. Am.* **66**, 1, 173-187.

Chib, S. and Greenberg (1995), Understanding the Metropolis-Hastings Algorithm, *Am. Stat.* **49**, 327-335, DOI: 10.1080/00031305.1995.10476177.

Curtis, A. and A. Lomax (2001), Prior information sampling distributions and the curse of dimensionality. *Geophysics* **66**, 2, 372-378, DOI:10.1190/1.1444928.

Dagum, L. and R. Menon (2002), OpenMP: an industry standard API for shared-memory programming, *Comput. Sci. Eng. IEEE* **5**, 1, 46-55, DOI:10.1109/99.660313.

Debski, W. (1996). *Location of Seismic Events A Quest for Accuracy*, Springer-Verlag, Berlin, DOI: 10.1007/BFb0011773.

Debski, W. (2004). Application of Monte Carlo techniques for solving selected seismological inverse problems, *Publs. Inst. Geophys. Pol. Acad. Sc.* **B-34**, 367, 1-207.

Debski, W. (2010), Probabilistic inverse theory, *Adv. Geophys.* **52**, 1-102, DOI:10.1016/S0065-2687(10)52001-6.

Debski, W. (2015), Using meta-information of a posteriori Bayesian solutions of the hypocenter location task for improving accurcy of location error estimation, *Geophys. J. Int.* **201**, 3, 1399-1408, DOI:1093/gji/ggv083.

Debski, W., B. Guterch, H. Lewandowska, and P. Labak (1997), Earthquake sequences in the Krynica region Western Carpathians 1992 - 1993, *Acta Geophys. Pol.* **XLV**, 4, 255-290.

Fink, M. (1997), Time reversed acoustic, *Physics Today* **50**, 3, 34-40, DOI:10.1063/1.881692.

Fink, M., D. Cassereau, A. Derode, C. Prada, P. Roux, M. Tanter, J.-L. Thomas, and F. Wu (2000), Time-reversed acoustics, *Reports on Progress in Physics* **63**, 12, 1933-1994.

Fink, M. and M. Tanter (2010), Multiwave imaging and super resolution. *Physics Today* **63**, 2, 28-33, DOI:10.1063/1.3326986.

Font, Y., H. Kao, S. Lallemand, C.-S. Liu, and L.-Y. Chiao (2004), Hypocentre determination offshore of eastern Taiwan using the Maximum Intersection method, *Geophys. J. Int.* **158**, 655-675, DOI:10.1111/j.1365-246X.2004.02317.x.

Gajewski, D. and E. Tessmer (2010), Reverse modelling for seismic event characterization, *Geophys. J. Int.* **163**, 1, 276-284, DOI: 10.1111/j.1365-246X.2005.02732.x.

Gibowicz, S. J. and A. Kijko (1994), *An Introduction to Mining Seismology*. San Diego: Academic Press.

Gibowicz, S. J. and S. Lasocki (2001), Seismicity Induced by Mining: Ten Years Later. *Adv. Geophys.* **44**, 39-181, DOI:10.1016/S0065-2687(00)80007-2.

Gilks, W., S. Richardson, and D. Spiegelhalter (1995), *Markov Chain Monte Carlo in Practice*, Chapman& Hall/CRC Press.

Giovambattista, R. and S. Barba (1997), An estimate of hypocentre location accuracy in a large network: possible implications for tectonic studies in Italy, *Geophys. J. Int.* **129**, 1, 124-132, DOI:10.1111/j.1365-246X.1997.tb00941.x.

Hauser, J., M. Sambridge, and N. Rawlinson (2008). Multiarrival wavefront tracking and its applications, *Geochem. Geophys., Geosys.* **9**, 11, DOI:10.1111/j.1365-246X.1997.tb00941.x.

Husen, S. and J. Hardebeck (2010). Earthquake location accuracy, *Community online resources for statistical seismicity analysis*. DOI: 10.5078/corssa-55815573.

Husen, S., E. Kissling, E. Flueh, and G. Asch (1999), Accurate hypocentre determination in the seismogenic zone of the subducting Nazca Plate in northern Chile using a combined on-/offshore network, *Geophys. J. Int.* **138**, 3, 687-701, DOI:10.1046/j.1365-246x.1999.00893.x.

Kennett, B. and A. Fichtner (2012), A unified concept for comparison of seismograms using transfer functions, *GJI* **191**, 3, 1403-1416, DOI: 10.1111/j.1365-246X.2012.05693.x.

Klein, F. (2002), *User's guide to HYPOINVERSE-2000: A Fortran program to solve for earthquake locations and magnitudes*, US Geological Survey.

Kloc, M. and T. Danek (2012), *The Multi GPU Accelerated Waveform Inversion in Distributed OpenCL Environment*, Volume 151 of *Lecture Notes in Electrical Engineering*, Springer New York, .

Kremers, S., A. Fichtner, G. Brietzke, H. Igel, C. Larmat, L. Huang, and M. Kaser (2011), Exploring the potentials and limitations of the time-reversal imaging of finite seismic sources. *Solid Earth* **2**, 1, 95-105, DOI:10.5194/se-2-95-2011.

Lahr, J. (1989), *HYPOELIPSE (revised); A computer program for determining local earthquake hypocentral parameters, magnitude and first motion pattern* , US Geological Survey.

Larmat, C., R. Guyer, and P. A. Johnson1 (2009), Tremor source location using time reversal: Selecting the appropriate imaging field, *Geophys. Res. Lett.* **36**, (L22304),DOI:10.1029/2009GL040099 .

Larmat, C., J. Tromp, Q. Liu, and J.-P. Montagner (2008), Time reversal location of glacial earthquakes, *J. Geophys. Res.* **113**, B09314, 1-9, DOI:10.1029/2008JB005607

Lee, W. and J. Lahr (1975), *HYPO71 (revised); A computer program for determining hypocenter, magnitude, and first motion pattern of local earthquakes*, US Geological Survey.

Lehmann, E. L. and G. Casella (1998). *Theory of Point Estimation*, Springer Texts in Statistics. New York: Springer-Verlag.

Lomax, A. (2005). A Reanalysis of the Hypocentral Location and Related Observations for the Great 1906 California Earthquake, *Bull. Seismol. Soc. Am.* **95**, 3, 861-877, DOI: 10.1785/0120040141.

Lomax, A., A. Michelini, and A. Curtis (2009), *Earthquake Location, Direct, Global-Search Methods*, Volume 5. New York: Springer, DOI: 10.1007/978-0-387-30440-3.

Lomax, A., J. Virieux, P. Volant, and C. Berge (2000). *Probabilistic earthquake location in 3D and layered models: Introduction of a Metropolis-Gibbs method and comparison with linear locations*, Amsterdam: Kluver, DOI:10.1007/978-94-015-9536-0_5.

Lomax, A., A. Zollo, P. Capunao, and J. Virieux (2001), Precise absolute earthquake location under Somma-Vesuvius volcano using a new three-dimensional velocity model, *Geophys. J. Int.* **146**, 2, 313-331, DOI:10.1046/j.0956-540x.2001.01444.x.

Masson, Y., P. Cupillard, Y. Capdeville, and B. Romanowicz (2014), On the numerical implementation of time-reversal mirrors for tomographic imaging. *Geophys. J. Int.*, **3**, 1-11, DOI:10.1093/gji/ggt459.

Matsu'ura, M. (1984), Bayesian estimation of hypocenter with origin time eliminated, *J. Phys. Earth.* **32**, 6, 469-483.

Menke, W. (1989), *Geophysical Data Analysis: Discrete Inverse Theory*, International Geophysics Series. San Diego: Academic Press.

Mosegaard, K. and M. Sambridge (2002). Monte Carlo analysis of invers problems. *Inv. Prob. 18*, 3, R29-45, DOI: 10.1088/0266-5611/18/3/201.

Mosegaard, K. and A. Tarantola (2002), *International Handbook of Earthquake & Engineering Seismology*, Volume 81 of *International Geophysics Series*. Academic Press.

O'Brien, G., J. Lokmer, L. D. Barros, C. Bean, G. Saccorotti, J.-P. Metaxian, and D. Patane (2011). Time reverse location of seismic long-period events recorded on Mt Etna. *Geophys. J. Int.* **184**, 1, 452-462, DOI:10.1111/j.1365-246X.2010.04851.x.

Pavlis, G. L. (1992), Appraising relative earthquake location errors, *Bull. Seismol. Soc. Am.* **82**, 2, 836-859.

Podvin, P. and I. Lacomte (1991), Finite-difference computtion of traveltimes in very contrasted velocity models: a massively paarallel approach and its associated tools, *Geophys. J. Int.* **105**, 1, 271-284, DOI:10.1111/j.1365-246X.1991.tb03461.x.

Quinn, M. (2008), *Parallel Programming in C with MPI and OpenMP*. McGraw-Hill Education.

Rawlinson, N. and M. Sambridge (2004), Multiple reflection and transmission phases in complex layered media using a multistage fast marching method. *Geophys.* **69**, 5, 2178-2193, DOI:10.1190/1.1801950.

Rudzinski, L. and W. Debski (2011). Extending the Double Difference location technique for mining applications - part I: Numerical study. *Acta Geophys.* **59**, 4, 785-814, DOI:10.2478/s11600-011-0021-5.

Rudzinski, L. and W. Debski (2012), Extending the Double Difference location technique - improving hypocenter depth determination. *J. Seismol.* **17**, 1, 83-94, DOI:10.1007/s10950-012-9322-7.

Sambridge, M. and K. Mosegaard (2002), Monte Carlo Methods in Geophysical Inverse Problems. *Rev. Geophys.* **40**, 3, 3.1-3.29, DOI:10.1029/2000RG000089.

Scalerandi, M., M. Griffa, and P. Johnson (2009), Robustness of computational time reversal imaging in media with elastic constant uncertainties. *J. Appl. Phys.* **106**, 114911, DOI:10.1063/1.3269718.

Sethian, J. A. (1999), *Level set methods and fast marching methods: evolving interfaces in computational geometry fluid mechanics computer vision and materials science*. Cambridge Monographs on Applied and Computational Mathematics. New York: Cambridge University Press.

Steiner, B. and E. Saenger (2012), Comparison of 2D and 3D time-reverse imaging - A numerical case study, *Comput & Geosci.* **46**, 174-182, DOI:10.1016/j.cageo.2011.12.005.

Tarantola, A. (2005), *Inverse Problem Theory and Methods for Model Parameter Estimation*, Philadelphia: SIAM.

Thurber, C. and N. Rabinowitz (2000). *Advances in Seismic Event Location*, Volume 18. Springer.

Tromp, J., C. Tape, and Q. Liu (2005), Seismic tomography, adjoints methods, time reversal and banana-doughnut kernels, *Geophys. J. Int.* **160**, 1, 195-216, DOI:10.1111/j.1365-246X.2004.02453.x.

Ulrich, T., K. V. D. Abeele, P.-Y. L. Bas, M. Griffa, B. Anderson, and R. Guyer (2009), Three component time reversal: Focusing vector components using a scalar source, *J. Appl. Physics* **106**, 11, 113504, DOI:10.1063/1.3259371.

Ulrich, T., A. Sutin, R. Guyer, and P. Johnson (2008), Time reversal and non-linear elastic wave spectroscopy (TR NEWS) techniques, *Int. J. of Non-Lin. Mech.* **43**, 3, 209-216, DOI:10.1016/j.ijnonlinmec.2007.12.017.

Vidale, J. (1990), Finite-difference calculation of traveltime in three dimensions. *Geophysics* **55**, 5, 521-526, DOI:10.1190/1.1442863.

Virieux, J., S. Operto, H. Ben-Hadj-Ali, R. Brossier, V. Etienne, and F. Sourbier (2009). Seismic wave modeling for seismic imaging. *Leading Eadge*,**28**, 5, 538-544, DOI:10.1190/1.3124928.

Waldhauser, F. and W. Ellsworth (2000), A double-difference earthquake location algorithm: method and application. *Bull. Seismol. Soc. Am.* **90**, 6, 1353-1368.

Wiejacz, P. and W. Debski (2001), New Observation of Gulf of Gdansk Seismic Events. *Phys. Earth Planet. Int.* **123**, 2-4, 233-245, DOI:10.1016/S0031-9201(00)00212-0.

Witten, B. and B. Artman (2011). Signal-to-noise estimates of time-reverse images. *Geophysics* **76**,2, MA1-MA10, DOI:10.1190/1.3543570.

Zhang, L., J. Rector, and G. Hoversten (2005), Eikonal solver in the celerity domain, *Geophys. J. Int.* **162**, 1, 1-8, DOI:10.1111/j.1365-246X.2005.02626.x.

Zhao, H. (2005), Fast Sweeping Method for Eikonal equations. *Math. Comput.* **74**, 603-627, DOI:10.1090/S0025-5718-04-01678-3.

Zhou, H. (1994), Rapid three-dimensional hypocentral determination using a master station method, *J. Geophys. Res.* **99**, B8, 715439-15455, DOI:10.1029/94JB00934.

Received 19 February 2015
Received in revised form 29 October 2015
Accepted 19 November 2015

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2410-2429
DOI: 10.1515/acgeo-2016-0094

Acoustic Emission Parameters of Three Gorges Sandstone during Shear Failure

Jiang XU, Yixin LIU, and Shoujian PENG

State Key Laboratory of Coal Mine Disaster Dynamics and Control, Chongqing University, Chongqing, China

State and Local Joint Engineering Laboratory of Methane Drainage in Complex Coal Gas Seam, Chongqing University, Chongqing, China
e-mails: jiangxu@cqu.edu.cn, yxliu@cqu.edu.cn (corresponding author), sjpeng@cqu.edu.cn

Abstract

In this paper, an experimental investigation of sandstone samples from the Three Gorges during shear failure was conducted using acoustic emission (AE) and direct shear tests. The AE count rate, cumulative AE count, AE energy, and amplitude of the sandstone samples were determined. Then, the relationships among the AE signals and shearing behaviors of the samples were analyzed in order to detect micro-crack initiation and propagation and reflect shear failure. The results indicated that both the shear strength and displacement exhibited a logarithmic relationship with the displacement rate at peak levels of stress. In addition, the various characteristics of the AE signals were apparent in various situations. The AE signals corresponded with the shear stress under different displacement rates. As the displacement rate increased, the amount of accumulative damage to each specimen decreased, while the AE energy peaked earlier and more significantly. The cumulative AE count primarily increased during the post-peak period. Furthermore, the AE count rate and amplitude exhibited two peaks during the peak shear stress period due to crack coalescence and rock bridge breakage. These isolated cracks later formed larger fractures and eventually caused ruptures.

Key words: sandstone, shear failure, acoustic emission, displacement rate.

1. INTRODUCTION

Rock slope stability estimations are required for a variety of civil, road and mining engineering projects not only in feasibility studies, but also in the excavation and operating stages (Taheri and Tani 2010). In these projects, numerous tunnels and caverns are created through brittle rock mass under high amounts of stress. Under these stressful conditions, rock structures become less stable, ultimately resulting in rock failure (Chang and Lee 2004, Stock *et al.* 2012). Previous studies concerning the behaviors of rock under various strain rates have shown that the compressive strength, Young's Modulus, and Poisson's Ratio of rock are highly dependent upon the strain rate (Kawamato and Saito 1974, Kranz 1979, Lajtai *et al.* 1991). However, a more detailed study of the mechanical behaviors of rock could provide valuable information regarding rock structure design as well as numerous mining operations, such as drilling, blasting, and crushing. Strain levels can change within seconds during blasting and earthquakes, or over many years during mining operations (Ray *et al.* 1999). Rock is a typical inhomogenous and anisotropic material that contains several natural defects with various scales, such as micro cracks, pores, joint inclusions, and precipitates. A large number of acoustic emission (AE) events occur when rock specimens are subjected to loading stages until failure. AE signals are associated with the initiation and propagation of micro-cracks, and provide a significant amount of information regarding the internal structural changes that occur in rock. Therefore, the behaviors of rock are reflected by their AE signals (Li *et al.* 2010, Lockner 1993, Katsuyama 1996).

Numerous studies concerning the acoustic emissions of rock have been published. Moradian *et al.* (2010) investigated the AE signal characteristics of various joints and concluded that AE signals could be used to effectively monitor the shear behaviors of joints. Chang and Lee (2004) investigated the fracture and damage mechanisms induced by micro-crack accumulation in rocks by conducting a moment tensor analysis and applying the moving point regression technique to acoustic emission (AE) and strain data obtained via triaxial compression tests. Xu *et al.* (2009) used the acoustic emission (AE) technique to continuously monitor micro-crack development and failure in rock samples in real-time. Majewska and Mortimer (2006) studied non-linear dynamics of acoustic emission (AE) generated in coal samples subjected to gas sorption-desorption. According to the results, the acoustic emission (AE) signal characteristics of the coal rocks accurately reflected their behaviors under different conditions. In the past few decades, many other studies concerning the mechanical behaviors of rocks have been conducted using acoustic emission location technology (Yang *et al.* 2012, Goszczyńska *et al.* 2012, Rao and Kusunose 1995).

However, the methods currently used to reflect rock failure using AE technology are based on the analysis of AE signal characteristics. In this paper, laboratory direct shear tests were conducted under constant normal loading conditions using various shear displacement rates and AE signals with different characteristics. The parameters included the AE count rate, cumulative AE count, AE energy, and amplitude. The AE signal most representative of the failure process was determined based on the relationships among the AE signals and shear failure characteristics.

2. ROCK SAMPLES AND EXPERIMENTAL TECHNIQUES

2.1 Rock samples

The sandstone samples were obtained from the Three Gorges region in ChongQing, China. The samples were primarily composed of quartz, feldspar, chert, and muscovite with a grain size distribution of 0.1 to 0.5 mm. Drilling cores without obvious fractures were selected and cut into cubes with dimensions of approximately $40 \times 40 \times 40$ mm. Then, the surfaces of the samples were ground in order to ensure the flatness, verticality, and parallelism standards provided by ISRM. Mesh sandpaper (600, 800, 1200, and 2000 grit) was used to further grind the surfaces in order to ensure a parallelism error of less than 0.02 mm. Five specimens were adopted in laboratory experiments of each conditions, and one typical specimen of each condition was chosen to display and detailed analysis. The Young's modulus is 11.89 GPa, the Poisson ratio is 0.37, the uniaxial compressive strength is 55.97 MPa, the density is 2.33 $g \cdot cm^{-3}$.

2.2 Experimental apparatus

Direct shear tests were performed on the rock specimens under different shear displacement rates using a direct shear apparatus (Fig. 1a). The apparatus consisted of a loading system and acoustic emission monitoring system. A more detailed description of the direct shear apparatus is provided by Xu *et al.* (2011).

In displacement rate loading, unlike stress loading, the applied load can be controlled and reduced to a value lower than the peak shear strength. This feature allows for the modelling of the strain-softening behaviors of rock. Thus, constant displacement rates were used to control the loading during the direct shear tests. The control system of the apparatus, which was entire-

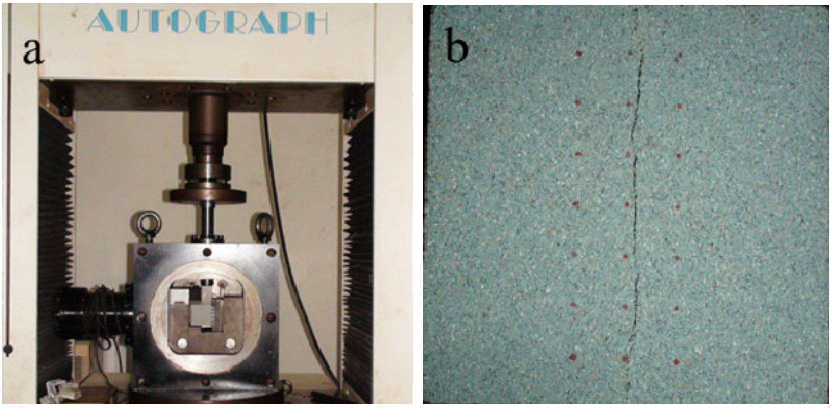


Fig. 1: (a) The direct shear apparatus, and (b) fractured rock sample after the experiment.

ly digital, performed consistently and exhibited a high loading rate accuracy. The loading process stopped automatically once the specimens were broken. The real-time AE signals were acquired using a PCI-2 acoustic emission system. A transducer frequency range of 20~400 kHz and sampling frequency of 1 MHz were used for the purposes of this study. In addition, a threshold of 40 dB was used in order to achieve a high signal/noise ratio. The AE count rate, cumulative AE count, AE energy, and amplitude were used as the acoustic emission parameters.

2.3 Experimental procedure

Constant displacement rates of 0.200, 0.020, and 0.002 mm/min were adopted during the experiments in order to determine the acoustic emission characteristics of the sandstone samples under various shear strain rates. The experiments were conducted on five specimens under different conditions. A specimen representative of each experimental condition was selected and analyzed in detail. In order to ensure synchronization during data acquisition, the loading system and acoustic emission system were initiated simultaneously. The length, width, and height, and other basic parameters of the samples were measured before the experiment. When installed the acoustic emission transducer on the back surface close to the pre shear failure surface, we first spread the face detection with butter and clung to the specimen (as shown in Fig. 2), at the same time fixing the transducer on the specimen with tape to avert the influence caused by transducer off. Figure 1b shows the fractured rock sample after the experiment.

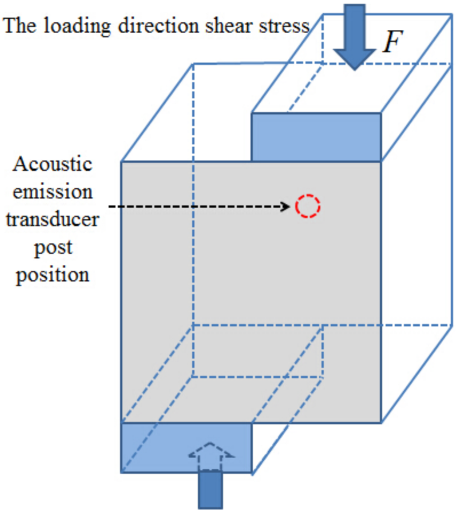


Fig. 2. Acoustic emission transducer paste position.

3. RESULTS

3.1 Deformation characteristics

Figure 3 displays the shear stress values of the specimens under different displacement rates. As shown in this figure, as the shear displacement rate increased, the shear stress also increased with no apparent periodical characteristics. This was likely because plastic deformation occurs at a more consistent rate during slow straining than during rapid loading (Lavrov 2001). The shear strength of the sandstone also increased with the displacement rate. Table 1 displays the shear strength of the sandstone samples under different displacement rates. Figures 4a and b display the relationships between the displacement rates and shear strength and displacement, respectively. As the displacement rate increased, both the shear strength and displacement decreased linearly. These results corresponded with those obtained by Li *et al.* (2010) and Zhukov (1965).

Table 1

Shear strength and displacement of the samples under different displacement rates

V [mm/min]	$-lg\ V$	τ_{max} [MPa]	S [mm]
0.002	2.700	14.460	0.477
0.020	1.700	13.398	0.447
0.200	0.700	11.718	0.346

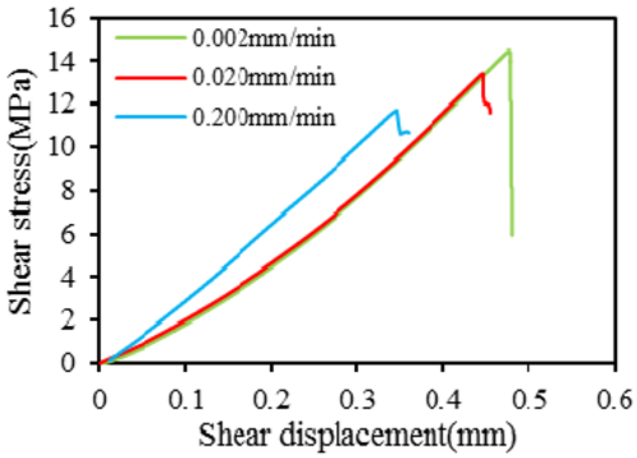


Fig. 3. Shear stress values of the samples under different displacement rates.

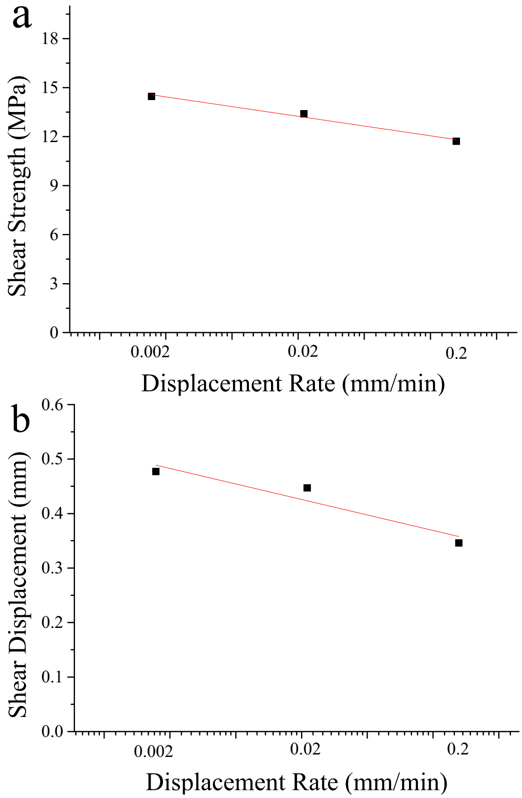


Fig. 4: (a) Displacement rate *versus* shear stress, and (b) displacement rate *versus* shear displacement.

3.2 AE count rate with time evolution characteristics

AE count rate can be used to identify internal transient damage, such as crack initiation and propagation. Figures 5a-c display the relationship between the AE count rate and shear stress under different displacement rates over time. The first acoustic emissions occurred at shear strength values of 0.39 τ_{max}, 0.47 τ_{max}, and 0.50 τ_{max}, where τ_{max} represents the peak shear stress, or shear strength. Thus, as the displacement rate increased, the amount of shear stress required for micro-crack initiation also increased. As shown in these figures, the displacement rate began to increase as the shear stress peaked, then rapidly decreased. In addition, the AE count rate peaked as the shear stress decreased and peaked earlier than the shear stress at higher displacement rates.

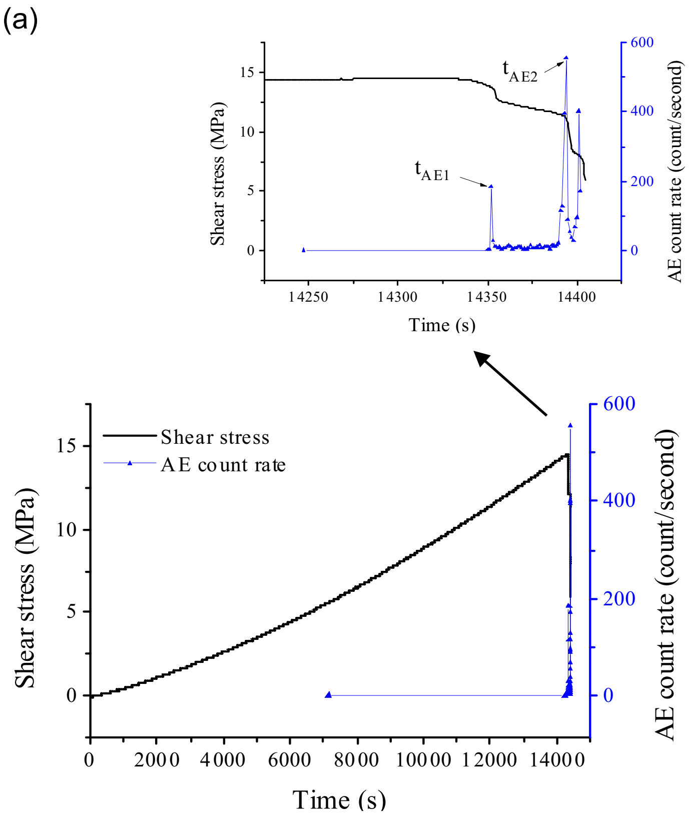
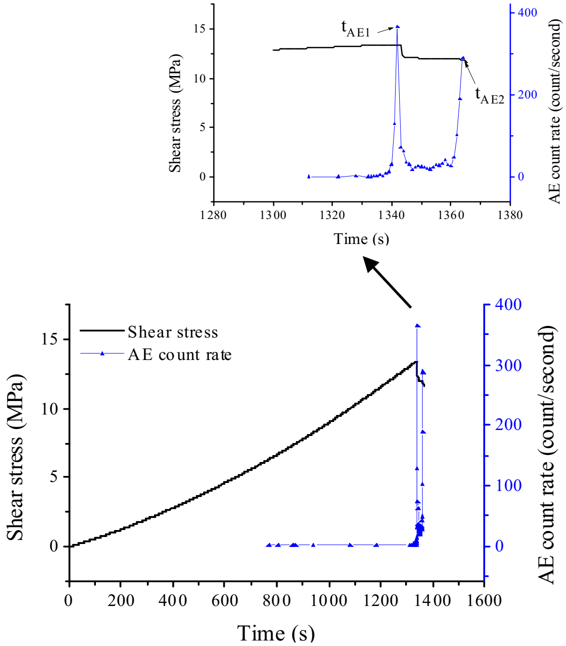


Fig. 5. Caption on next page.

(b)



(c)

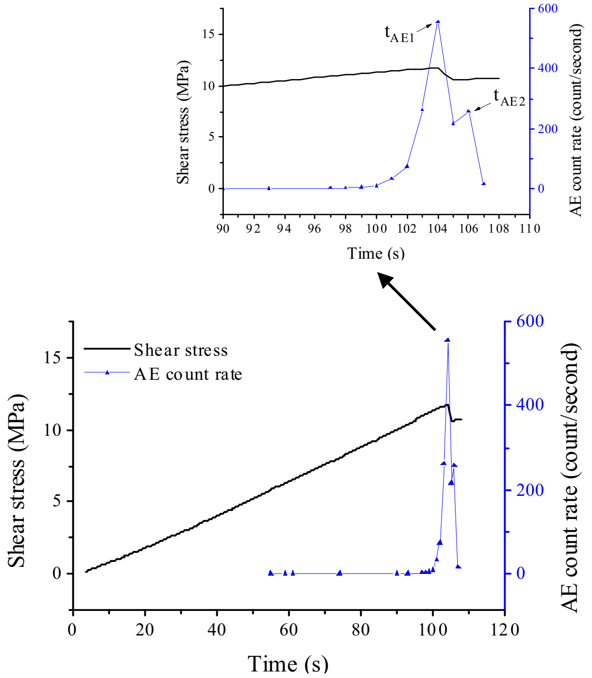


Fig. 5. Shear stress *versus* AE count rate under different displacement rates over time: (a) 0.002 mm/min, (b) 0.02 mm/min, and (c) 0.2 mm/min.

3.3 Time evolution characteristics of the cumulative AE count

In this study, the cumulative AE count was used to reflect the internal damage to the sandstone specimens under shear loading. As shown in Figs. 6a-c, the AE count rate varied as the cumulative AE count periodically increased. The AE counts predominantly occurred during the slowing stage, suggesting that the cracks within the sandstone specimens primarily occurred during this time period. Figure 6d displays the cumulative AE counts under the different displacement rates. This result demonstrates that with the increase of displacement rate, the cumulative AE count decreased.

3.4 Time evolution characteristics of AE Energy

AE energy represents the amount of elastic energy released as a result of crack initiation and propagation over time. Figures 7a-c display the relationship between the AE energy and shear stress under different displacement rates over time. As shown in these figures, as the displacement rate increased, the AE energy began to increase earlier in time. In addition, the AE energy peaked before the shear stress at a displacement rate of 0.200 mm/min. These results indicated that fewer micro-cracks propagated as the displacement rate increased, resulting in the accumulation of elastic energy. This elastic energy was released earlier in time at a relatively high rate. In contrast, the AE count rate only peaked once as the shear stress increased. Thus, during shear failure, only a large amount of elastic energy was released, resulting in the rough formation of a single, large crack. Figure 7d displays the time evolution curve of the AE energy under different displacement rates. As shown in this figure, the AE energy increased as the displacement rate increased.

3.5 Time evolution characteristics of the amplitude

Amplitude, which reflects the size of acoustic emission events, is used to evaluate the sources and magnitude of acoustic emissions. As shown in Figs. 8a-c, the acoustic emission signals exhibited low amplitude values during the pre-peak period. As each acoustic emission signal peaked, the amplitude increased significantly. Sandstone is composed of conglomerated particles. The cracks that occurred in the samples primarily occurred along these grains. Thus, low cement strength was associated with weak AE signals. During peak periods of stress, the shear stress increased, resulting in the propagation of micro-cracks throughout the particles. The formation of these cracks resulted in the rupture of particles and, thereby, the release of strong acoustic emission signals. The amplitude exhibited two peaks, which were determined based on the shear stress values.

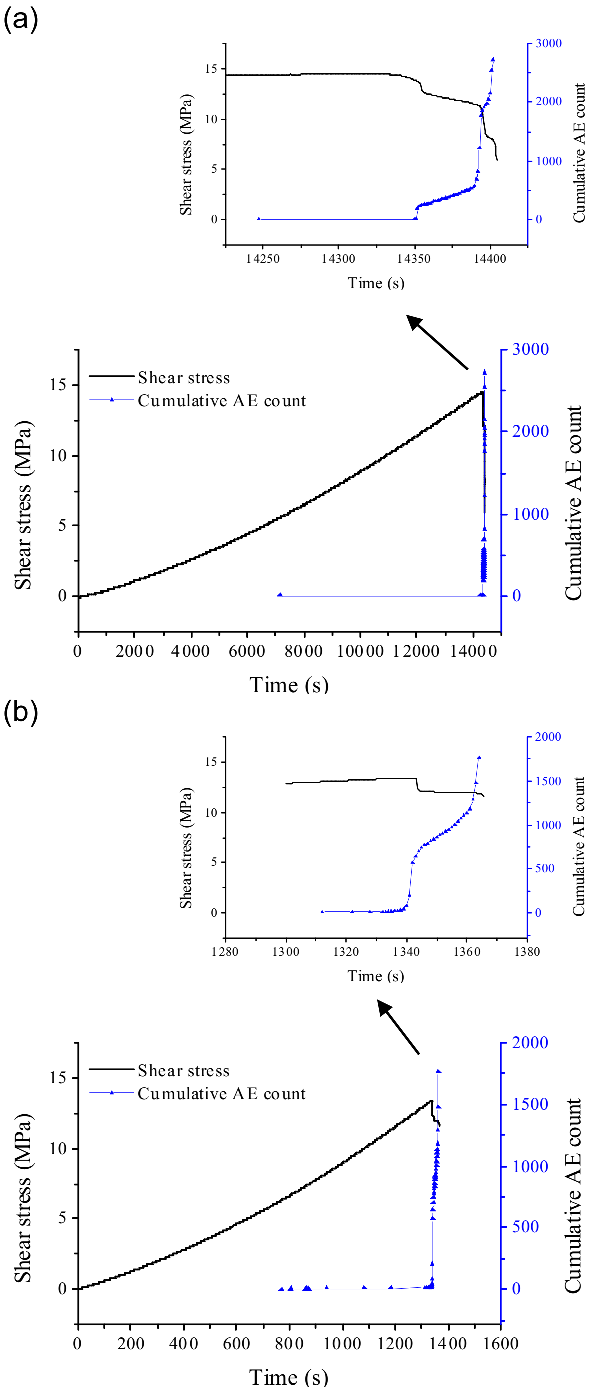
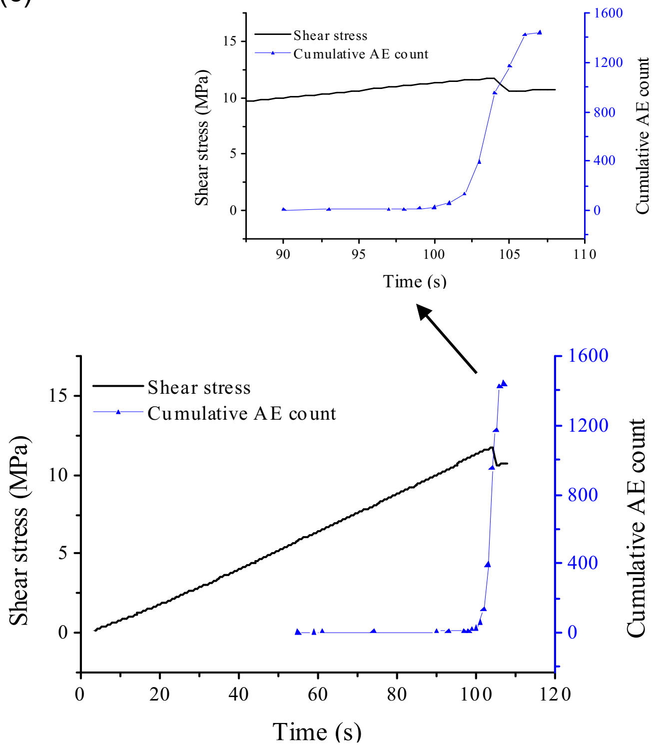


Fig. 6. Caption on next page.

(c)



(d)

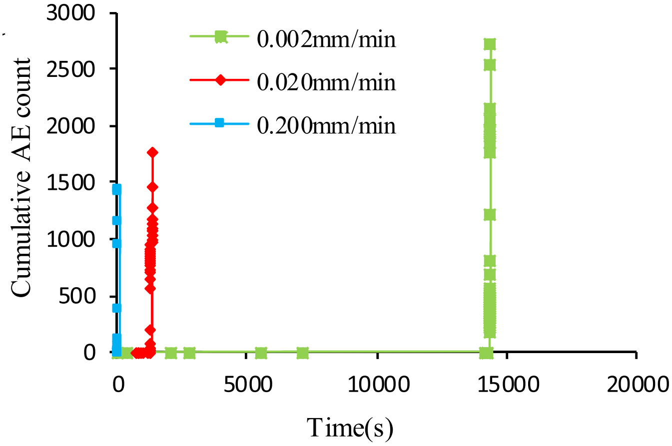


Fig. 6. Shear stress *versus* cumulative AE count under different displacement rates over time: (a) 0.002 mm/min, (b) 0.02 mm/min, (c) 0.2 mm/min, and (d) contrast curve.

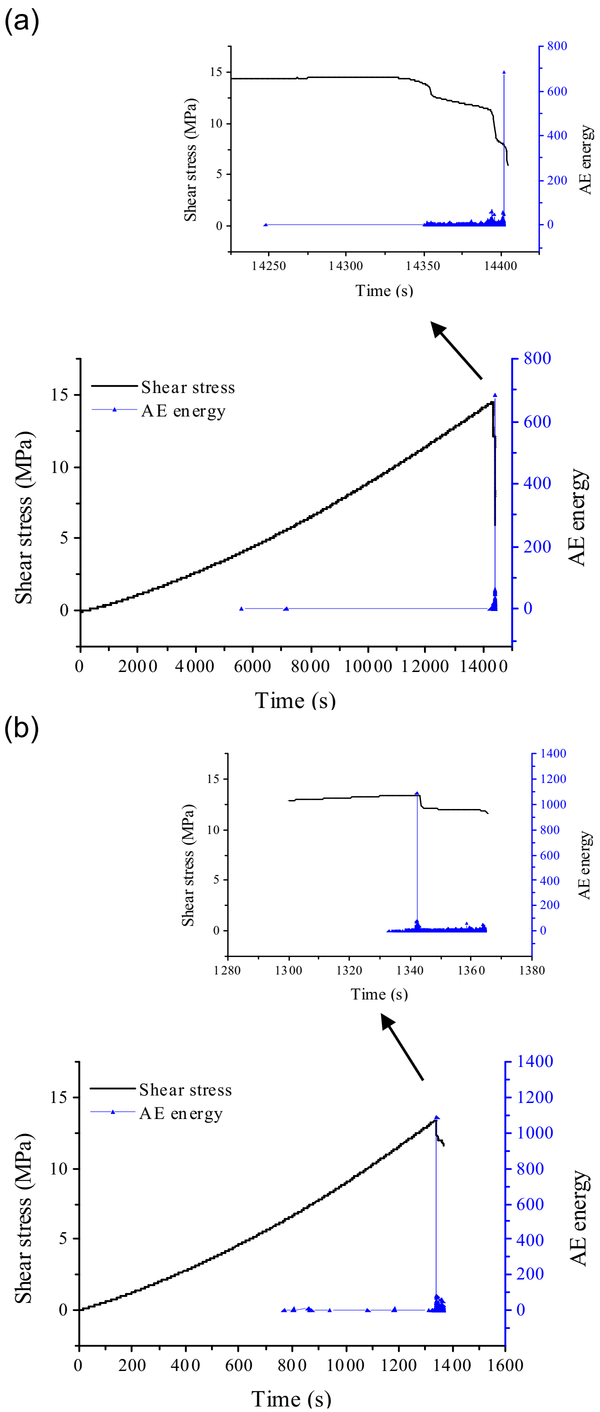
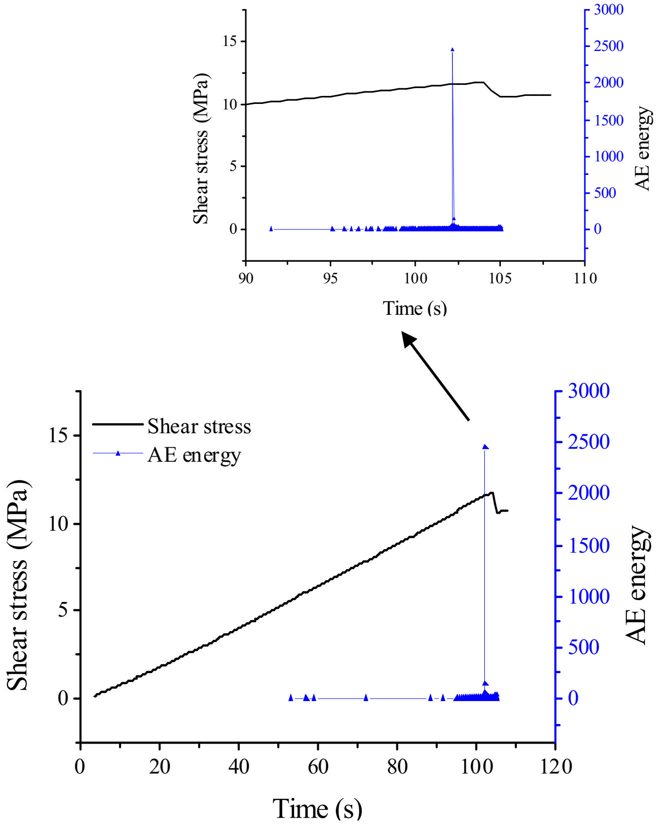


Fig. 7. Caption on next page.

(c)



(d)

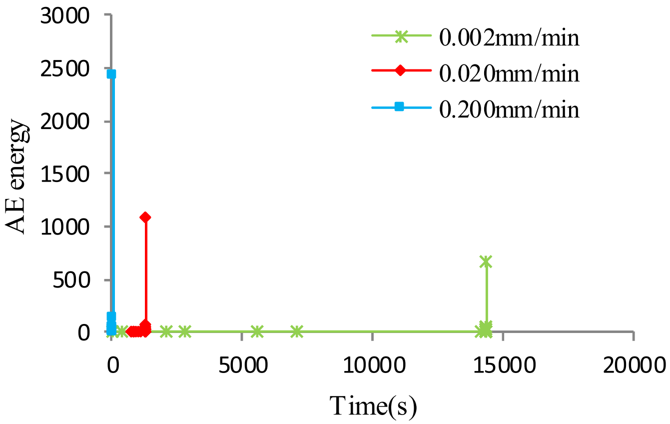


Fig. 7. Shear stress *versus* AE energy under different displacement rates over time: (a) 0.002 mm/min, (b) 0.02 mm/min, (c) 0.2 mm/min, and (d) contrast curve.

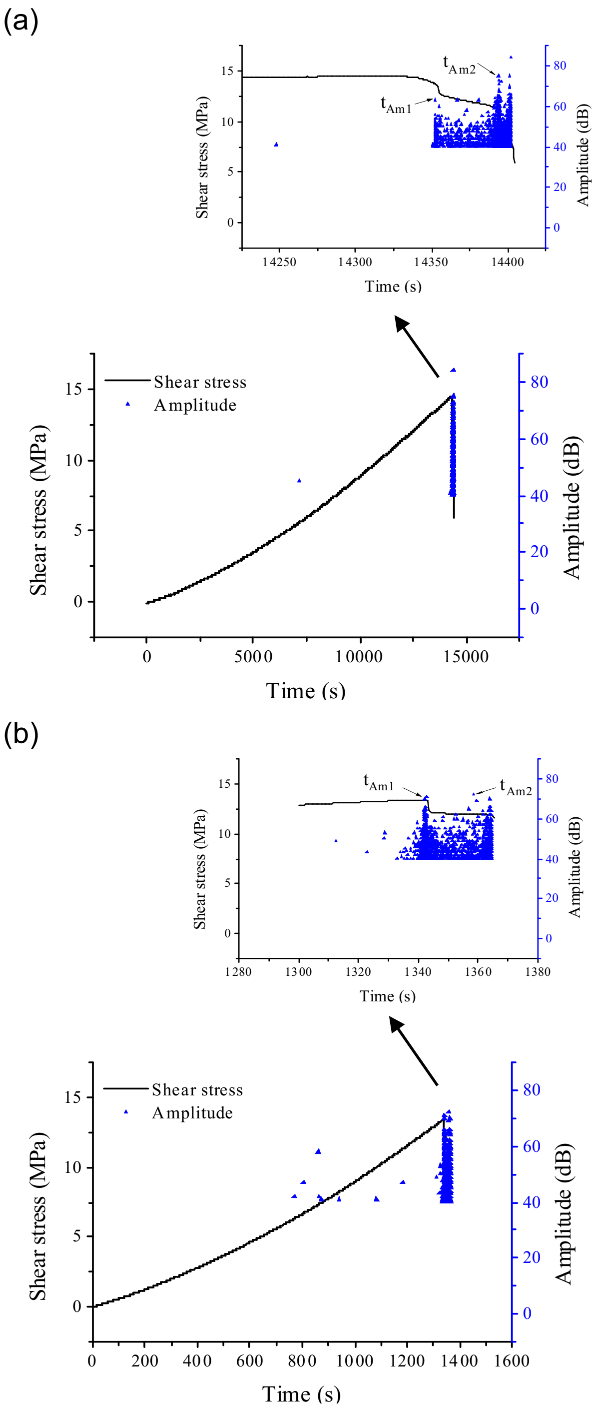


Fig. 8. Caption on next page.

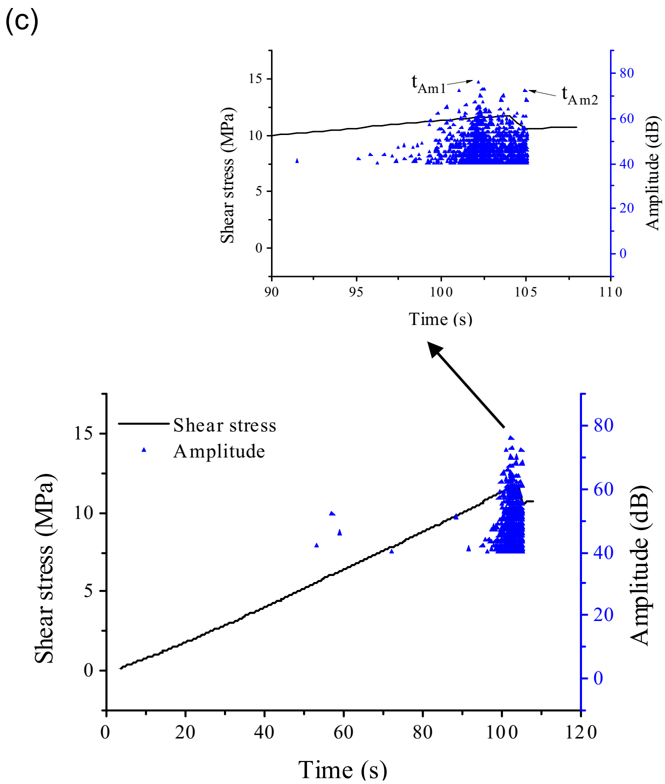


Fig. 8. Shear stress and amplitude *versus* displacement rate over time: (a) 0.002 mm/min, (b) 0.02 mm/min, and (c) 0.2 mm/min.

4. DISSCUSION

Under shear loading, internal stress redistributions in sandstone inevitably result in concentrated stress and, thereby, micro-crack initiation and propagation accompanied by the release of energy. Thus, internal damage can be analyzed based on the elastic waves received by acoustic emission transducers. Elastic waves are caused by the release of energy. In practical engineering applications, monitoring the stress that occurs in rocks as a result of routine operations is difficult. However, rock instability and failure could be reflected by monitoring these changes in acoustic emission signals. Table 2 displays the relationship between the AE signal peak time and specimen fracture time. As shown in this table, as the displacement rate decreased, the amount of time between the first AE signal peak and fracture increased. In addition, the second AE signal peak was more closely correlated with the time of fracture. This second AE signal peak could be used to identify specimen failure. However, in practical engineering applications, failure

Table 2

Acoustic emission signal peak times and specimen fracture times

V [mm/min]	t_C [s]	t_{AE1}/t_C-t_{AE1} [s]	t_{AE2}/t_C-t_{AE2} [s]	t_{Am1}/t_C-t_{Am1} [s]	t_{Am2}/t_C-t_{Am2} [s]	t_E/t_C-t_E [s]
0.002	14401	14352/+49	14394/+7	14352/+49	14402/–1	14402/–1
0.020	1363	1342/+21	1364/–1	1342/+21	1358/+5	1342/+21
0.200	107	102/+5	104/+3	102/+5	105/+2	104/+3

Notes: t_{AE1} denotes the first AE count rate peak, t_{AE2} denotes the second AE count rate peak, t_{Am1} denotes the first amplitude peak, t_{Am2} denotes the second amplitude peak, t_E denotes the AE energy peak, t_C denotes the fracture time, + denotes values earlier than the fracture time, and – denotes values later than the fracture time.

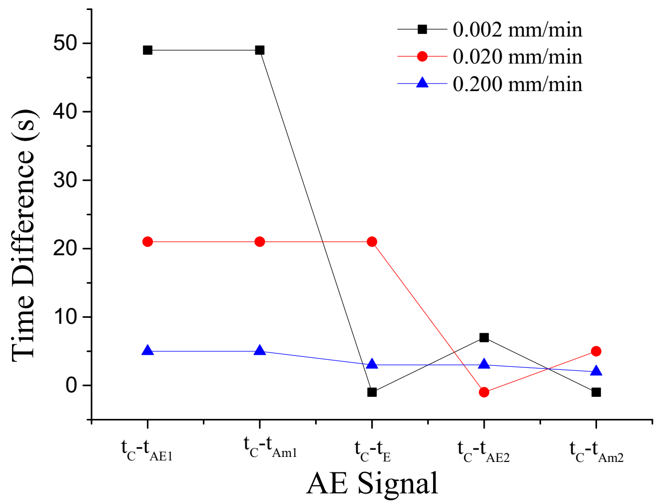


Fig. 9. Relationship between the acoustic emission signal peak time and specimen fracture time.

prevention is more important than failure identification. Therefore, according to the data presented in Fig. 9, the first AE signal peak would be more appropriate for reflecting rock failure in practical applications.

As shown in Table 3, the differences between the shear displacements of the two AE count rate and amplitude peaks were small. The grain size distribution of the sandstone samples ranged from 0.1 to 0.5 mm. When the displacement rate was 0.002 mm/min, the ratios between the shear displacements of the different acoustic emission signal peaks and their particle diameters ranged from 0.2~1.7%. As the displacement rate increased, these ratios

Table 3

Shear displacement values of the acoustic emission signal peaks

V [mm/min]	t_{AE1} [s]	t_{AE2} [s]	Δs_{AE} [mm]	t_{Am1} [s]	t_{Am2} [s]	Δs_{Am} [mm]
0.002	14352	14394	0.0014	14352	14402	0.0017
0.020	1342	1364	0.0073	1342	1358	0.0053
0.200	102	104	0.0067	102	105	0.0100

Notes: t_{AE1} denotes the first AE count rate peak, t_{AE2} denotes the second AE count rate peak, Δs_{AE} denotes the difference between the shear displacement values of the two AE count rate peaks, t_{Am1} denotes the first amplitude peak, t_{Am2} denotes the second amplitude peak, and Δs_{Am} denotes the difference between the shear displacement values of the two amplitude peaks.

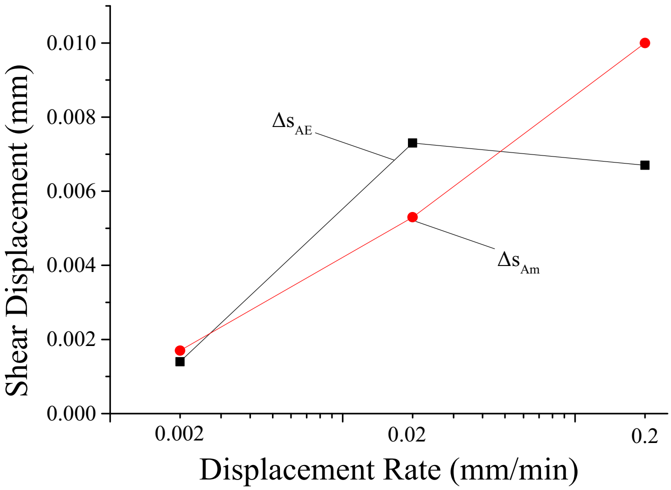


Fig. 10. Shear displacement values of the acoustic emission signal peaks.

also increased. According to Fig. 10 and the stress propagation mechanism (Havaej *et al.* 2013), the acoustic emission signal of the first peak in sandstone results from the propagation and coalescence of cracks, which cause slight deformations. These deformations result in instantaneous reductions in shear stress. As the shear displacement increased, the isolated cracked formed larger fractures, eventually resulting in rupture. The difference between the shear displacements of the two amplitude peaks also increased linearly as the displacement rate increased.

5. CONCLUSIONS

In this paper, acoustic emission signals were used to effectively monitor the shear behaviors of sandstone specimens under different displacement rates. Direct shear tests were conducted in order to investigate the acoustic emission characteristics of the sandstone under various shear loading conditions. The results indicated that AE signals could be applied to the real-time monitoring and reflection of sandstone rock failure.

The shear stress of the sandstone specimens decreased significantly in a relatively short amount of time during shear failure. These changes were reflected by the AE count rates and amplitude curves of the specimens. The AE signals corresponded with the amount of shear stress under various displacement rates. In addition, as the displacement increased, the amount of cumulative damage to each specimen decreased, and the AE energy peaks became larger and occurred earlier in time. The cumulative AE count primarily increased during the post-peak period. This indicated that higher displacement rates were associated with larger cracks in the samples (Kranz 1979). In contrast, lower displacement rates were primarily associated with the initiation and propagation of micro-cracks.

According to the comprehensive analysis of the AE signals, the time at which the second peak in amplitude occurred corresponded closely with the time of fracture. Thus, count rate peaks and initial peaks in amplitude could be a reference in practical engineering applications to provide warnings regarding rock failure.

Acknowledgments. The authors would like to thank the General Project of the National Natural Science Foundation of China (51474040, 51434003) for their financial support.

References

Chang, S.H., and C.I. Lee (2004), Estimation of cracking and damage mechanisms in rock under triaxial compression by moment tensor analysis of acoustic emission, *Int. J. Rock Mech. Mining Sci.* **41**, 7, 1069-1086, DOI: 10.1016/j.ijrmms.2004.04.006.

Goszczyńska, B., G. Świt, W. Trąmpczyński, A. Krampikowska, J. Tworzewska, and P. Tworzewski (2012), Experimental validation of concrete crack identification and location with acoustic emission method, *Arch. Civil Mech. Eng.* **12**, 1, 23-28, DOI: 10.1016/j.acme.2012.03.004.

Havaej, M., A. Wolter, D. Stead, Z. Tuckey, L. Lorig, and E. Eberhardt (2013), Incorporating brittle fracture into three-dimensional modelling of rock slopes. **In:** *Proc. Int. Symp. Slope Stab., Brisbane, Australia.*

Katsuyama, K. (1996), *Application of AE Techniques*, Metallurgy Industry Press, Bejing.

Kawamato, T., and T. Saito (1974), The behaviour of rock like materials in some controlled strain rates. **In:** *III Congress ISRM, Denver, USA*, Vol. 2, 161-166.

Kranz, R.L. (1979), Crack growth and development during creep of Barre granite, *Int. J. Rock Mech. Mining Sci. Geomech. Abstr*. **16**, 1, 23-35, DOI: 10.1016/0148-9062(79)90772-1.

Lajtai, E.Z., E.J.S. Duncan, and B.J. Carter (1991), The effect of strain rate on rock strength, *Rock Mech. Rock Eng.* **24**, 2, 99-109, DOI: 10.1007/BF01032501.

Lavrov, A. (2001), Kaiser effect observation in brittle rock cyclically loaded with different loading rates, *Mech. Mater.* **33**, 11, 669-677, DOI: 10.1016/S0167-6636(01)00081-3.

Li, Y.H., J.P. Liu, X.D. Zhao, and Y.J. Yang (2010), Experimental studies of the change of spatial correlation length of acoustic emission events during rock fracture process, *Int. J. Rock Mech. Min. Sci*. **47**, 8, 1254-1262, DOI: 10.1016/j.ijrmms.2010.08.002.

Lockner, D. (1993), The role of acoustic emission in the study of rock fracture, *Int. J. Rock Mech. Min Sci. Geomech. Abstr*. **30**, 7, 883-899, DOI: 10.1016/0148-9062(93)90041-B.

Majewska, Z., and Z. Mortimer (2006), Chaotic behaviour of acoustic emission induced in hard coal by gas sorption-desorption, *Acta Geophys*. **54**, 1, 50-59, DOI: 10.2478/s11600-006-0005-z.

Moradian, Z.A., G. Ballivy, P. Rivard, C. Gravel, and B. Rousseau (2010), Evaluating damage during shear tests of rock joints using acoustic emissions, *Int. J. Rock Mech. Min. Sci*. **47**, 4, 590-598, DOI: 10.1016/j.ijrmms.2010.01.004.

Rao, M.V.M.S., and K. Kusunose (1995), Failure zone development in andesite as observed from acoustic emission locations and velocity changes, *Phys. Earth Planet Int*. **88**, 2, 131-143, DOI: 10.1016/0031-9201(94)02967-G.

Ray, S.K., M. Sarkar, and T.N. Singh (1999), Effect of cyclic loading and strain rate on the mechanical behaviour of sandstone, *Int. J. Rock Mech. Min. Sci.* **36**, 4, 543-549.

Stock, G.M., S.J. Martel, B.D. Collins, and E.L. Harp (2012), Progressive failure of sheeted rock slopes: the 2009-2010 Rhombus Wall rock falls in Yosemite Valley, California, USA, *Earth Surf. Proc. Landforms* **37**, 5, 546-561, DOI: 10.1002/esp.3192.

Taheri, A., and K. Tani (2010), Assessment of the stability of rock slopes by the slope stability rating classification system, *Rock Mech. Rock Eng*. **43**, 3, 321-333, DOI: 10.1007/s00603-009-0050-4.

Xu, J., S. Li, Y. Tao, X. Tang, and X. Wu (2009), Acoustic emission characteristic during rock fatigue damage and failure, *Proc. Earth Planet. Sci*. **1**, 1, 556-559, DOI: 10.1016/j.proeps.2009.09.088.

Xu, J., S. Peng, G. Yin, H. Yang, and W. Wang (2011), Development of mesoshear test equipment for coal rock containing gas and its application, *Chin. J. Rock Mech. Eng.* **30**, 4, 677-685.

Yang, S.Q., and H.W. Jing, and S.Y. Wang (2012), Experimental investigation on the strength, deformability, failure behavior and acoustic emission locations of red sandstone under triaxial compression, *Rock Mech. Rock Eng.* **45**, 4, 583-606, DOI: 10.1007/s00603-011-0208-8.

Zhao, X., Y. Li, R. Yuan, T. Yang, J. Zhang, and J. Liu (2007), Study on crack dynamic propagation process of rock samples based on acoustic emission location, *Chin. J. Rock Mech. Eng.* **26**, 5, 944-950.

Zhurkov, S.N. (1965), Kinetic concept of the strength of solids, *Int. J. Fract.* **4**, 311-323.

Received 25 November 2015
Received in revised form 12 April 2016
Accepted 5 May 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2430-2448
DOI: 10.1515/acgeo-2016-0104

Applying the General Regression Neural Network to Ground Motion Prediction Equations of Induced Events in the Legnica-Głogów Copper District in Poland

Jan WISZNIOWSKI

Institute of Geophysics, Polish Academy of Sciences, Warsaw, Poland
e-mail: jwisz@igf.edu.p

Abstract

This paper presents a study of the nonlinear estimation of the ground motion prediction equation (GMPE) using neural networks. The general regression neural network (GRNN) was chosen for its high learning rate. A separate GRNN was tested as well as a GRNN in cascade connection with linear regression (LR). Measurements of induced seismicity in the Legnica-Głogów Copper District were used in this study. Various sets of input variables were tested. The basic variables used in every case were seismic energy and epicentral distance, while the additional variables were the location of the epicenter, the location of the seismic station, and the direction towards the epicenter. The GRNN improves the GMPE. The best results were obtained when the epicenter location was used as an additional input. The GRNN model was analysed for how it can improve the GMPE with respect to LR. The bootstrap resampling method was used for this purpose. It proved the statistical significance of the improvement of the GMPE. Additionally, this method allows the determination of smoothness parameters for the GRNN. Pa-

Ownership: Institute of Geophysics, Polish Academy of Sciences

rameters derived through this method have better generalisation capabilities than the smoothness parameters estimated using the holdout method.

Key words: ground motion prediction equation, artificial neural network, general regression neural network.

1. INTRODUCTION

Since the beginning of seismic hazard analysis, ground motion prediction equations (GMPEs) have been effectively used to estimate ground motions in deterministic and probabilistic seismic hazard analyses. The basic method of estimation of GMPE is linear regression (LR), where magnitude and the logarithm of distance are independent variables, and the logarithm of peak ground acceleration is the dependent variable (Cornell 1968). Instead of magnitude, the logarithm of energy can also be used (*e.g.*, Golik and Mendecki 2012, Lasocki 2013).

While there are GMPEs with more expanded formulae (*e.g.*, Douglas 2011), which incorporate higher powers of magnitude (*e.g.*, Trifunac and Brady 1976, Joyner and Boore 1988, Akkar and Bommer 2010) and complex functions incorporating source models (*e.g.*, Abrahamson and Silva 2008, Boore and Atkinson 2008, Campbell and Bozorgnia 2008), the use of artificial neural networks (ANNs) to estimate the GMPE does not require knowledge of these models.

Basically, GMPEs vary considerably depending on the seismic zone (Douglas 2011), and it is common to perform GMPE estimation for each particular case individually. For induced seismicity, the most effective GMPEs are often basic LR models (Golik and Mendecki 2012, Lasocki 2013).

ANNs have been chosen as a tool to answer the question of whether using nonlinear regression can help create a model that results in a significantly better description of the predicted mean of the GMPE.

For nonlinear estimation, three types of neural networks are used most frequently: multilayer perceptrons (MLPs) with one or two hidden layers (Pozos-Estrada *et al.* 2014), networks with radial based functions (RBFs), and general regression neural networks (GRNNs). All three solutions have been applied to the GMPE (Günaydin and Günaydin 2008), with MLP being used most often (Güllü and Erçcelebi 2007, Derras *et al.* 2012, Hong *et al.* 2012, Arjun 2013). The magnitude and epicentral or hypocentral distance have been used as input data for the GMPE estimation with the help of ANNs. The additional input data that have also been used are: focal depth (Pozos-Estrada *et al.* 2014); shear wave velocity down to 30 m (V_{S30}) and focal depth (Hong *et al.* 2012, Derras *et al.* 2012); the thickness of the sedimentary layers below the site down to a shear wave velocity equal to

800 m/s, and the corresponding resonant frequency (Derras and Bekkouche 2011); the fundamental resonance frequency, as determined by the horizontal-to-vertical (H/V) spectral ratio technique (Derras *et al.* 2012); site conditions (Güllü and Erçelebi 2007, Günaydin and Günaydin 2008); and the average values of shear wave velocity, primary wave velocity, standard penetration test blow count, and the density of soil (Arjun 2013).

For this paper we chose the GRNN method, because it is fast-learning and trends toward the optimal regressional surface. In the case of testing the usefulness of ANN for GMPE estimation, using the GRNN gives a quicker result, without the need to analyse whether the ANN is optimally trained.

Recently, GRNN networks have been applied for solving different problems in the fields of seismic hazard and earthquake engineering. García *et al.* (2003) and Derras and Bekkouche (2011) applied GRNN together with other ANN methods for estimation of the GMPE. Yaghmaei-Sabegh and Tsang (2011, 2014) applied the GRNN as well as probabilistic neural network (PNN) for site classification based on an H/V spectral ratio technique. Yaghmaei-Sabegh (2012) also employed the GRNN for ranking and weighting the GMPE models in the logic tree. Hanna *et al.* (2007) employed the GRNN for assessing liquefaction in soil deposits.

Although earlier studies have not shown any better effectiveness of the GMPE estimated by GRNN than the MLP (García *et al.* 2003, Derras and Bekkouche 2011), the possibility of quick learning and the uniqueness of training allow the examination of the distribution of GMPE model error using the bootstrap method, and testing the statistical significance of the improvement of the GMPE. The effectiveness of the GMPE estimated by GRNN is improved by applying the GRNN in connection with the LR, where the GRNN reduces the error of LR.

Data on induced seismicity in the Legnica-Głogów Copper District (LGCD) in Poland were used in this study. Tremors with energy of up to 10^9 J occur in this region, and generate motions of up to 3 m/s^2. The GMPEs for this region have been previously estimated based on LR methods (Lasocki 2013).

2. THE GRNN

The concept of the GRNN (Specht 1991) is based on computing the conditional mean of y, given X, according to the formula

$$E(y \mid X) = \frac{\int_{-\infty}^{+\infty} y f(X \mid y)\, dy}{\int_{-\infty}^{+\infty} f(X \mid y)\, dy} , \tag{1}$$

where X is a particular measured value of the random variable x and $f(X|y)$ is the condition probability density function of X given y. Because $f(X|y)$ is not known, usually the estimator proposed by Parzen (1962) is used.

In the implementation of the GRNN, the integration is performed by summation, according to the formula:

$$\hat{Y}(X) = \frac{\sum_{i=1}^{n} Y_i p(X \mid Y_i)}{\sum_{i=1}^{n} p(X \mid Y_i)}, \tag{2}$$

where n is the number of samples used for learning.

The GRNN consists of four layers (Fig. 1): input, pattern, summation and output. The input layer is made up of units corresponding to each of the independent variables used to estimate the ground motion. The number of units in the pattern layer is the same as the number of samples in the training set. The i^{th} unit in the pattern layer, which corresponds to the i^{th} sample in the training set, computes $p(\boldsymbol{x}|y_i)$ according to the formula:

$$p(x \mid y_i) = \exp\left(\frac{D_i^2}{\sigma^2}\right), \tag{3}$$

where σ is the smoothness parameter of the GRNN and D_i is the distance between the i^{th} sample and the value in the input layer. The summation layer consists of two units that compute the numerator

$$S = \sum_{i=1}^{n} y_i p(x \mid y_i) \tag{4}$$

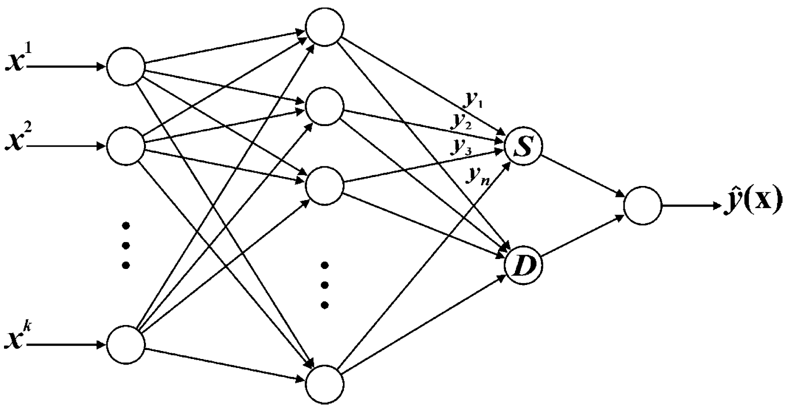


Fig. 1. Diagram of the GRNN.

and the denominator:

$$D = \sum_{i=1}^{n} p\left(x \mid y_i\right). \tag{5}$$

The output layer realises the division:

$$\hat{y}\left(x\right) = \frac{S}{D}. \tag{6}$$

The smoothness parameter σ is the only value selected during the network learning process. The holdout method proposed by Specht (1991) consists of removing one sample at a time and constructing a network based on all of the other samples. Then, the network is used to calculate the squared error between the expected Y and the estimated $\hat{Y}$ for the removed sample. By repeating this process for each sample, and for the particular value of σ, the sum of squared errors (SSE) for all samples is calculated. The value of σ with the smallest SSE is used in the GRNN. In the following sections, the other method for determining σ is proposed. It is based on the bootstrap method.

As a preprocessing step, we scale all input variables such that they have approximately the same ranges or variances. The need for this process stems from the fact that the probability density function $f(X|y)$ is to be estimated using (3) with only one smoothness parameter (Specht 1991). Therefore, the metric in the form

$$D_i^2 = \sum_{j=1}^{k} \frac{\left(x^j - x_i^j\right)^2}{\mathrm{Var}\left(X^j\right)} \tag{7}$$

was used, where k is the number of input variables (Fig. 1).

The GRNN has advantages and disadvantages when compared to MLPs. The GRNN has a solid mathematical background (Wasserman 1993). It is a one-pass learning algorithm (*i.e.*, lazy learning). Prediction is unique, which means that it is not dependent on the training procedure or the initial conditions. On the other hand, the GRNN requires more memory space to store the model and is slower than MLPs when classifying new cases.

3. DATA USED TO TEST THE GRNN

Copper ore extraction by three underground mines in the LGCD in south-west Poland, which takes place in hard rocks at depths of 800-1100 m, generating earthquakes that can exceed 4.5 of local magnitude and 10^9 J of energy, often significantly affects buildings and other surface structures. The

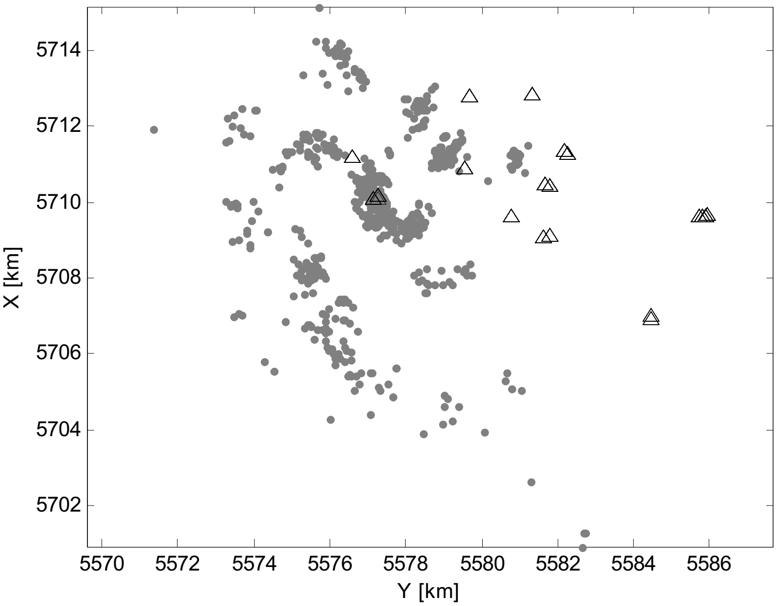


Fig. 2. Distribution of earthquakes recorded in the LGCD, 2001–2012. Dots – earthquakes, triangles – stations.

strongest events are capable of producing a peak ground acceleration (PGA) of more than 3 m/s^2.

A GMPE is a function of values describing an event, site, and source-to-site route. For the LGCD, only data on ground motion, seismic energy, the location of the epicenter, and the location of the station are provided and can be used for estimation of the GMPE. The data consisted of 2991 ground motion measurements of 904 events recorded by up to 11 accelerometers (Fig. 2) in the period from April 2001 to November 2012. For the GMPE estimation, we used values that were functions of the original values:

- $\log E$ – the logarithm of energy,
- $\log R$ – the logarithm of distance, where

$$R = \sqrt{r^2 + h_0^2}\ , \tag{8}$$

and r is the source-receiver epicentral distance calculated from the coordinates of the epicenter and station, whereas the coefficient h_0 is the common depth factor introduced by Joyner and Boore (1993). The value $h_0 = 800$ was chosen by minimising the error of linear regression of $\log E$ and $\log R$. The additional input variables are:

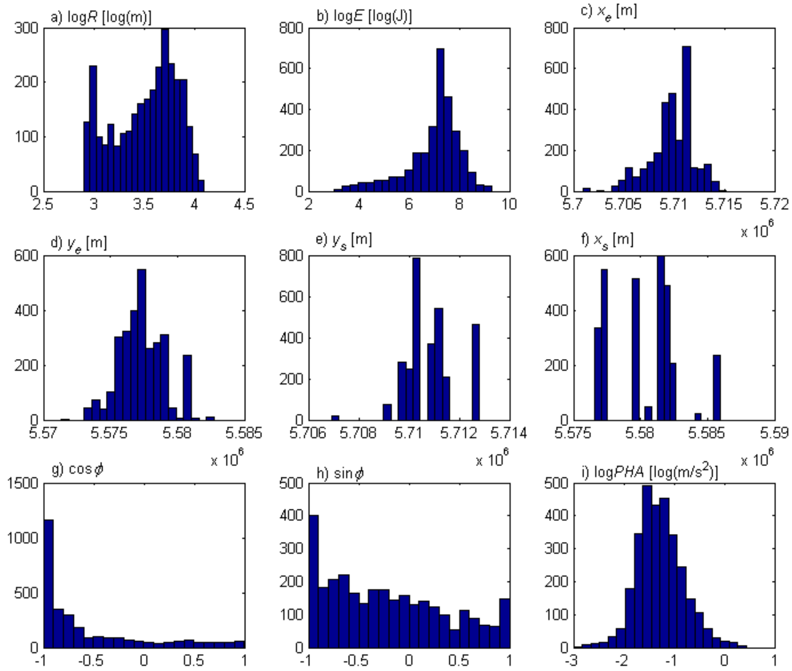


Fig. 3. Histograms of inputs (a-h) and output of the GRNN (i): (a) logarithms of distances; (b) logarithms of energies; (c & d) coordinates of the epicenters; (e & f) coordinates of the seismic stations; (g) cosines of radial directions; (h) sines of radial directions; (i) logarithms of peak horizontal accelerations.

- x_e, y_e – the coordinates of the epicenters,
- x_s, y_s – the coordinates of the seismic stations,
- $\cos\varphi$, $\sin\varphi$ – the cosines and sines of radial direction – instead of $\cos\varphi$ and $\sin\varphi$, angles φ could be used as well; however, the cosine and sine functions vary smoothly when the angle changes from 360° to 0°.

The output value was the logarithm of peak horizontal component of ground acceleration (*i.e.*, Peak Horizontal Acceleration, *PHA*). Figure 3 shows individual histograms of these values, while Fig. 2 better illustrates the distribution of coordinates (x_e, y_e and x_s, y_s) as they are, in fact, 2D variables.

The dataset is divided randomly into training and test sets during the bootstrap test. Any outliers are not removed.

4. APPLYING THE GRNN TO THE GMPE

The GRNN was tested both separately and in cascade with an LR estimator. Several GMPE models were compared (Fig. 4).

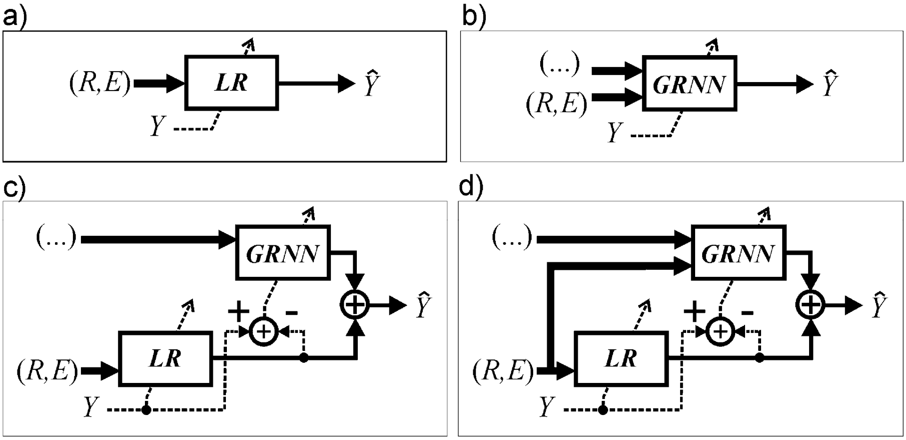


Fig. 4. Diagrams of the GMPE models: (a) LR of $\log R$ and $\log E$; (b) separate GRNN, where $\log R$ and $\log E$ are some of the inputs; (c) GRNN in cascade with LR, where $\log R$ and $\log E$ are not inputs of the GRNN; (d) GRNN in cascade with LR, where $\log R$ and $\log E$ are inputs of the GRNN.

The reference model is the LR model (Fig. 4a), which is commonly used for GMPE (Douglas 2011), especially in the case of mining induced seismicity (Golik and Mendecki 2012, Lasocki 2013). The effectiveness of all GRNN models is compared to that of the LR. The considered form of the LR model is:

$$\log a = \alpha + \beta \log E + \gamma \log R \ , \tag{9}$$

where a is the PHA in m/s^2 estimated from the model. The form of LR was chosen because it was applied to the GMPE in the LGCD by Lasocki (2013). The estimated parameters of LR were:

$$\begin{gathered} \alpha = 0.41, \\ \beta = 0.47, \\ \gamma = -1.42, \\ R^2 = 0.731. \end{gathered} \tag{10}$$

The value of the R^2 is typical of mining induced GMPEs. The coefficient of determination obtained by Lasocki (2013), for the LR model covering different parts of the LGCD and estimated from a smaller amount of data (1818 records), was lower (R^2 = 0.532). In the case of induced seismicity in the Upper Silesian Coal Basin in Poland (Golik and Mendecki 2012), the coefficients of determination of GMPE models are in the range of 0.64 (region of "Bielszowice" coal mine) to 0.79 (region of "Ziemowit" coal mine). The error of LR (10) model does not have a normal distribution (Fig. 5). It did not pass Lilliefors's normality test ($p = 0.007$).

The next tested model was the separate GRNN (Fig. 4b). Its inputs were log R, log E, and combinations of the following pairs of variables (described as '(…)' in Fig. 4):

$$(x_e, y_e), (x_s, y_s), (\cos\varphi, \sin\varphi). \tag{11}$$

However, log R, cos φ and sin φ are functions of x_e, y_e and x_s, y_s. Therefore, another GMPE model, in the form of the GRNN(log E, x_e, y_e, x_s, y_s), was also analysed. It was the only model in which the distance was not given as log R.

The models in Figs. 4c and 4d are combinations of LR and the GRNN. Both models consist of two components in cascade. The first component is the LR. Its inputs are only log R and log E. Its output is the linear prediction of ground motion. The second component is the GRNN. It is trained by residuals of the LR. The output of the whole model is the sum of the outputs of the LR and GRNN. The GMPE model, in which the GRNN improves results of the LR, is referred to as GRNN_R. In the model presented in Fig. 4c, the GRNN input incorporates only combinations of inputs (11), whereas in the

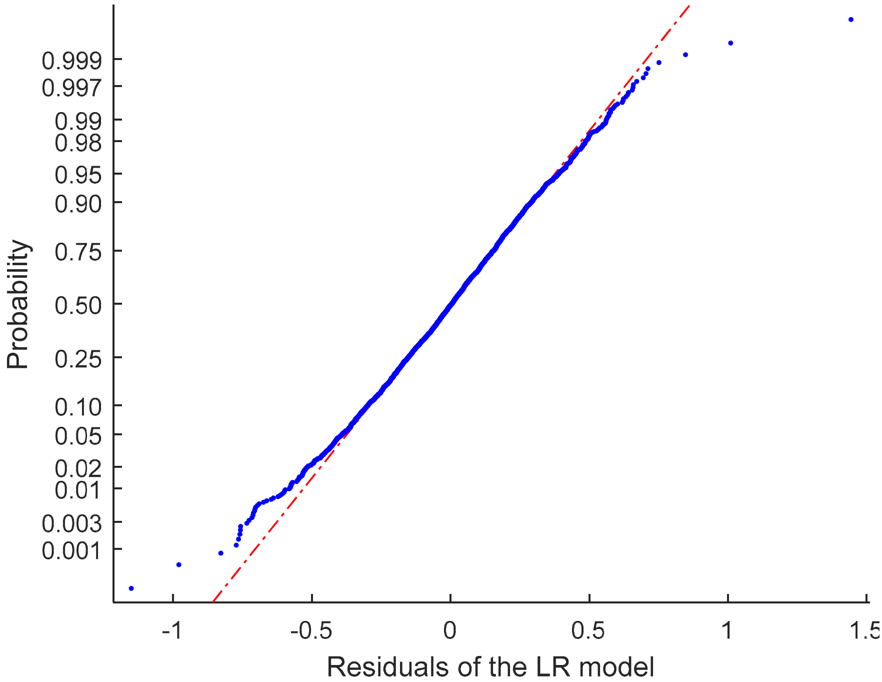


Fig. 5. Normal probability plot for the LR model.

model presented in Fig. 4d, the inputs log R and log E are also included in GRNN. As the smoothness parameter increases, the output of the GRNN approaches the mean value of log PHA, whereas $GRNN_R$ approaches the result of the LR. Therefore, the $GRNN_R$ is less sensitive to large values of σ.

Table 1

Results of holdout method Optimum smoothness parameter values obtained with the help of the holdout method and corresponding residual sum of squares, and the coefficient of determination for various GMPE models utilising the GRNN and various combinations of inputs

GMPE model		σ	SS_{res}	R^2
Separate GRNN (Fig. 4b)	GRNN(log E, log R)	0.096	142.84	0.77
	GRNN(log E, log R, x_e, y_e, x_s, y_s, cosφ, sinφ)	0.235	115.23	0.81
	GRNN(log E, log R, x_e, y_e, x_s, y_s)	0.175	110.12	0.82
	GRNN(log E, log R, x_e, y_e, cosφ, sinφ)	0.224	113.57	0.82
	GRNN(log E, log R, x_s, y_s, cosφ, sinφ)	0.222	118.76	0.81
	GRNN(log E, log R, x_e, y_e)	0.138	104.14	0.83
	GRNN(log E, log R, x_s, y_s)	0.113	119.93	0.81
	GRNN(log E, log R, cosφ, sinφ)	0.166	129.77	0.79
$GRNN_R$ with inputs log E and log R (Fig. 4d)	GRNN(log E, x_e, y_e, x_s, y_s)	0.147	109.87	0.82
	$GRNN_R$(log E, log R)	0.100	142.19	0.77
	$GRNN_R$(log E, log R, x_e, y_e, x_s, y_s, cosφ, sinφ)	0.294	106.22	0.83
	$GRNN_R$(log E, log R, x_e, y_e, x_s, y_s)	0.180	104.21	0.83
	$GRNN_R$(log E, log R, x_e, y_e, cosφ, sinφ)	0.271	105.32	0.83
	$GRNN_R$(log E, log R, x_s, y_s, cosφ, sinφ)	0.264	111.41	0.82
	$GRNN_R$(log E, log R, x_e, y_e)	0.141	100.93	0.84
	$GRNN_R$(log E, log R, x_s, y_s)	0.114	120.16	0.81
	$GRNN_R$(log E, log R cosφ, sinφ)	0.183	125.05	0.80
$GRNN_R$ without inputs log E, and log R (Fig. 4c)	$GRNN_R$(x_e, y_e, x_s, y_s, cosφ, sinφ)	0.102	115.48	0.82
	$GRNN_R$(x_e, y_e, x_s, y_s,)	0.069	111.24	0.82
	$GRNN_R$(x_e, y_e, cosφ, sinφ)	0.097	115.80	0.82
	$GRNN_R$(x_s, y_s, cosφ, sinφ)	0.088	147.43	0.76
	$GRNN_R$(x_e, y_e)	0.012	112.13	0.82
	$GRNN_R$(x_s, y_s)	0.039	161.03	0.74
	$GRNN_R$(cosφ, sinφ)	0.107	163.99	0.74

The quality of methods is assessed based on the residual sum of squares, SS_{res}, and the coefficient of determination, R^2. SS_{res} is defined as

$$SS_{res} = \sum_{i=1}^{n} \left(\hat{y}_i - y_i\right)^2 , \tag{12}$$

where y_i is the measured output and $\hat{y}_i$ is the output of the GMPE model. R^2 is defined as

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} , \tag{13}$$

where SS_{tot} is the total sum of squares defined as:

$$SS_{tot} = \sum_{i=1}^{n} \left(y_i - \bar{y}\right)^2 , \tag{14}$$

and:

$$\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i . \tag{15}$$

Table 1 shows the SS_{res} and R^2 for the smoothness parameter that was selected using the holdout method (Specht 1991) for the analysed models and for the analysed GMPE inputs. The best prediction improvement for all the GRNN models was obtained when the epicenter of the earthquake was incorporated as an input value. Prediction was also improved when the only GRNN inputs were $\log R$ and $\log E$. The model with no $\log R$ as input (GRNN($\log E$, x_e, y_e, x_s, y_s)) also yielded good results. Using too many input parameters did not improve the GMPE. In the case of epicenter location (x_e, y_e), the quality of the results was reduced when the station location (x_s, y_s), or the angular coefficients ($\cos\varphi$, $\sin\varphi$), were included. However, the best results were obtained for $\mathrm{GRNN_R}$ with inputs $\log E$ and $\log R$, which suggests that adding the inputs $\log R$ and $\log E$ had no negative impact on prediction.

5. TESTING THE SIGNIFICANCE OF GMPE IMPROVEMENT FOR THE GRNN

Due to the differences in the LR and GRNN methodologies, a method based on the bootstrap (Efron 1979) was chosen to compare them. The full list of records was randomly divided into the subset for training the network and the test subset. The training subset contained one-quarter of all records. The remaining validation subset was used for estimation of the prediction error (Moody 1994). This operation was performed 1000 times for each GRNN model and for each set of inputs. The results are presented in Fig. 6. The bootstrap tests were performed for various values of σ that gave the distribution of R^2 as a function of σ. The mean values of R^2 of GRNNs are shown as solid lines, whereas the 5th and 95th percentiles of R^2 are shown as dotted lines.

The LR was estimated and tested in an analogous way, and the following result was obtained:

- The mean value of R^2 was 0.7281;
- The 5th percentile of R^2 was 0.7202;
- The 95th percentile of R^2 was 0.7334.

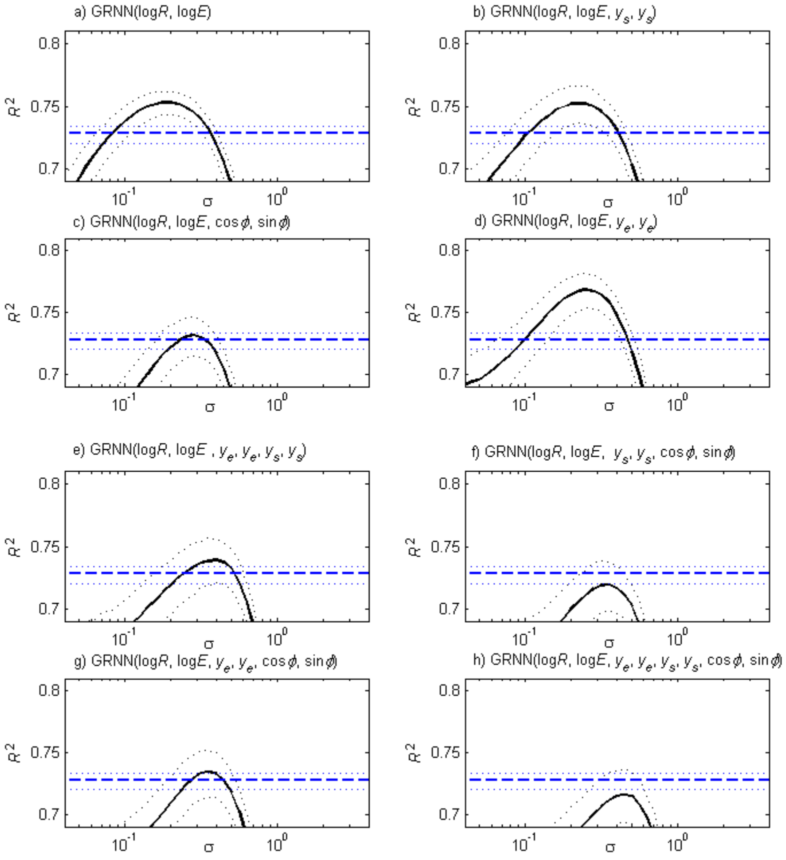


Fig. 6. Values of the coefficient of determination for bootstrap testing of the GRNN as a function of the smoothness parameter. Solid lines – mean R^2 calculated for the testing set (data not used for training the GRNN); dashed lines – mean R^2 calculated using LR; dotted lines – 5th and 95th percentiles of R^2. a-h: results for the separate GRNN (Fig. 4b). Continued on next page.

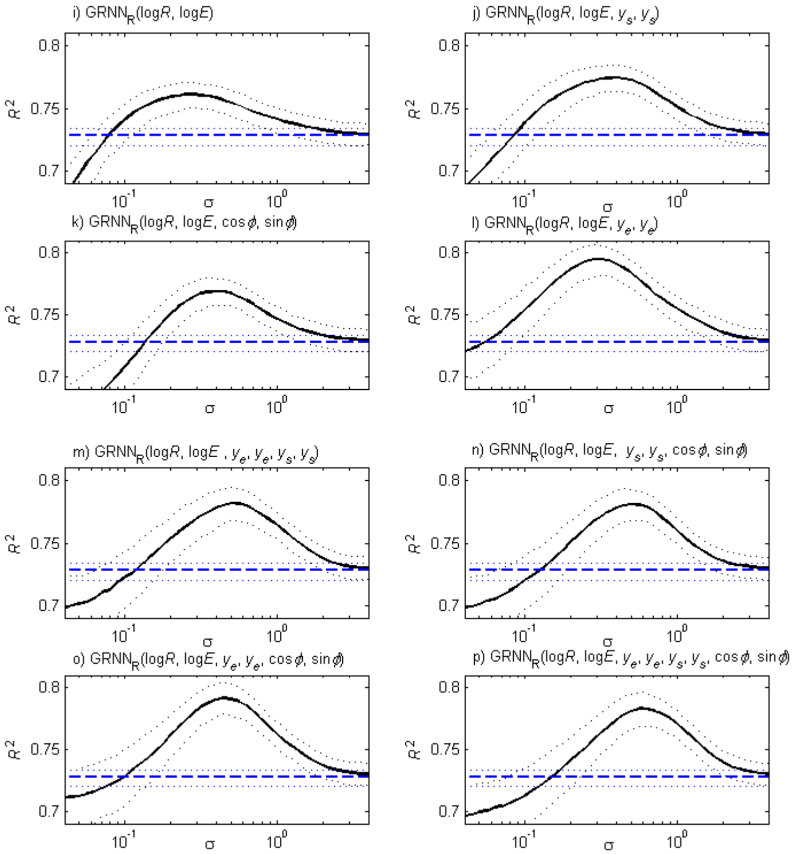


Fig. 6 continuation: Solid lines – mean R^2 calculated for the testing set (data not used for training the GRNN); dashed lines – mean R^2 calculated using LR; dotted lines – 5th and 95th percentiles of R^2. i-p: Results for the cascade GRNN with log E and log R as GRNN$_R$ inputs (Fig. 4d). Continued on next page.

The mean value of R^2 of the LR is shown as dashed lines in Fig. 6, whereas the 5th and 95th percentiles of R^2 of the LR are shown as dotted lines.

Marking the 5th percentile for the GRNN and the 95th percentile for the LR in Fig. 6 allows us to conclude whether or not the estimation made using the GRNN is statistically significantly better than that made using LR, and for which values of σ. If we take an optimum σ that maximises the 5th percentile of R^2, the bootstrap method can be used to assess the minimum GRNN improvement of the GMPE estimation with the probability is greater than 95%, and to select the corresponding value of σ.

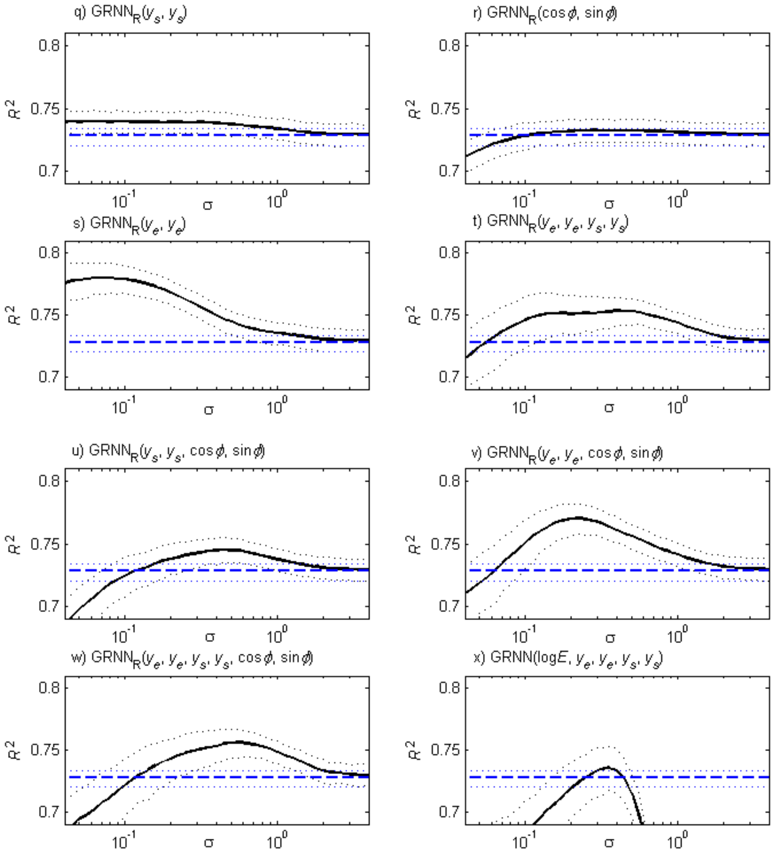


Fig. 6 continuation: Solid lines – mean R^2 calculated for the testing set (data not used for training the GRNN); dashed lines – mean R^2 calculated using LR; dotted lines – 5th and 95th percentiles of R^2. q-w: Results for the cascade GRNN without $\log E$ and $\log R$ as $GRNN_R$ inputs (Fig. 4c); x: Results for the GRNN replacing $\log R$ with station location and epicentral location as inputs.

The bootstrap analysis has confirmed that the best GMPE improvement is achieved when the epicenter location is used as an input (Figs. 6d, 6l, and 6s). The improvement of the GMPE is also significant when only $\log E$ and $\log R$ are inputs for the GRNN, both for the separate GRNN (Fig. 6a) and for the cascade mode with the LR (Fig. 6i). The information about the station location improves the GMPE most when it is combined with $\log E$ and $\log R$ (Figs. 6b and 6j).

The test using only the station location in cascade with LR ($GRNN_R(x_s, y_s)$) shows a small but statistically significant improvement of the GMPE

(Fig. 6q). To assess this case, a GMPE model with relative local amplification factors (*e.g.* Lasocki 2013) was analysed and estimated using LR according to the formula:

$$\log a_i = \alpha + \beta \log E + \gamma \log R + \sum_{j=1}^{J} w_j \delta_{i,j} \ , \tag{16}$$

where i is the station number, a_i are the *PHA* values recorded by the i^{th} station, w_i is the estimated amplification factor at the i^{th} station location, $\delta_{i,j}$ is Kronecker's delta, and $J = 19$ is the number of stations. The other symbols are the same as in (9). $R^2 = 0.747$ was calculated for the tested data using a GMPE model with relative local amplification factors. The result is comparable with $\text{GRNN}_{\text{R}}(x_s, y_s)$.

In particular, the bootstrap test did not show a significant improvement for the model GRNN(log E, x_{e}, y_{e}, x_s, y_s), which did not consider the distance from the earthquake as an input (Fig. 6x). The results yielded by the holdout method were very good, whereas the bootstrap test did not show a significant improvement. This model either has low generalisation capability, or it is very sensitive to the reduction of the training set.

The application of the angular coefficients was based on the assumption that the GMPE is related to direction. An improvement of the GMPE in models with angular coefficients as the only inputs would suggest a constant trend for the whole zone. On the other hand, an improvement of the model in the case of angular coefficients in combination with the epicenter location would confirm that the earthquake mechanisms depend on location. The test results (Figs. 6o and 6v) show an improvement in the GMPE when x_{e} and y_{e} are added as inputs along with cosφ and sinφ; however, the results are worse than for x_{e} and y_{e} alone. We can see, however, that $\text{GRNN}_{\text{R}}(\cos\varphi, \sin\varphi)$ requires a larger smoothness parameter (Fig. 6r) than $\text{GRNN}_{\text{R}}(x_{\text{e}}, y_{\text{e}})$ (Fig. 6s). In addition, R^2 for larger values of σ – optimum for $\text{GRNN}_{\text{R}}(x_{\text{e}}, y_{\text{e}}, \cos\varphi, \sin\varphi)$ (Fig. 6v) – is a little larger than R^2 in the model $\text{GRNN}_{\text{R}}(x_{\text{e}}, y_{\text{e}})$. Thus, it cannot be ascertained whether the poorer results of the GMPE for the inputs x_{e}, y_{e}, cosφ and sinφ are a result of the lack of dependency, or of the unmatched inputs for one smoothness parameter.

Applying more inputs does not improve the results. The GRNN is most effective when only one additional input is incorporated to log R and log E. In particular, combined information about the epicenters and directions did not improve the GMPE. It can be concluded that either focal directions in one place do not have any influence on the ground motion, or that this model of GRNN does not work properly with this combination of input variables. This could be overcome by using the input-dependent smoothing parameter, σ, which would further complicate the training of the GRNN, or by searching for a better metric D to calculate a distance between samples of GRNN inputs.

6. CONCLUSIONS

Only one method of nonlinear estimation of GMPE by artificial neural networks was tested – the GRNN. Nonetheless, this approach proved that nonlinear regression modelled by artificial neural networks improves the GMPE.

The abilities of the GRNN were first analysed through learning, by employing a classic selection of the smoothness parameter using the holdout and bootstrap methods. The bootstrap tests showed that the results obtained from the classic selection using the holdout method are too optimistic, and the generalisation is in some cases not rewarding. The bootstrap method showed that some of the GRNN models significantly improved the GMPE, whereas others do not. This was not dependent on the value of the coefficient of determination. The bootstrap method has proved successful for verifying the GMPE model for some input sets, and for some of the GRNN models.

Various input variables for the GRNN were tested. The best results were obtained when one of the inputs was the location of the epicenter. The GMPE was also improved by applying to the GRNN inputs only the logarithms of distance and energy, which are commonly used in LR. A small but significant improvement in the GMPE was achieved when the station location was applied; the result was comparable to the GMPE calculated using the LR method with amplification factors. No significant improvement was achieved when the direction towards the earthquake source was taken as the input for the GRNN.

Out of all the models that were tested, the best results were obtained for a cascade connection of the GRNN with an LR of the logarithms of energy and distance, when these two values were also inputs of the GRNN. This model is superior to the separate GRNN, because the result is not as influenced by large values of the smoothness parameter. It is also superior to the cascade model without $\log E$ and $\log R$ as GRNN inputs, due to the nonlinear relationship between $\log PHA$ and the logarithms of energy and distance, which can be estimated by the two former models only.

The GRNN does not exhaust the possibilities of nonlinear methods of predicting ground motion. It was chosen to test whether using these kinds of methods is reasonable, because it is the fastest of the ANN methods, and the prediction is unique. Further studies are required in order to identify the most suitable method of estimating the GMPE in the LGCD.

Data and resources

The data and ground motions used in this study were collected using a classified network of the KGHM Polska Miedź S.A, and cannot be released to the public.

Our calculations and figures were made using the Matlab program.

Acknowledgements. I wish to express gratitude to KGHM Polska Miedź S.A. – Oddział Zakład Hydrotechniczny for releasing the data for the study.

This work was partially supported within the statutory activity No. 3841/E-41/S/2015 of the Ministry of Science and Higher Education of Poland.

References

Abrahamson, N., and W. Silva (2008), Summary of the Abrahamson & Silva NGA ground-motion relations, *Earthq. Spectra* **24**, 1, 67-97, DOI: 10.1193/1.2924360.

Akkar, S., and J.J. Bommer (2010), Empirical equations for the prediction of PGA, PGV, and spectra accelerations in Europe, the Mediterranean Region, and the Middle East, *Seismol. Res. Lett.* **81**, 2, 195-206, DOI: 10.1785/gssrl.81.2.195.

Arjun, C.R. (2013), *Neural Network-Based Estimation of Strong Ground Motion Parameters*, Lambert Academic Publishing, Saarbrücken.

Boore, D.M., and G.M. Atkinson (2008), Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01 s and 10.0 s, *Earthq. Spectra* **24**, 1, 99-138, DOI: 10.1193/1.2830434.

Campbell, K.W., and Y. Bozorgnia (2008), NGA ground motion model for the geometric mean horizontal component of PGA, PGV, PGD and 5% damped linear elastic response spectra for periods ranging from 0.01 to 10 s, *Earthq. Spectra* **24**, 1, 139-171, DOI: 10.1193/1.2857546.

Cornell, C.A. (1968), Engineering seismic risk analysis, *Bull. Seismol. Soc. Am.* **58**, 5, 1583-1606.

Derras, B., and A. Bekkouche (2011), Use of the Artificial Neural Network for Peak Ground Acceleration estimation, *Lebanese Sci. J.* **12**, 2, 101-115.

Derras, B., P.-Y. Bard, F. Cotton, and A. Bekkouche (2012), Adapting the neural network approach to PGA prediction: An example based on the KiK-net data, *Bull. Seismol. Soc. Am.* **102**, 4, 1446-1461, DOI: 10.1785/0120110088.

Douglas, J. (2011), Ground-motion prediction equations 1964-2010, BRGM/RP-59356-FR, 444 pp.

Efron, B. (1979), Bootstrap methods: another look at the jackknife, *Ann. Stat.* **7**, 1, 1-26, DOI: 10.1214/aos/1176344552.

García, S.R., M.P. Romo, and N. Sarmiento (2003), Modeling ground motion in Mexico City using artificial neural networks, *Geofís. Int.* **42**, 2, 173-183.

Golik, A., and M.J. Mendecki (2012), Ground-motion prediction equations for induced seismicity in the main anticline and main syncline, Upper Silesian Coal Basin, Poland, *Acta Geophys.* **60**, 2, 410-425, DOI: 10.2478/s11600-011-0070-9.

Güllü, H., and E. Erçelebi (2007), A neural network approach for attenuation relationships: An application using strong ground motion data from Turkey, *Eng. Geol.* **93**, 3, 65-81, DOI: 10.1016/j.enggeo.2007.05.004.

Günaydin, K., and A. Günaydin (2008), Peak ground acceleration prediction by artificial neural networks for northwestern Turkey, *Math. Problems Eng.* **2008**, article ID 919420, 1-20, DOI: 10.1155/2008/919420.

Hanna, A.M., D. Ural, and G. Saygili (2007), Neural network model for liquefaction potential in soil deposits using Turkey and Taiwan earthquake data, *Soil Dyn. Earthq. Eng.* **27**, 6, 521-540, DOI: 10.1016/j.soildyn.2006.11.001.

Hong, H., T. Liu, and C.-S. Lee (2012), Observations on the application of artificial neural network to predicting ground motion measures, *Earthq. Sci.* **25**, 2, 161-175, DOI: 10.1007/s11589-012-0843-5.

Joyner, W.B., and D.M. Boore (1988). Measurement, characterization, and prediction of strong ground motion. **In:** *Proc. Earthquake Engineering & Soil Dynamics II. Geotechnical Division*, ASCE Special Publication 20, 43-102.

Joyner, W.B., and D.M. Boore (1993), Methods for regression analysis of strong-motion data, *Bull. Seismol. Soc. Am.* **83**, 2, 469-487.

Lasocki, S. (2013), Site specific prediction equations for peak acceleration of ground motion due to earthquakes induced by underground mining in Legnica-Głogów Copper District in Poland, *Acta Geophys.* **61**, 5, 1130-1155, DOI: 10.2478/s11600-013-0139-8.

Moody, J. (1994), Prediction risk and architecture selection for neural networks. **In**: J. Cherkassy, J.H. Friedman, and H. Wechsler. (eds.), *From Statistics to Neural Networks: Theory and Pattern Recognition Applications*, NATO ASI Series, Springer, Berlin, 147-165, DOI: 10.1007/978-3-642-79119-2.

Parzen, E. (1962), On estimation of a probability density function and mode, *Ann. Math. Statist.* **33**, 3, 1065-1076, DOI: 10.1214/aoms/1177704472.

Pozos-Estrada, A., R. Gómez, and H.P. Hong (2014). Use of neural network to predict the peak ground accelerations and pseudo spectral accelerations for Mexican Inslab and Interplate Earthquakes, *Geofís. Int.* **53**, 1, 39-57, DOI: 10.1016/S0016-7169(14)71489-8.

Specht, D.F. (1991), A general regression neural network, *IEEE Trans. Neural Network* **2**, 6, 568-576, DOI: 10.1109/72.97934.

Trifunac, M.D., and A.G. Brady (1976), Correlations of peak acceleration, velocity and displacement with earthquake magnitude, distance and site conditions, *Earthq. Eng. Struct. Dyn.* **4**, 5, 455-471, DOI: 10.1002/eqe.4290040504.

Wasserman, P.D. (1993). *Advanced Methods in Neural Computing*, John Wiley & Sons, Inc. New York.

Yaghmaei-Sabegh, S. (2012). A new method for ranking and weighting of earthquake ground-motion prediction models, *Soil Dyn. Earthq. Eng.* **39**, 78-87, DOI: 10.1016/j.soildyn.2012.03.006.

Yaghmaei-Sabegh, S., and H.-H. Tsang (2011), A new site classification approach based on neural networks, *Soil Dyn. Earthq. Eng.* **31**, 7, 974-981, DOI: 10.1016/j.soildyn.2011.03.004.

Yaghmaei-Sabegh, S., and H.-H. Tsang (2014), Site class mapping based on earthquake ground motion data recorded by regional seismographic network, *Nat. Hazards* **73**, 3, 2067-2087, DOI: 10.1007/s11069-014-1177-5.

Received 3 December 2015
Received in revised from 18 April 2016
Accepted 26 July 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2449-2470
DOI: 10.1515/acgeo-2016-0109

New Ground Motion Prediction Equation for Peak Ground Velocity and Duration of Ground Motion for Mining Tremors in Upper Silesia

Jacek CHODACKI

Department of Geology and Geophysics, GIG, Central Mining Institute, Katowice, Poland; e-mail: jchodacki@gig.eu

Abstract

This article presents a method of predicting the peak horizontal velocity of ground motion, *PHV,* and the duration of vibration, t_H, for strong seismic events ($E \geq 5 \cdot 10^6$ J, $M_L > 2.5$) in the Upper Silesian Coal Basin (USCB). For the prediction of *PHV*, a model proposed by Si and Midorikawa was used. The regression method takes into account the impact of the local geology under seismic stations on the ground motion according to the Eurocode 8 classification. The ground classification was based on the results of a seismic survey conducted near the seismometer stations. This method is of great practical use because it allows the degree of vibration intensity to be determined on the basis of the Mining Seismic Instrumental Intensity Scale MSIIS-15 (acronym GSI_{GZW} in Polish version) at any distance from the epicentre of the seismic events induced or triggered by mining.

Key words: Ground Motion Prediction Equation, GMPE, peak ground velocity, Upper Silesian Coal Basin, mining seismicity.

1. INTRODUCTION

Hard coal mining, which has been conducted in Silesia for centuries, often in highly urbanized areas, poses serious hazards and significantly affects the safety of the surface. Mining tremors, *i.e.*, dynamic phenomena resulting from rock mass cracking and displacement, are one of sources of such a hazard. Mining tremors cause additional dynamic loads on buildings, which are designed to withstand static loads. These dynamic loads therefore can damage the buildings, weaken their structure and lower their durability and value (Tatara 2012, Zembaty *et al.* 2015).

In Poland, tremors are induced by mining operations in the Upper Silesian Coal Basin (USCB; Mutke and Stec 1997, Lasocki and Idziak 1998, Idziak *et al.* 1999, Marcak and Mutke 2013), Bełchatów open pit mine (Wiejacz and Rudziński 2010) and Legnica-Głogów Copper District (LGCD; Orlecka-Sikora 2010, Lasocki 2013, Lizurek *et al.* 2014), which are associated with the mining of hard coal, brown coal, and copper ore, respectively. The article concerns the USCB and seismicity in the area associated with underground hard coal mining. The seismicity is of two types – the first type, induced seismicity, includes typical mining tremors directly related to mining operations in the area of operating workings and mining (Stec 2012) and the second, tectonic type, includes tremors triggered by a combination of mining and tectonic factors, resulting in disturbance of the rock mass stress equilibrium on a regional scale (Kozłowska *et al.* 2016). Tremors of the first type are much more numerous and fall within an energy range of up to 10^7 J (so-called mining tremors), which corresponds to a magnitude of $M_L = 2.7$. The other group are tremors occurring over a dozen times a year with energies of 10^7-$5\cdot10^9$ J (M_L within the range of 2.7 and 4.2), and these are regional-scale tremors. Tremors with a seismic energy of 10^5 J ($M_L = 1.7$) are felt by inhabitants of the epicentral zone. Every year, in the USCB area, there are between a few hundred and 2 000 tremors with energies between 10^5 and 10^7 J and several to a few dozen tremors with energy of $E \geq 10^8$ J ($M_L \geq 3.3$) (Stec 2007).

The seismic events that cause most of damage to buildings are the triggered ones (Mutke *et al.* 2015, COMEX 2012-2015), despite their lower vibration velocity and acceleration in the epicentral zone relative to the induced ones. The essential differences in the characteristics of the vibrations triggered by regional seismic events are the lower frequency of the vibrations and the longer signal duration. Consequently, higher levels of dynamic loads occur in construction elements of buildings (Zembaty *et al.* 2015). This confirms the importance of the duration of the main phase of vibrations as one of the basic parameters that fundamentally influence the credible assessment of the effects of vibrations on the surface and influence of the mining tremors on buildings (observed intensity).

The protection of the surface in the USCB requires considering the influence of mining seismic events, especially the strongest ones, with seismic energies of 10^7-10^9 J. To assess this influence, seismic intensity scales are used. They describe directly the impact of vibration on surface. The scales of vibration intensity that were applied until recently for such assessments were formulated for earthquakes or other types of paraseismic vibrations, and the assessment criteria they contained were not sufficiently credible for mining tremors. The example includes the most popular Medvedev–Sponheuer–Karnik scale (MSK-64) (Drzeźla *et al.* 2002) and Modified Mercalli Intensity Scale used by United States Geological Survey (Wald *et al.* 1999).

For the last few years, the effects of vibrations in the USCB have been assessed with a special intensity scale, named Mining Seismic Instrumental Intensity Scale, MSIIS-15 (GSI_{GZW} in Polish version) (Mutke *et al.* 2015). It relates the two parameters: peak horizontal velocity of ground motion, *PHV*, and vibration duration time, t_H, to the observed intensity of the surface environment and therefore its use requires knowledge of these two parameters. MSIIS-15 scale can be easily used for measuring these parameters during the seismic measurements, yet it is difficult to apply in areas located at any distance from the epicentre. To do this it is necessary to define GMPE that enables to predict the parameters *PHV* and t_H. This way we may determine the degree of intensity of vibrations in the MSIIS-15 scale for a point on the surface at any distance from the epicentre for an induced or predicted seismic event with a given seismic energy *E*. Moreover, knowledge of the right model of propagation of seismic waves around an epicentre of seismic event enables drawing seismic hazard maps for a given area without installing extensive measuring equipment.

The article presents a new regional GMPE that can predict the *PHV* and t_H in the USCB by considering the ground type according to the Eurocode 8 standard.

2. REVIEW OF PUBLICATIONS CONCERNING GMPE AND DURATION OF VIBRATIONS

Many authors have researched empirical models of seismic wave propagation. Most of the models are based on seismological research and describe the tremors of strong earthquakes.

Seismological observations in Europe, aimed at creating a model of vibrations in a medium to determine the seismic hazard, were conducted by Ambraseys and researchers at Imperial College in London (Ambraseys *et al.* 1996, 2005; Ambraseys and Simpson 1996). These observations enabled them to formulate GMPE for Europe and the Middle East. Other researchers focused on smaller areas of heightened seismic activity, such as Greece (Skarlatoudis *et al.* 2004), Italy (Rinaldis *et al.* 1998, Zonno and Montaldo 2002), Turkey (Kalkan and Gulkan 2004) and Romania (Sokolov *et al.* 2008). Verified GMPE for European and some of the Middle East countries were formulated for both vertical and horizontal peak values of ground vibration acceleration (Ambraseys *et al.* 2005)

In Poland, research was conducted into GMPE of seismic waves in the rock mass. However, this research, unlike the examples of research presented above, was not associated with earthquakes but with mining induced and triggered seismicity (Dubiński and Gerlach 1983, Mutke 1991, Lasocki 2002, Olszewska 2008, Mutke and Stec 2010), mainly in the area of the USCB and LGCD.

Dubiński and Gerlach (1983) determined general relationships that enabled the determination of the degree of intensity of vibrations induced by a rock mass tremor at any distance from the epicentre. The qualitative and quantitative assessment of the influence of rock mass tremors on the natural environment was based on relationships between the intensity of vibrations and their influence on given forms of the environment contained in the MSK scale.

Mutke (1991) formulated regional GMPE for the whole area of the USCB. These relationships can be used to quantitatively determine the peak horizontal velocity and acceleration of ground motion and the dominant frequency of the horizontal components of the vibration acceleration and velocity in Carboniferous and Triassic hard rock. Mutke and Dworak (1992) also demonstrated the importance of local geological structures in the Quaternary overburden, which increases the recorded amplitudes of vibration acceleration and velocity (the phenomenon of vibration amplification).

Signal duration is one of basic parameters describing vibrations in the ground caused by mining seismic events and plays an important role in investigating the effects of the tremors on the surface. Bolt (1973), and Trifunac and Brady (1975), among others, made important contributions to the research into the dependence of signal duration on selected seismic parameters. Although they studied the signal duration of strong earthquakes, their results, especially defining the factors that determine signal duration, can be applied to investigate tremors induced by mining operations.

Esteva and Rosenblueth (1964) describe signal duration as a function of earthquake magnitude and distance from the epicentre. Bolt (1973), based on results of Housner (1965), determined the dependence of signal duration on magnitude, but this relationship is for strong earthquakes and great distances from the epicentre.

Trifunac and Brady (1975) also demonstrated the influence of other factors on signal duration. In their article, apart from magnitude and distance from the epicentre, they considered the influence of local geological structure and investigated a correlation between signal duration and the 12-degree modified Mercalli scale.

3. MATHEMATICAL MODELS

3.1 GMPE for ground motion

GMPE is a relationship between a selected vibration parameter and factors determining its value. It allows for forecasting the seismic effect on the land surface at any distance from the source and assessing the impact of seismic events on surface environment. The vibration intensity scales are used for such assessment. Seismic effects can be represented by various vibration parameters. *PHV* and t_H is such a parameter for MSIIS-15 scales. The factors that influence the magnitude of the parameters are seismic energy, distance from the seismic source, and the local geological and topographical conditions.

GMPE can be determined using measurements of ground vibrations (seismic measurements) in association with known seismic parameters. With respect to the influence of tremors on buildings, strong seismic events play a key role, because they are the only ones that can exert forces that exceed the limit values of the buildings and may be harmful.

3.2 Si-Midorikawa model

Si and Midorikawa (2000) proposed a novel approach to the issue of GMPE. It applies regressions, in which factors associated with the seismic source parameters (volume of energy and source mechanisms) and the path of wave propagation are analysed together with local geological conditions.

For the purpose of this article, this model was modified so that the ground type forming the near-surface zone of the rock medium is an additional factor influencing the amount of vibration recorded on the surface. This model was selected due to the fact that there was a good correlation between the forecast obtained with this model and empirical data. Additionally, the forecast in the epicentral zone gave better results than forecast obtained based on other previously used GMPE.

The proposed in the article model for *PHV* prediction is written as follows:

$$\log(PHV) = b - \log(R) - k \cdot R \ , \tag{1}$$

where *PHV* is the peak horizontal velocity of ground motion, *R* is the source distance,

$$R = \sqrt{d^2 + h^2} \ ,$$

d is the epicentral distance, and *h* is the depth.

Offset parameter *b*, determined for each seismic phenomenon, is the first component. The second one is responsible for non-linear geometric dispersion, and the third one is responsible for non-elastic damping. Coefficient k in the equation assumes the value of 0.002.

Regression analysis is conducted in two stages. In the first step of each recorded phenomenon, the offset parameter b is determined. In the following step, having determined parameter b, we can determine parameters of a regression equation for seismic parameter A.

By transforming Eq. 1, the value of offset parameter b for each recorded phenomenon can be determined. Then, by applying regression analysis, the coefficients a, c, d, and e in the equation were determined:

$$b = a \cdot \log(E) + c \cdot R + \sum d_i S_i + e + \varepsilon \ , \tag{2}$$

where S_i represents the soil classification, assuming a value of 1 for a given site or 0 for others (qualitative variable). This is how we finally obtain the distribution of values for vibration velocity on the surface depending on tremor energy, distance from the seismic source and local geological conditions.

3.3 Influence of the local geological structure on the magnitude of vibrations

Seismological observations show that the magnitude of vibrations on the surface depends on seismic energy, epicentral distance, and geological conditions of the near-surface zone. The amplitudes of vibrations are amplified by a near-surface low-velocity zone. Passage of a seismic wave through the zone also results in changes in the characteristics of the vibrations, including longer durations and a change in their frequency characteristics (Savarienskij 1959). We observe a similar effect in the area of the USCB, which tends to amplify vibrations induced by mining tremors.

The measured amplitude of vibrations can be amplified several times, which has a significant influence on the predicted magnitude of ground vibrations and assessment of the influence of tremors on buildings on the surface (Mutke and Dworak 1992).

There are several methods for determining the level of amplification of vibrations: an analytical method, a method involving direct measurement of the amplification effect, and a method involving studying the spectral relationships of recorded seismic signals known as the HVSR method (Nakamura 1989). In the presented study, the amplification phenomenon was considered directly in a GMPE by applying an additional parameter representing the ground type at the seismometer site in a regression equation.

The classification of the ground was made using the European construction standard "seismic design of buildings", *i.e.*, Eurocode 8. In the classification, the ground is divided into types depending on average wave propagation velocity in 30-m-thick overburden (V_{S30}). This is a convenient tool for engineering applications because methods of surface seismic measurements enable direct measurement of the parameter. Table 1 shows division of the ground into types according to the standard.

Table 1

Classification of the ground according to European standard Eurocode 8

Ground type	Stratigraphic description	V_{S30} [m/s]
A	Rock or other rock-like geological formation, including at most 5 m of weaker material at the surface	> 800
B	Deposits of very dense sand, gravel, or very stiff clay, at least several tens of meters thick, characterized by a gradual increase in mechanical properties with depth	360-800
C	Deep deposits of dense or medium dense sand, gravel or stiff clay with thickness from several tens to many hundreds meters	180-360
D	Deposits of loose-to-medium cohesionless soil (with or without some soft cohesive layers) or predominantly soft-to-firm cohesive soil	< 180

3.4 Study of duration of ground motion

To determine signal duration, the influence of the factors tremor energy, epicentral distance and local geological structure on the parameter was analysed. Regression analysis was used again, applying the following form of a regression model:

$$t_H = a \cdot \log(E) + c \cdot R + \sum d_i S_i + \varepsilon \ , \tag{3}$$

where E is the seismic energy, R is the source distance, S_i is the soil classification parameter, $\Sigma d_i S_i$ is the component responsible for influence of geological conditions on signal duration, and a, c, d are regression coefficients.

According to the guidelines to conduct surface seismometric measurements, the duration of the horizontal component of vibration velocity t_H was determined with an integral of the sum of squares of the horizontal components of vibrations. Signal duration represents the time between the moments when the intensity of the vibrations $I_V(t_k)$ reaches 5 and 95% of the maximum value, as determined with the following equation:

$$I_V(t_k) = \int_0^{t_k} \left(v_x^2(t) + v_y^2(t) \right) dt \ , \tag{4}$$

where v_x^2 is the squared velocity value of x-component, and v_y^2 is the squared velocity value of y-component.

In the equation, the variable t_k determines the dependence of intensity of vibration in time.

4. CHARACTERISTICS OF THE MEASUREMENT DATA

Determining a model of propagation of seismic waves in a rock requires knowledge of the ground vibrations in the site as well as the parameters of the seismic events inducing the vibrations. Moreover, because the propagation model must consider the local geological structure, it is necessary to determine the ground type.

Numerous seismometric, seismological and seismic measurements collected for 25 years at the Department of Geology and Geophysics of the Central Mining Institute were used in the analyses in this study. The seismic data included records of ground vibrations measured by surface seismic stations located all over the USCB area. In total, there were 51 AMAX-GSI-type seismic stations recording amplitudes of ground vibration velocity and acceleration in three perpendicular planes. The distribution of the sites is presented in Fig. 1.

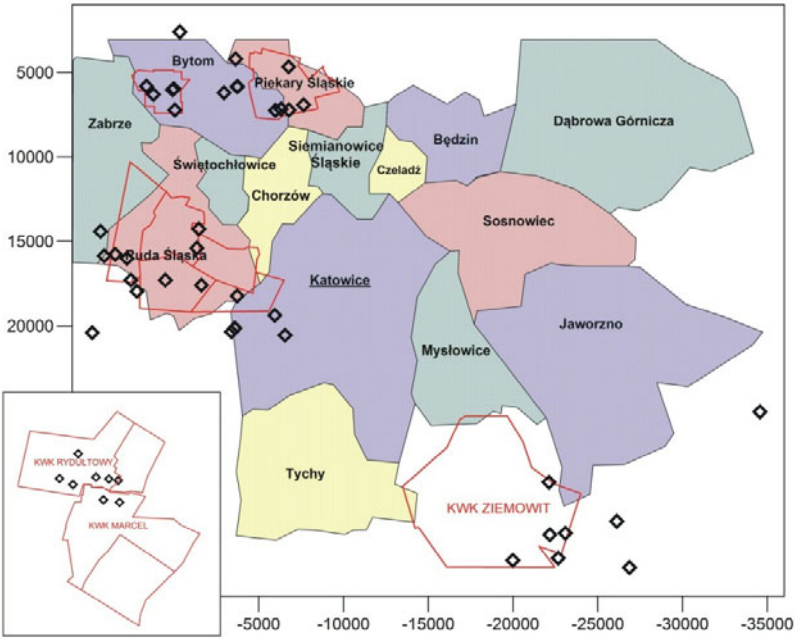


Fig. 1. Location of seismic stations. The main map shows the area of the USCB with boundaries of cities. Additional map shows the area of Rydułtowy and Marcel mines, located about 60 km to the south-west.

This article considers 350 records of tremors with energies between $5\cdot10^6$ J ($M_L = 2.1$) and $3\cdot10^9$ J ($M_L = 4.1$), because only the so-called high-energy seismic events are important in terms of buildings security. In this group, there were 186 records of tremors with energies of 10^6 J, 137 records of tremors with energies of 10^7 J, 16 records of tremors with energies of 10^8 J, and 11 records of tremors with energies of 10^9 J. The recorded values of ground vibration velocity in given sites were between 0.1 and 44 mm/s, and the recorded values of acceleration were between a few and nearly 1300 mm/s^2.

The analysis of the seismological data involved determining the basic parameters, *i.e.*, energy and location of seismic events, as well as time correlation between the seismological and seismometric data. The ground was classified according to the Eurocode 8 standard based on the results of seismic tests conducted using the multichannel analysis of surface waves (MASW) method (Xia *et al.* 1999). The method can determine the distribution of *S* wave velocity in depth scale, *i.e.*, the parameter that determines the ground type in the Eurocode 8 standard, based on the analysis of Rayleigh-type surface waves. An example of the application of this method within the area of USCB can be found in the paper of Siata and Chodacki (2005). It presents the results of seismic measurements taken in Ruda Śląska. The research performed and calculations made allowed for detailing the information concerning the near-surface layers. They constituted the basis for the classification of low-velocity near-surface layers according to Eurocode 8.

5. GMPE OF VIBRATION PARAMETERS

This section presents results of the regression analysis made using models 1, 2 and 3. Their aim was to determine *PHV* and t_H parameters, considering the local geological structure. The analysis was made according to the classical regression method, in which the model is linear (or can be transformed into a linear one).

5.1 GMPE for peak horizontal velocity considering local geological conditions

The construction of a model of wave propagation in rock medium for the USCB area with regression analysis was preceded by calculation of the offset parameter *b* for each recorded phenomenon using Eq. 1. Then, based on Eq. 2, a function describing the decrease in amplitudes of the ground vibration velocity depending on the tremor energy, epicentral distance and ground type at the measuring site was determined. The influence of the given variables, independent of the value of *PHV*, was tested.

The first step was to determine the value of the depth parameter h in Eq. 2. As selection of value of the parameter means selecting the regression coefficients with the lowest value of standard estimation error, a regression analysis for parameter h from the range 100-2000 m was conducted. The lowest value of standard estimation error, S_e, was obtained for $h = 525$ m. Therefore, further analysis of the model was conducted using this depth value.

Finally, the GMPE for the PHV values is as follows:

$$\log(PHV) = 0.209 \cdot \log(E) - 0.035 \cdot R - \log(R) + d_i \ , \tag{5}$$

where PHV is the peak ground velocity [mm/s], E is the seismic energy [J], and R is the source distance [km].

Table 2 presents values of coefficient d_i for given ground types.

Table 2

Values of coefficient d_i for ground types A, B, and C

Coefficient	A	B	C
d_i	–0.814	–0.659	–0.598

The data used in the analysis did not contain records from sites located in ground type D (according to Eurocode 8); thus, the determined model considers only types A, B, and C.

The coefficient of determination $R^2 = 0.86$, which indicates that the model explains variation of PHV in 86%. The standard estimation error $S_e = 0.314$, and the standard estimation errors for the given regression coefficients (which are estimates of the regression coefficients for the whole population) are as follows:

$$S_{\log(E)} = 0.0298, \quad S_R = 0.0086, \quad S_A = 0.2283, \quad S_B = 0.2146, \quad S_C = 0.2097.$$

The significance of the relationship between variables was investigated on the basis of Student's t-test. It shows that there is a strong relationship between variables linear, because all the regression coefficients are highly significant:

$$p_{\log(E)} = 0.000000, \ \ p_R = 0.000135, \ \ p_A = 0.000418, \ \ p_B = 0.002322, \ \ p_C = 0.004622.$$

The significance of the obtained model was tested by using an analysis of variance (Fisher–Snedocor distribution). The null hypothesis for the lack of significance was rejected at the level of 0.000.

A tool which allows for a quick and effective detection of deviations from the correct analysis of regression is the analysis of residuals and that is

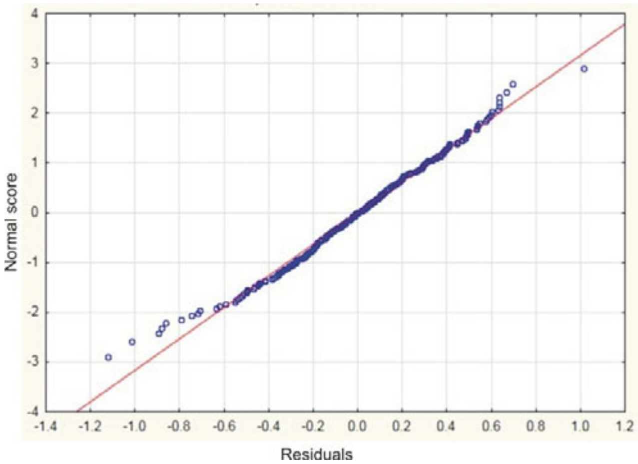


Fig. 2. Normal probability probe for model 5.

why it should constitute one of the most important stages of verification of the regression model. First of all, it allows to verify the assumptions of the classical method of least squares. The first is the assumption stating that the residuals of the model are normally distributed. Normal probability probe for the model 5 has been shown in Fig. 2.

The obtained graph allows to evaluate quickly the conformity of the residuals with the normal distribution. If the residuals are arranged along a straight line, it means that they are normally distributed.

A histogram of residuals provides similar information. At best, the normal curve should go through the centers of upper edges of bars. A slight deviation from the normality, in particular for the samples large in number, does not significantly affect the obtained results. The histogram for the model 5 has been shown in Fig. 3.

Another essential assumption concerns a random component and states that the variance of a random component is the same for all observations. The best method to check this assumption is to create an appropriate scatter plot. If we expect various values of the variance of the random component (σ^2) for different $E(y_i)$, it would be best to create a scatter plot of residuals (which are the estimators of random components) in relation to the expected values (which are the estimators $E(y_i)$). The following scatter plot was obtained for the analyzed model. The scatter plot obtained for the analyzed model has been shown in Fig. 4.

The visible points cloud without a clear trend of increase (decrease) of the residual variation with an increase of the expected values of residuals shows that the assumption of a constant variance of the random component is met.

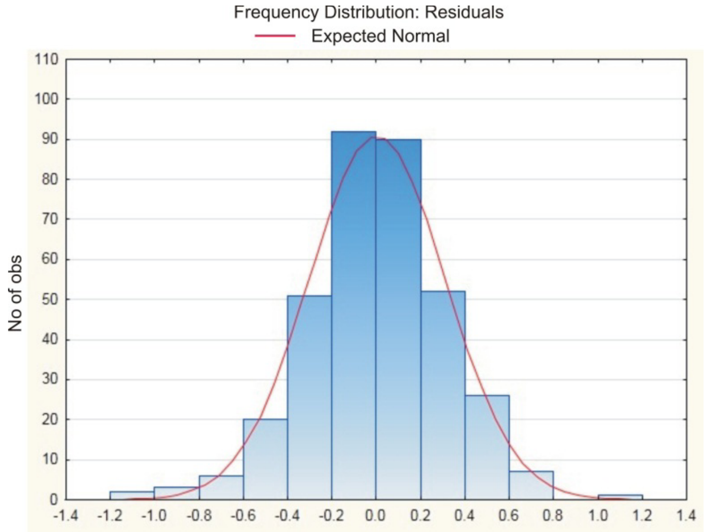


Fig. 3. Histogram of residuals for model 5.

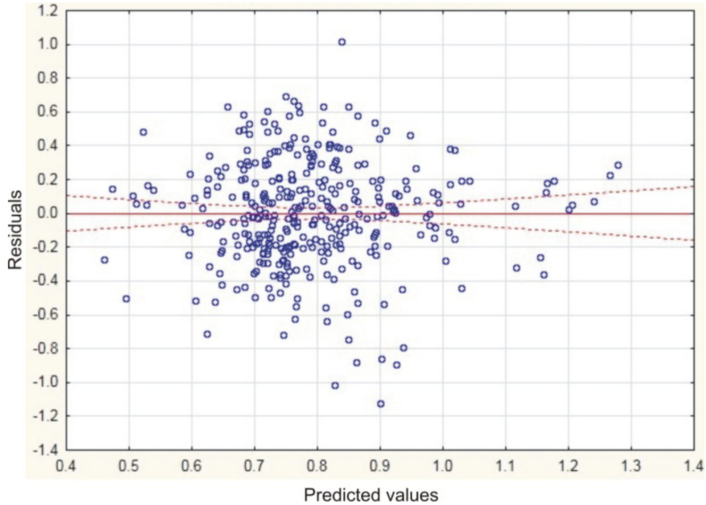


Fig. 4. Plot of residuals against predicted values for model 5.

Once a model of propagation of a seismic wave is ready and verified, it can be used to predict seismic effects on the surface. The predicted values of *PHV* for a tremor with an energy of $1{\cdot}10^8$ J ($M_L = 3.3$) and the 90% upper confidence interval are presented in Fig. 5. In the graphs, the *x*-axis represents the epicentral distance.

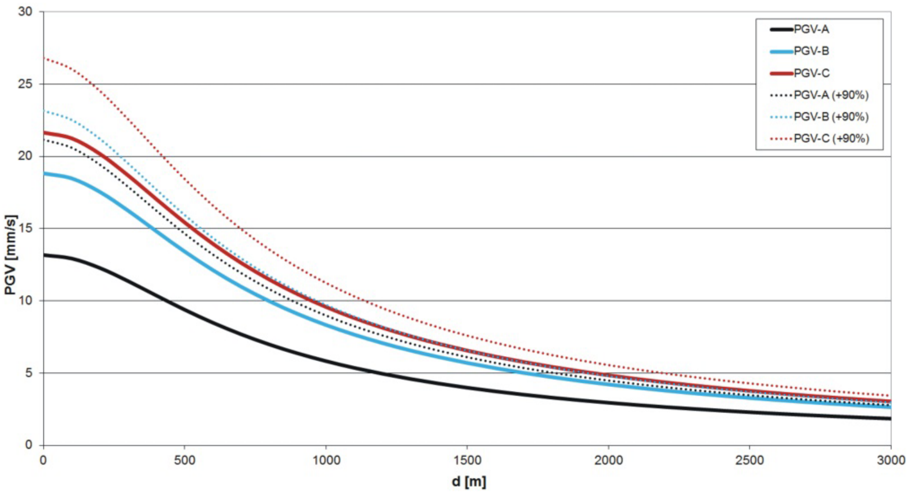


Fig. 5. redicted values of *PHV* with the 90% upper confidence interval for a seismic event with an energy of $1{\cdot}10^8$ J.

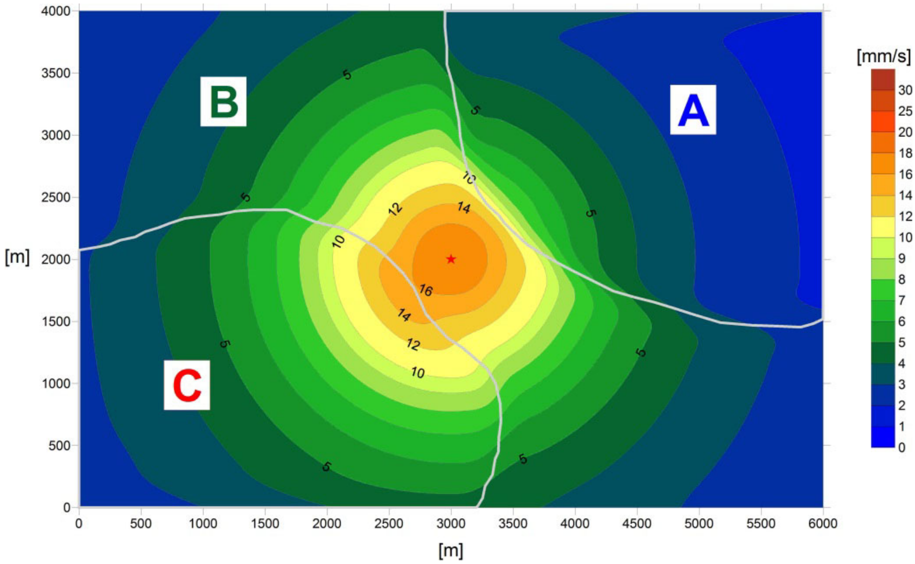


Fig. 6. Distribution map of parameter *PHV* for a seismic event with an energy of $1{\cdot}10^8$ J.

Comparing the *PHV* curves corresponding to different ground types, the influence of a near-surface low-velocity layer on vibrations recorded at the surface is clearly observed. The lower the average wave propagation velocity in the layer, the larger the recorded vibrations. This pattern is confirmed by observations and is a result of the already mentioned phenomenon of increased amplitudes of vibrations as the wave propagates through a low-velocity zone.

The values of the predicted ground vibration velocity of a tremor with an energy of $1\cdot10^8$ J in the epicentral zone are within the range of between approximately 13 mm/s for ground type A and approximately 22 mm/s for ground type C. For ground type B, the velocity is approximately 19 mm/s. In the epicentral zone, there is a rapid decrease in the amplitudes of the vibrations (approximately 50% at a distance of 900 m from the epicentre), while for epicentral distances greater than approximately 1500 m, the curves are flatter.

Based on Eq. 5, it is also possible to draw a map of *PHV* around the tremor epicentre. Figure 6 presents such a map. The epicentre of a tremor with an energy of $1\cdot10^8$ J is located in the centre of the area. The grey lines mark the boundary areas of ground types A, B, and C.

5.2 Determining the dependence between signal duration and epicentral distance

Signal duration was determined on the basis of Eq. 3. For this relationship, the factor associated with tremor energy (for tremors with energies between $5\cdot10^6$ J and $3\cdot10^9$ J) turned out to be statistically non-significant (on the basis of Student's *t*-test). During regression analysis, the most extreme values were discarded (residua analysis enabled us to identify and discard them). Overall, the form of the relationship between signal duration, epicentral distance and ground type is as follows:

$$t_H = 3.417 \cdot \log(R) + c_i \ , \tag{6}$$

where R is the source distance [km].

The values of the coefficient c_i for the given ground types are presented in Table 3.

Table 3

Values of coefficient c_i for ground types A, B, and C

Coefficient	A	B	C
c_i	1.9218	2.3503	3.136

The coefficient of determination for Eq. 6 is $R^2 = 0.92$. The standard estimation error is $S_e = 1.12$, and the standard estimation errors for given regression coefficients are as follows:

$$S_{\log(R)} = 0.216,\quad S_A = 0.412,\quad S_B = 0.122,\quad S_C = 0.114.$$

The significance of the obtained model was tested by using analysis of variance (Fisher–Snedocor distribution). The null hypothesis for the lack of significance was rejected at the level of 0.000.

Normal probability probe, histogram of residuals and plot of residuals against predicted values for model 6 are shown in Figs. 7-9.

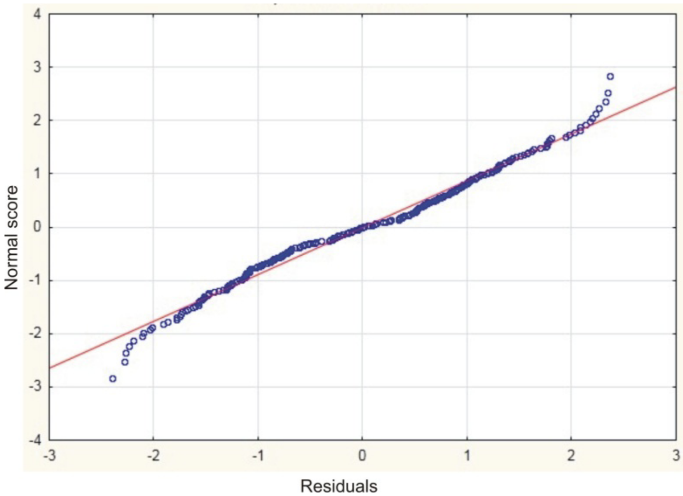


Fig. 7. Normal probability probe for model 6.

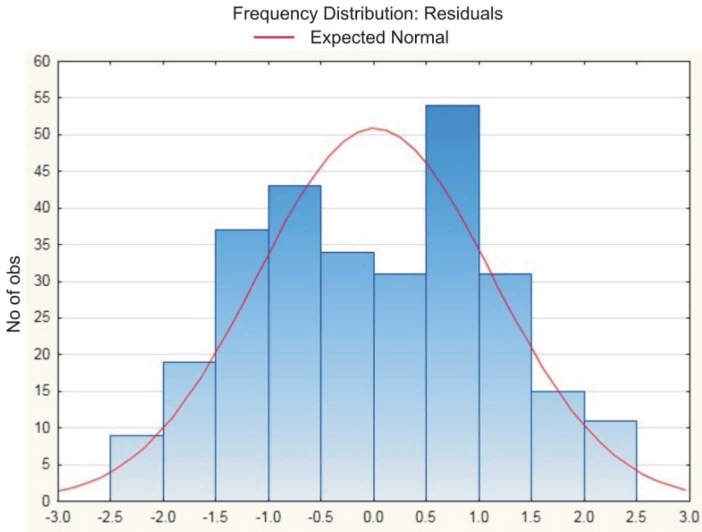


Fig. 8. Histogram of residuals for model 6.

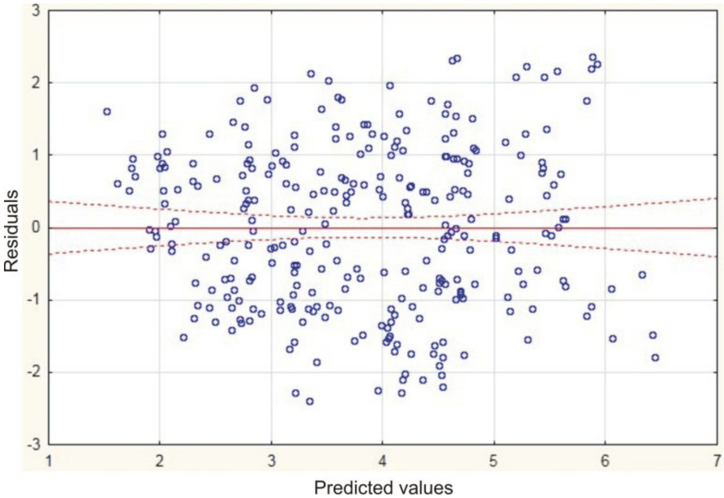


Fig. 9. Plot of residuals against predicted values for model 6.

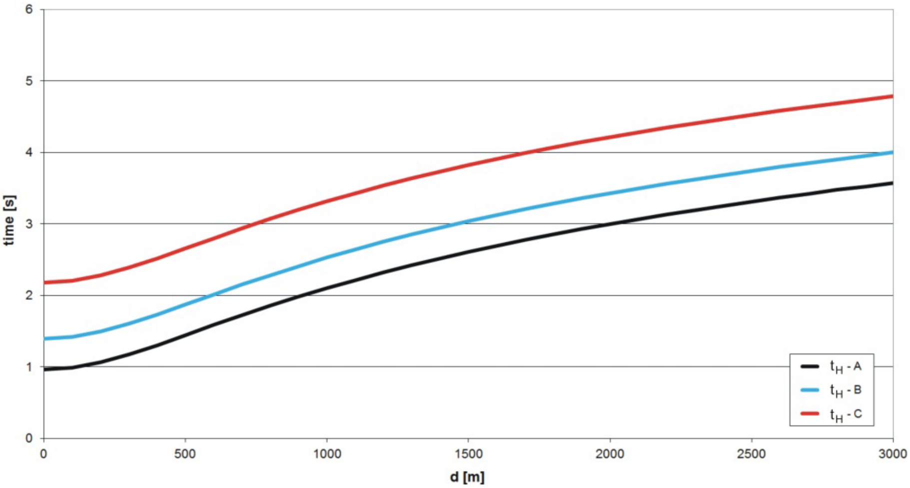


Fig. 10. The relationships between signal duration, the epicentral distance and ground type.

A graph of dependence 6 is presented in Fig. 10. The curves – a solid line, a dashed line, and a dotted line – present the dependence of signal duration on the epicentral distance for ground types A, B, and C, respectively. With an increase in epicentral distance, the time increases. There is also a visible influence of geological conditions on the signal duration: the weaker the formations, the lower the velocity of wave propagation *S* in the 30-metre-thick overburden and the longer the signal duration for a given epicentral distance.

6. DISCUSSION AND SUMMARY

There are various models of GMPE and the choice of the best is not obvious. The commonly used models come from the regression model of Joyner-Boore (Lasocki *et al.* 2000), but there are also non-linear models (Mutke 1991). The selection of factors influencing the vibrations which we analyse also plays an important role. In this article, in addition to commonly applied parameters such as tremor energy and distance from the seismic source, the influence of the local geological structure was also investigated. For the first time, the model proposed by Si and Midorikawa was used to investigate the influence of mining tremors on the surface in the USCB. The results show that the model well represents the physical characteristics of seismic wave propagation in a rock medium. An increase in tremor energy causes an increase in the ground vibration velocity at the surface, whereas an increase in the epicentral distance results in a decrease in vibration velocity. The relationships between values of tremors for different ground types also agree with actual measurements and analytical solutions (Savarienskij 1959). According to these relationships, a seismic wave passing through weaker ground (of lower seismic wave propagation velocity) results in amplified vibrations at the surface. The same phenomenon occurs in the predicted signal duration. Passing through a low-velocity medium results in a change in frequency characteristics and a longer signal duration.

So far, two equations for ground motion prediction have been used in the USCB. The first one is a non-linear model developed by Mutke (1991) which describes the dependence between the amplitude of a horizontal component of acceleration and velocity of bedrock vibrations, and the epicentral distance and seismic energy. This model is used for forecasting vibration resulting from high-energy seismic events and it allows to specify the values of acceleration and velocity of vibrations at the surface of hard rock layers. If weaker soil layers which form a zone of low velocities of seismic wave propagation are above a hard rock, the vibration amplitudes can be strengthened. This phenomenon is called the amplification of vibrations and it is considered analytically in the subject model. The results of the forecast obtained based on model of Mutke coincides generally with the results obtained on the basis of GMPE for the soil class A presented in this paper (Fig. 11).

For the epicentral distances above 1000 m, these results are almost the same. On the other hand, the model designated by Mutke has more "flat" course in the epicentral area. It may result from the adopted depth of the source which was 600 m for Mutke model as well as the energy of seismic events which were used to develop the model (model of Mutke refers to seismic events with energy greater than $1 \cdot 10^5$ J).

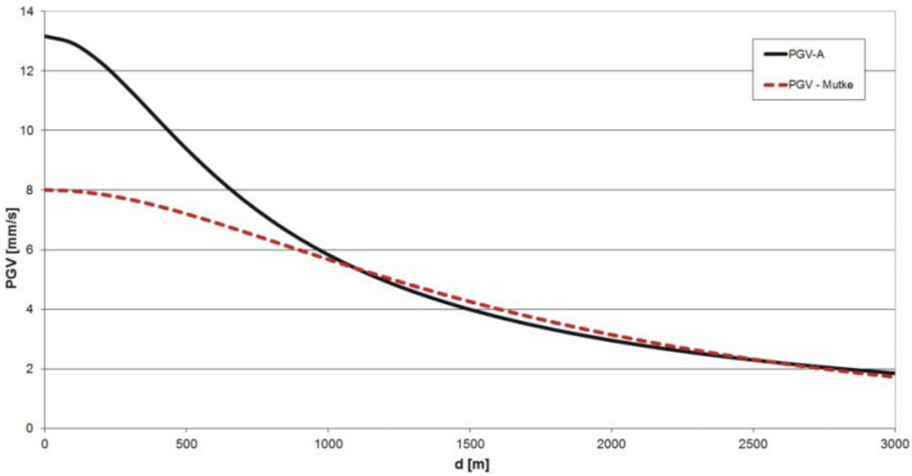


Fig. 11. Comparison of the forecast results for model of Mutke and for model presented in article.

The second model is the linear Joyner-Boore model (Joyner and Boore 1981), mostly used locally by coalmines having a network of seismic stations. In this case, the coefficients in GMPE are designated for particular mining areas and they take into account the local seismic nature for a certain coalmine. These models are developed based on mining seismic events and the possibility of their application in forecasting vibrations resulting from regional events is limited. The peak value of ground motion is determined in this model for the average geological conditions in the place where the seismometer stations are located.

As far as the forecast of the parameter t_H is concerned, it is the first paper concerning USCB presenting how the distance from the seismic source and the structure of the rock influence the duration of vibrations, and thus the parameter which plays a key role in the study of the seismic effects on the surface.

The GMPE proposed in the article is of great practical importance and will be applied to solve problems of seismic engineering in mining areas; the obtained results will be used to assess hazards associated with the seismic influence of actual tremors, thereby helping to minimize the effects of future mining operations in the USCB.

References

Ambraseys, N.N., and K.A. Simpson (1996), Prediction of vertical response spectra in Europe, *Earthq. Eng. Struct. Dyn.* **25**, 4, 401-412, DOI: 10.1002/(SICI)1096-9845(199604)25:4<401::AID-EQE551>3.0.CO;2-B.

Ambraseys, N.N., K.A. Simpson, and J.J. Bommer (1996), Prediction of horizontal response spectra in Europe, *Earthq. Eng. Struct. Dyn.* **25**, 4, 371-400, DOI: 10.1002/(SICI)1096-9845(199604)25:4<371::AID-EQE550>3.0.CO;2-A.

Ambraseys, N.N., J. Douglas, S.K. Sarma, and P.M. Smit (2005), Equations for the estimation of strong ground motions from shallow crustal earthquakes using data from Europe and the Middle East: Horizontal peak ground acceleration and spectra acceleration, *Bull. Earthq. Eng.* **3**, 1, 1-53, DOI: 10.1007/s10518-005-0183-0.

Bolt, B.A. (1973), Duration of strong ground motion. **In:** *5th World Conference on Earthquake Engineering, Rome*, 6-D, Paper No. 292.

COMEX (2012-2015), Complex Mining Exploitation: optimizing mine design and reducing the impact on human environment (2012-2015), COMEX Project (RFCS-PR-11012), Deliverable 1.4.: New European Mining Seismic Intensity Scale – MSIIS-15.

Drzeźla, B., J. Dubiński, and G. Mutke (2002), Macroseismic scales – their essence and principles of using in assessments of the effects of mining tremors**. In:** A. Sroka and R. Wittenburg (eds.), *Geokinematischer Tag des Institutes fur Markscheidewesen und Geodasie and der TU Bergakademie, Freiberg*, 47-55.

Dubiński, J., and Z. Gerlach (1983), Assessment of impact of mining seismicity on the natural environment, *Prz. Górn.* **3**, 135-142 (in Polish).

Esteva, L., and E. Rosenblueth (1964), Espectros de temblores a distancias moderadas y grandes, *Bol. Soc. Mex. Ing. Sism.* **2**, 1, 1-18.

Housner, G.W. (1965), Intensity of ground shaking near the causative fault. **In:** *Proc. Third World Conference on Earthquake Engineering*, Vol. 1, New Zealand.

Idziak, A.F., L. Teper, and W.M. Zuberek (1999), Seismicity and tectonics of the Upper Silesian Coal Basin, *University of Silesia Press Katowice* **1999**, 1793 (in Polish).

Joyner, W.B., and D.M. Boore (1981), Peak horizontal acceleration and velocity from strong-motion records including records from the 1979 Imperial Valley, California, earthquake, *Bull. Seismol. Soc. Am.* **71**, 6, 2011-2038.

Kalkan, E., and P. Gulkan (2004), Empirical attenuation equations for vertical ground motion in Turkey, *Earthq. Spectra* **20**, 3, 853-882, DOI: 10.1193/1.1774183.

Kozłowska, M., B. Orlecka-Sikora, L. Rudziński, and G. Cielesta Mutke (2016), A typical evolution of seismicity patterns resulting from the coupled natural, human-induced and coseismic stresses in a longwall coal mining environment, *Int. J. Rock Mech. Mining Sci.* **86**, 5-15, DOI: 10.1016/j.ijrmms. 2016.03.024.

Lasocki, S. (2002), Attenuation relation for horizontal component of peak ground acceleration below 10 Hz frequency for the Polk, *Publs. Inst. Geophys. Pol. Acad Sci*. **M-27**, 352, 79-90 (in Polish).

Lasocki, S. (2013), Site specific prediction equations for peak acceleration of ground motion due to earthquakes induced by underground mining in Legnica-Głogów Copper District in Poland, *Acta Geophys.* **61**, 5, 1130-1155, DOI: 10.2478/s11600-013-0139-8.

Lasocki, S., and A. Idziak (1998), Dominant directions of epicenter distribution of regional mining-induced seismicity series in Upper Silesian Coal Basin in Poland, *Pure Appl. Geophys.* **153**, 1, 21-40, DOI: 10.1007/s000240050183.

Lasocki, S., M. Szybiński, J. Matuszyk, J. Mirek, and A. Pielesz (2000), Prediction of ground motion caused by seismic events from mines: a critical review. **In:** *Workshops 2000 "Natural Hazards in Mining"*, Conference publication, IGSMiE PAN, 261-279 (in Polish).

Lizurek, G., Ł. Rudziński, and B. Plesiewicz (2014), Mining induced seismic event on an inactive fault, *Acta Geophys.* **63**, 1, 176-200, DOI: 10.2478/s11600-014-0249-y.

Marcak, H., and G. Mutke (2013), Seismic activation of tectonic stresses by mining, *J. Seismol.* **17**, 4, 1139-1148, DOI: 10.1007/s10950-013-9382-3.

Mutke, G. (1991), Prediction of parameters of ground motion caused by mining tremors in the Upper Silesian Coal Basin, Ph.D. Thesis, CMI (in Polish).

Mutke, G., and J. Dworak (1992), Factors determining the effect of mining seismic events on buildings in the Upper Silesian Coal Basin. Selected issues of the geophysical studies in the mines – Lubiatów 1991, *Publs. Inst. Geophys. Pol. Acad. Sc.*, **M-16**, 245, 115-130 (in Polish).

Mutke, G., and K. Stec (1997), Seismicity in the Upper Silesian coal basin, Poland: Strong regional seismic events. **In:** S.J. Gibowicz and S. Lasocki (eds.), *Proc. 4th Int. Symp. Rockburst and Seismicity in Mines*, Balkema, Rotterdam, 213-218.

Mutke, G., and K. Stec (2010), Prediction of parameters of ground motion for intensity assessment of mining tremor impact using the GSI-GZW scale, *Res. Rep. CMI Min. Environ.* **4/1/2010** (in Polish).

Mutke, G., S. Kremers, J. Chodacki, R. Fritschen, and L. Muszyński (2015), Mining seismic instrumental intensity scale MSIIS-15 – verification in coal basins. **In:** *5th Inter. Symp. Mineral Resources and Mine Development, Aachen*.

Nakamura, Y. (1989), A method for dynamic characteristics estimations of subsurface using microtremors on the ground surface, *QR RTRI* **30**, 25-33.

Olszewska, D. (2008), Analysis of site effects and frequency spectrum of signals in order to improvement accuracy of attenuation relation of ground motion caused by mining induced seismic events in Legnica Głogow Copper District, Ph.D. Thesis, AGH, Kraków, Poland (in Polish).

Orlecka-Sikora, B. (2010), The role of static stress transfer in mining induced seismic events occurrence, a case study of the Rudna mine in the Legnica-Glogow Copper District in Poland, *Geophys. J. Int.* **182**, 2, 1087-1095, DOI: 10.1111/j.1365-246X.2010.04672.x.

Rinaldis, D., R. Berardi, N. Theodulidis, and B. Margaris (1998), Empirical predictive models based on a joint Italian and Greek strong-motion database: I. Peak ground acceleration and velocity. **In:** *Proc. 11th European Conference on Earthquake Engineering*, Balkema, Rotterdam.

Savarienskij, E.F. (1959), Evaluation of the impact of nearsurface layer on the amplitude of ground motion, *Izv. Akad. Nauk SSSR Geofizyka* **10**, 1441-1447 (in Russian).

Si, H., and S. Midorikawa (2000), New attenuation relations for peak ground acceleration and velocity considering effects of faulty type and site condition. **In:** *12th World Conference on Earthquake Engineering*.

Siata, R., and J. Chodacki (2005), Application of MASW method for determining the velocity profile of the subsurface layers. **In:** *Mining Workshops 2005*, Conference publication, IGSMiE PAN, Kraków, Poland (in Polish).

Skarlatoudis, A., N. Theodulidis, C. Papaioannou, and Z. Roumelioti (2004), The dependence of peak horizontal acceleration on magnitude and distance for small magnitude earthquakes in Greece. **In:** *Proc. Thirteenth World Conference on Earthquake Engineering*, paper no. 1857.

Sokolov, V., K.P. Bonjer, F. Wenzel, B. Grecu, and M. Radulian (2008), Ground-motion prediction equations for the intermediate depth Vrancea (Romania) earthquakes, *Bull. Earthq. Eng.* **6**, 3, 367-388, DOI: 10.1007/s10518-008-9065-6.

Stec, K. (2007), Characteristics of seismic activity of the Upper Silesian Coal Basin in Poland, *Geophys. J. Int.* **168**, 2, 757-768, DOI: 10.1111/j.1365-246X.2006.03227.x.

Stec, K. (2012), Focal mechanisms of Mine induced seismic events an explanation of geomechanical processes in the area of longwall 6, seam 510 in hard coal mine Bobrek-Centrum, *Arch. Mining Sci.* **57**, 4, 871-886, DOI: 10.2478/v10267-012-0057-7,.

Tatara, T. (2012), Dynamic resistance of buildings in terms of mining tremors, Cracow University of Technology, Cracow, Poland (in Polish).

Trifunac, M.D., and A.G. Brady (1975), A study of the duration of strong earthquake ground motion, *Bull. Seismol. Soc. Am.* **65**, 3, 581-626.

Wald, D.J., V. Quitoriano, T.H. Heaton, and H. Kanamori (1999), Relationship between peak ground acceleration, peak ground velocity, and modified mercalli intensity in California, *Earthq. Spectra* **15**, 3, 557-564, DOI: 10.1193/1.1586058.

Wiejacz, P., and Ł. Rudziński (2010), Seismic event of January 22, 2010 near Bełchatów, Poland, *Acta Geophys.* **58**, 6, 988-994, DOI: 10.2478/s11600-010-0030-9.

Xia, J., R.D. Miller, and C.B. Park (1999), Estimation of near-surface shear-wave velocity by inversion of Rayleigh waves, *Geophysics* **64**, 3, 691-700, DOI: 10.1190/1.1444578.

Zembaty, Z., S. Kokot, F. Bozzoni, L. Scandella, C.G. Lai, J. Kuś, and P. Bobra (2015), A system to mitigate deep mine tremor effects in the design of civil infrastructure, *Int. J. Rock Mech. Mining. Sci.* **74**, 81-90, DOI: 10.1016/j.ijrmms.2015.01.004.

Zonno, G., and V. Montaldo (2002), Analysis of strong ground motions to evaluate regional attenuation relationships, *Ann. Geophys.* **45**, 3-4, DOI: 10.4401/ag-3518.

Received 16 December 2015
Received in revised form 18 July 2016
Accepted 24 August 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2471-2486
DOI: 10.1515/acgeo-2016-0093

Distribution Characteristics of TOC, TN and TP in the Wetland Sediments of Longbao Lake in the San-Jiang Head Waters

Sujin LU[1], Jianhua SI[2], Yue QI[1], Zhanqing WANG[1], Xiaocui WU[1], and Chuanying HOU[1]

[1]Eco-environmental Engineering College, Qinghai University, Xi'ning, Qinghai Province, China; e-mail: lusujin88@163.com (corresponding author)

[2]College of Agriculture and Animal Husbandry, Qinghai University, Xi'ning, Qinghai Province, China

Abstract

The study deals with the distribution of nutrients in wetland sediments, which provide the basis for revealing the wetland eutrophication processes and mechanisms of internal pollution sources. The total organic carbon (TOC), total nitrogen (TN), and total phosphorus (TP) contents and distribution characteristics of sediment samples were examined. The results showed that the TOC concentration ranged from 3.81 to 15.6 g/kg, the TN concentration ranged from 0.21 to 1.18 g/kg with a mean concentration of 0.66 g/kg, and the TP concentration ranged from 0.16 to 0.35 g/kg with a mean of 0.23 g/kg. Statistical analysis showed close correlations between TOC and TN ($R^2 = 0.96$), and TN and TP ($R^2 = 0.97$), which indicated that the TN and TP in the sediments were from similar sources. The concentrations of TOC, TN, and TP in Longbao Lake wetland sediments were too low for eutrophication to occur. Our investigation indicated that Longbao Lake undergoes natural evolution rather than anthropogenic activities.

Key words: Longbao Lake, sediments, organic carbon, nitrogen, phosphorus, high altitude wetlands.

1. INTRODUCTION

Wetlands are referred to as “the kidneys of the Earth”, with the richest biodiversity, important for mankind. Wetlands play important roles in climate regulation, water conservation, protection of biological diversity and provision of resources for human production and life (Barbier *et al.* 1997; Mitch and Gosselink 2000). Alpine wetlands are important wetland types, which are generally distributed in cold zones with high altitudes (Sun and Liu 2006). The important distribution area for alpine wetlands in China are Qinghai Tibetan Plateau, covering 13 300 000 km^2, with an average elevation of more than 3000 m. The San-Jiang-Yuan Nature Reserve in the Qinghai Tibetan Plateau is located in the headwaters of the Yangtze, Yellow and Lancang Rivers, considered as a source for rivers and Chinese water towers, where rivers are densely developed and numerous lakes, swamps, snow-covered mountains and glaciers are widely distributed (Ma *et al*. 2009). Wetland ecosystems with the San-Jiang-Yuan Nature Reserve, including alpine swamps, alpine swampy meadows and alpine lakes, play crucial roles in water storage, recharging for surrounding areas and moderation of local and global climate (Lin 2008). Thus, the San-Jiang-Yuan Nature Reserve is of significant interest for ecological safety in China and around the world.

However, the simple structure of cold wetland ecosystems manifests itself in a relatively uniform composition, a weak system regulation ability as well as low ecological capacity and resilience. Therefore, cold wetland ecosystems are particularly sensitive to climate and environmental changes that cause wetland degradation in San Jiang Yuan Nature Reserve (Yang *et al*. 2008). The areas of forest and grasslands in the area of Longbao Lake in San-Jiang-Yuan Nature Reserve have been decreasing, and the permanent area of snow glaciers is currently only 24276.6 hm^2, which represents a reduction of 85%. The lake and swamp area, as a representative area for wetlands, was shrinking, and the area of swamp decreased by 0.93% a^{-1} (487.3 hm^2 a^{-1}) (Zheng 2011). Considering recent global warming and environmental damage and pollution, the ecological systems and water resources in the wetlands in San Jiang Yuan are seriously threatened. Furthermore, the water quality is a great concern for the wetland. The excess of nutrients into lakes can result in the destruction of the lake’s ecosystem and water quality deterioration (Li *et al*. 2001). P and N pollutants are released from sediments and can enter the water over a long period, which results in a significant amount of P in the water column (Ding *et al*. 2015). High internal P loading is frequently reported as a major factor that is responsible for delaying lake recovery after reducing external P loading (Ding *et al*. 2015, Søndergaard *et al*. 2007). Nutrient salt in sediments is co-related to exogenous pollution in wetland ecosystems. Although many researchers have recently conducted extensive studies of nitrogen and phosphorus sediments in lake ecosystems, most of these studies have focused on the release of nitrogen and phosphorus from sediments. The release of nitrogen and phosphorus from sediments contributes to exogenous pollution and water quality deterioration, which impact the organic matter in the sediments and the synergistic reactions with other biogenic elements in the lake. However, studies regarding the biogenic elements and environmental elements of lakes are lacking.

Currently, studies of the San Jiang Yuan headwaters mainly involve soil water heat processes, the deterioration of alpine wetlands, the effects of grazing on grassland reserves, and the responses of wetlands to regional climate changes (Lin 2003, Ma *et al*. 2009, Chen *et al*. 2012). However, studies on the distribution characteristics and potential variation tendencies of nutrient salts in sediments, with alpine lakes in the San-Jiang head source, have not been reported (Lu *et al*. 2009). Research regarding the TOC, TP, and TN contents with sediment quality in alpine lakes is necessary to understand the eutrophication potential in such lakes, and protect alpine wetland, which are important for maintaining a healthy, reasonable, sustainable and ecological environment in the plateau. The objectives of this study were to measure the TOC, TN, and TP contents and their distribution characteristics in sediments in Longbao Lake to reveal changes in nutrient (N and P) contents, determine the mechanisms of endogenous pollution in the Longbao Lake wetland, and provide a scientific basis for preventing eutrophication in Longbao Lake wetlands and references for damaged wetland restoration.

2. MATERIALS AND METHODS

2.1 Study site description

Longbao Lake covers an area of approximately 50 km^2 (18.7 km × 3 km) and is located in the Longbao Nature Reserve in the Yushu Tibetan Autonomous Prefecture of Yushu county in Qinghai Province (N 33°09′~33°16′, E°96°24′~96°37′) (Fig. 1). The Longbao Lake Nature Reserve is located on an average altitude of 4500 m covering a total area of 147.2 km^2. The reserve is located from the center to the surrounding regions, followed by lakes, swamps, marshes, meadows, alpine meadows, alpine mountain meadows, and bare rock lands. Thus, this region is considered to have a typical plateau continental climate (Lin 2003). Rainfall is concentrated from June to August, and approximately 2300 h of sunshine occur throughout the year. The annual precipitation in this region is 730 mm. The lake freezes in November and the ice is melted over the next May.

There are 33 types of birds in the Longbao nature reserve, including black-necked crane (*Grus nigricollis*) (Shi *et al*. 2009). The ecosystem involves lakes, marshes, swamp meadow, with Charophytes (*Chara fragilis*

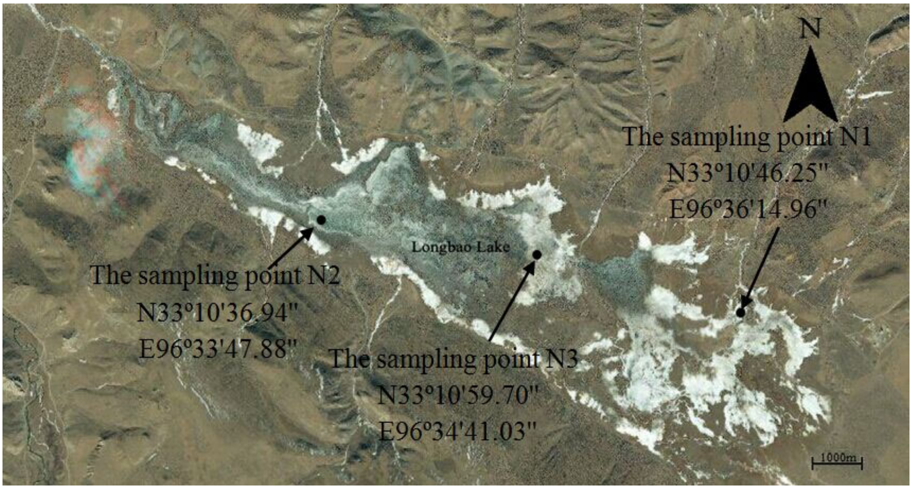


Fig. 1. Study area location and sampling station distribution.

Desv), Cedrus folium algae (*Hippuris vulgaris* L), *etc*., which are the main aquatic plants. The dominant plants in the Longbao Lake Nature Reserve are Aqua Buttercup (*Batrachium bungei (Steud.)* L. Liou), Siberia Polygonum (*Polygonum sibiricum* Laxm), Humilis (*Kobresia myosuroides* Fiori), Dwarf Nasturtium (*Trollius farreri* Stapf), and Water Winter Wheat (*Triglochin palustre*). The total population of the Long Lake Nature Reserve is 370 000, and animal husbandry is the main economic pillar in the region (Qing 2013). Currently, the entire periphery of the wetland is surrounded by chain link fence and human grazing is not permitted, with the purpose of reducing pollution and destruction of the wetland by human and grazing.

The upstream of Longbao Lake is the largest inflow river Yiqu, with flow of 21 m^3/s, annual runoff of 6.61×10^9 m^3. Yiqu river, Dengeyongqu river and Kequ river are large flowing rivers, with flow of 20.5, 20.8, and 20.9 m^3/s, respectively, while the annual runoff was 5.34×10^9, 5.27×10^9, and 5.47×10^9 m^3. The water temperature was 13.4 °C, pH was 7.9, conductivity was 0.66 ms/cm, DO was 7.43 mg/L, BOD_5 was 2.92 mg/L, and COD was 14.68 mg/L in Yiqu river. In Dengeyongqu river, the water temperature was 13.6 °C, pH was 8.0, conductivity was 0.69 ms/cm, DO was 7.64 mg/L, BOD_5 was 2.96 mg/L, and COD was 14.99 mg/L. In Kequ river, the water temperature was 13.2 °C, pH was 7.8, conductivity was 0.63 ms/cm, DO was 7.22 mg/L, BOD_5 was 2.94 mg/L, and COD was 15.11 mg/L.

The Longbao town, the road and wetland animals are the major sources of pollution surrounding the Longbao Lake. Longbao town is situated west of Longbao Lake, 17 km from the lake, with a population of 8623. Untreated sewage from residents of Longbao town will input into the lake by runoff, causing water pollution. The road is about 500 m away, to the north of the lake. So, grazing animals, vehicles and tourists passing by could result in automobile exhaust, dust, manure, garbage, which may pollute the lakes.

2.2 Basin shape

The lake is supplied by waters from surrounding mountains. Recent years, the permanent glaciers around the Longbao Lake area decreased with global climate changing, the area of wetland and swamp presents a declining trend, the water exchange cycle of Longbao Lake extended, the water input to the Longbao Lake decreased, and the self cleaning capacity weakened.

2.3 Experimental design and sampling

In order to avoid interference of water level, which is too high in the Lake for sampling, samples were collected during the dry season. To compare the related data from each sampling point and to consider the integrity of the watershed layout, three representative locations were selected (N1, N2, and N3) and stratified sediment samples were collected at sites N1, N2, and N3 (Fig. 1). Sampling was conducted in October 2012, May 2013, October 2013, May 2014, and October 2014. Each sampling site was located using a global satellite navigation and positioning system for precise positioning. For sampling, an open type stainless steel cylindrical sediment sampler was used.

Longbao Lake sediment profile sample hierarchy is as follows: three points were taken from the 50 cm mud column; use the mud collector to collect the overlying water; try to reduce the disturbance of surface sediments. Each 5 cm mud column was considered as one sample of the profile. All the samples were taken back to the lab and dried by indoor natural air, and different levels of sediment were sieved according to the corresponding determination and analysis methods (Guo 2007).

Before analyzing total N, total P, and TOC contents, the sediments were air-dried and passed through a 1 and 0.25 mm sieve and were stored in cloth bags to keep out the light. The total P content was determined by colorimetry using the phospho-molybdate reduction method, which is briefly described in the publication quoted as: ISSCAS (1987). Firstly, weigh 0.5 g sediment sample, and put it in a triangular flask. Add 6 mL concentrated sulfuric acid and 10 drops of perchlorate. Put it on the electric furnace to be digested until a colorless. The sediment sample was decomposed. Then a standard curve was drawn. Then, the sediment samples were placed into a colorimetric tube. Add digested sediment samples and 2~3 drops of 2-, 4-nitro phenol indicator. Adjust the pH. Next, the 1 ml 10% ascorbic acid and 2 ml 0.22 mol/L molybdate solutions were added and the tubes were shaken for 15 min. The absorbance at 700 nm was measured using a Spectrophotometer (manufactured by Agilent Technologies Inc.). The total N content was measured on soil samples that were digested in 98% H_2SO_4 by using the modified semimicro-Kjeldahl procedure. The potassium bichromate-concentrated sulfuric acid oxidation process was used to measure the TOC content which is briefly described here. First, 0.25~0.40 g of the prepared samples (weighting error of to 0.0001 g) was mixed with 10 ml of a 4 mol/L potassium dichromate-sulfuric acid standard solution and heated to 200~230 °C for 5±0.5 min. At the treatment, add 10 ml 4 mol/L the standard potassium dichromate into concentrated sulfuric acid medium. Under heating at 200~230 °C conditions for 5±0.5 min, the organic carbon in the sample is oxidized to carbon dioxide. That's it. Organic carbon sedimented on the bottom of the flask was oxidized and the concentration of potassium dichromate remaining in the supernatant was measured by titration with a 0.2 mol/L ferrous sulfate standard solution. The TOC content was calculated from the amount of consumed oxidant (ISSCAS 1987).

2.4 Statistical analysis

We used the Excel statistical software package version 2007 to sort data. The vertical trend charts of TOC, TN, and TP were made using the Origin Software Package Version 8.1. In this paper, the regression approach was used to the analysis of variance (ANOVA) in the SPSS software package version 17.0.

3. RESULTS

3.1 The spatial characteristics of the TOC, TP, and TN contents in the sediments

The TOC, TP, and TN concentrations varied in the three sampling sites from the Longbao Lake wetlands. With 0-50 cm sediments in Longbao Lake, the TOC, TP, and TN content were 3.81-15.6 g/kg (averaging 8.502 g/kg), 0.21-1.18 g/kg (averaging 0.614 g/kg), and 0.16-0.35 g/kg (averaging 0.216 g/kg), respectively (Fig. 2). In the three sampling sites, the TOC, TP, and TN content of N2 were higher than those of N1, and the N3 were the lowest (Fig. 2). The average TOC concentrations of N1, N2, and N3 were 10.062, 9.715, and 5.728 g/kg, respectively, while for TP were 0.795, 0.69, and 0.357 g/kg, respectively, and for TN they were 0.238, 0.222, and 0.190 g/kg, respectively (Fig. 2).

The TOC, TP, and TN in N1, N2, and N3 reduced gradually with the increase of sediment depth (Figs. 2-4). In the three sampling sites, the concentrations of TOC, TP, and TN in October were slightly higher than that in

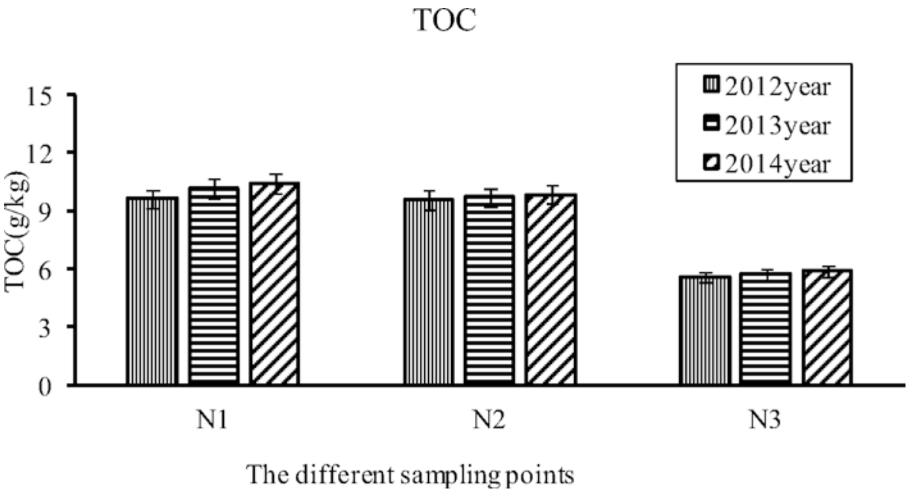


Fig. 2. Interannual variability of TOC content in sediments of different sampling sites.

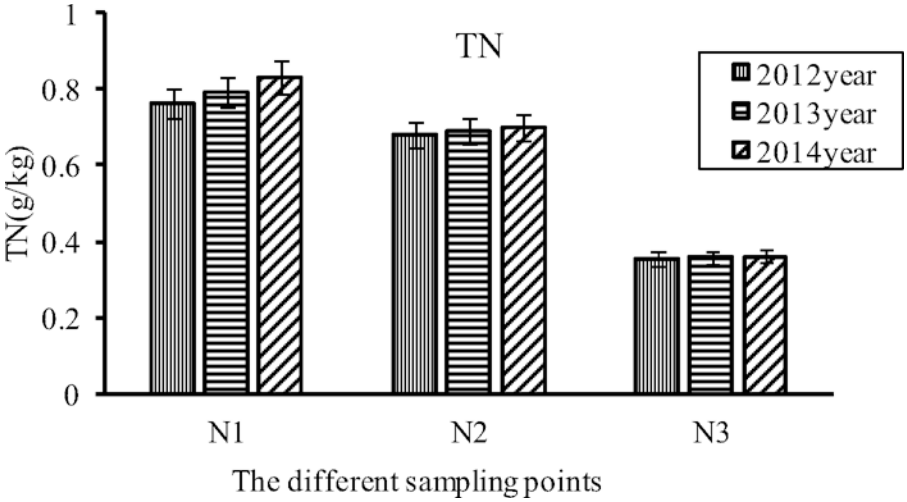


Fig. 3. Interannual variability of TN content in sediments of different sampling sites.

May (Figs. 2-4). In the three sampling sites, no significant difference was detected for TOC, TN, and TP, from 2011 to 2013. However, the average values on TOC, TN, and TP were in a continuous increasing tendency. In the three sampling sites, the TOC , TN, and TP in 2013 were increased by 3-8, 2-9, and 8-10%, respectively, compared to these in 2011 (Figs. 5-7).

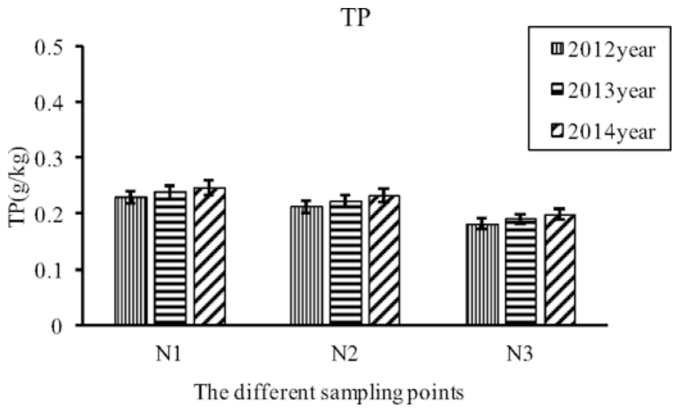


Fig. 4. Interannual variability of TP content in sediments of different sampling sites.

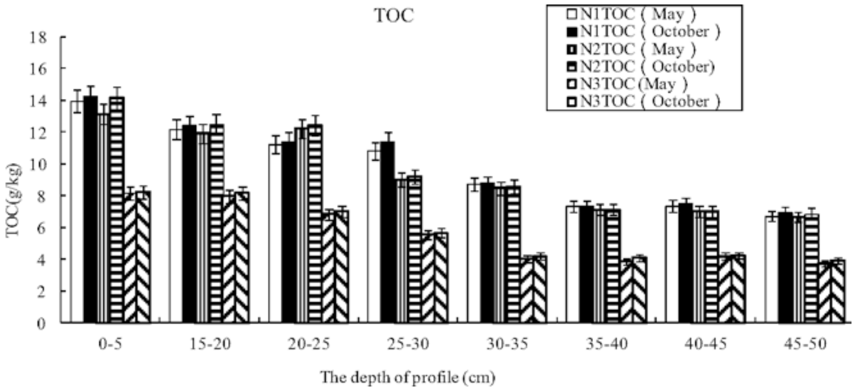


Fig. 5.Comparison of TOC content in October and May at different sediment depths of N1, N2, N3 sampling points.

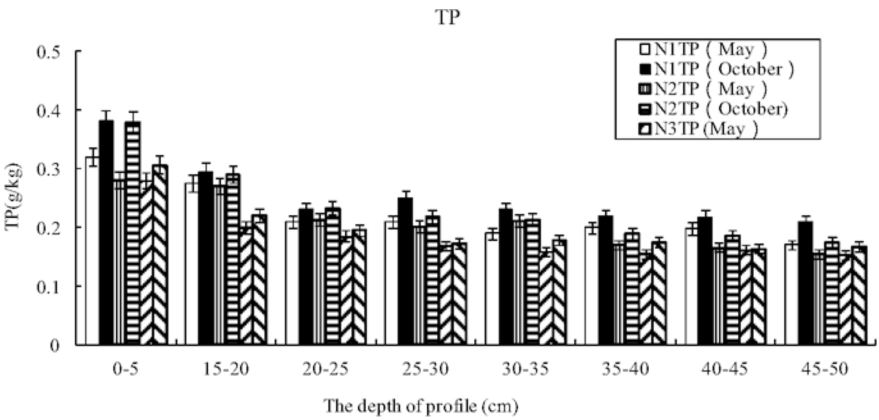


Fig. 6. Comparison of TP content in October and May at different sediment depths of N1, N2, N3 sampling points.

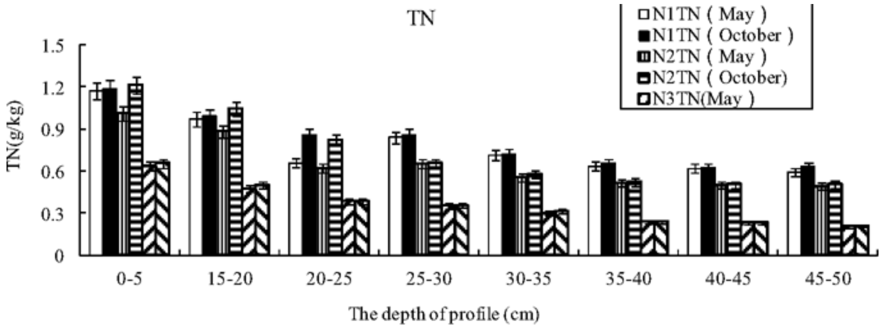


Fig. 7. Comparison of TN content in October and May at different sediment depths of N1, N2, N3 sampling points.

3.2 Correlation of TOC and TN, TN and TP

The TOC and TN contents varied from 3.81-15.6 g kg^{-1} to 0.21-1.18 g kg^{-1}, respectively (Figs. 8-11). Although the TOC and TN concentrations varied with soil depth, a strong correlation was observed between the TOC and TN concentrations (r^2 = 0.960). The TOC/TN ratio in the sediment, as an effective indicator of the source of organic matter (Li and Song 2005), varied from 11.14 to 18.14 in Longbao Lake, with an average of 14.31 (Fig. 12).

The TN and TP concentrations followed nearly the same trends through the soil profile (Figs. 9-11). The TN was positively correlated with TP in the soil profile (R^2=0.968), which indicated a strong correlation between the two nutrients in the wetland sediments of Longbao Lake. The TN/P ratios in the sediments from Longbao Lake varied from 1.31 to 3.76, with an average of 2.79 (Fig. 13). Thus, the same source contributed to the N and P pollution in Longbao Lake. The TN and TP concentrations were relatively higher in the surface sediments, which indicated that the Longbao Lake Wetland was surrounded by meadows, increasing TN and TP to the lake.

4. DISCUSSION

4.1 The TOC, TP, and TN contents in the sediments

The water quality with alpine lake is a great concern for protecting the ecosystem in San Jiang Yuan, which is the head water source of the Yangtze and Yellow Rivers. The documents revealed that with the water physical and chemical indexes in San Jiang Yuan, the wetlands had a great water quality with the type I standard for (GB3838-2002 standards for surface water environment quality) (Lu *et al*. 2009, Shi *et al*. 2009). In our study, with low nutrients inputs in 0-50 cm sediments (the 8.502 g/kg TOC, 0.614 g/kg TP, 0.216 g/kg TN), the water in the Longbao Lake had a good quality without eutrophication. Some research reported 12.9 ueq/L TN and 5.6 ug/L TP in sediment in European alpine lakes without eutrophication (Catalan *et al*.

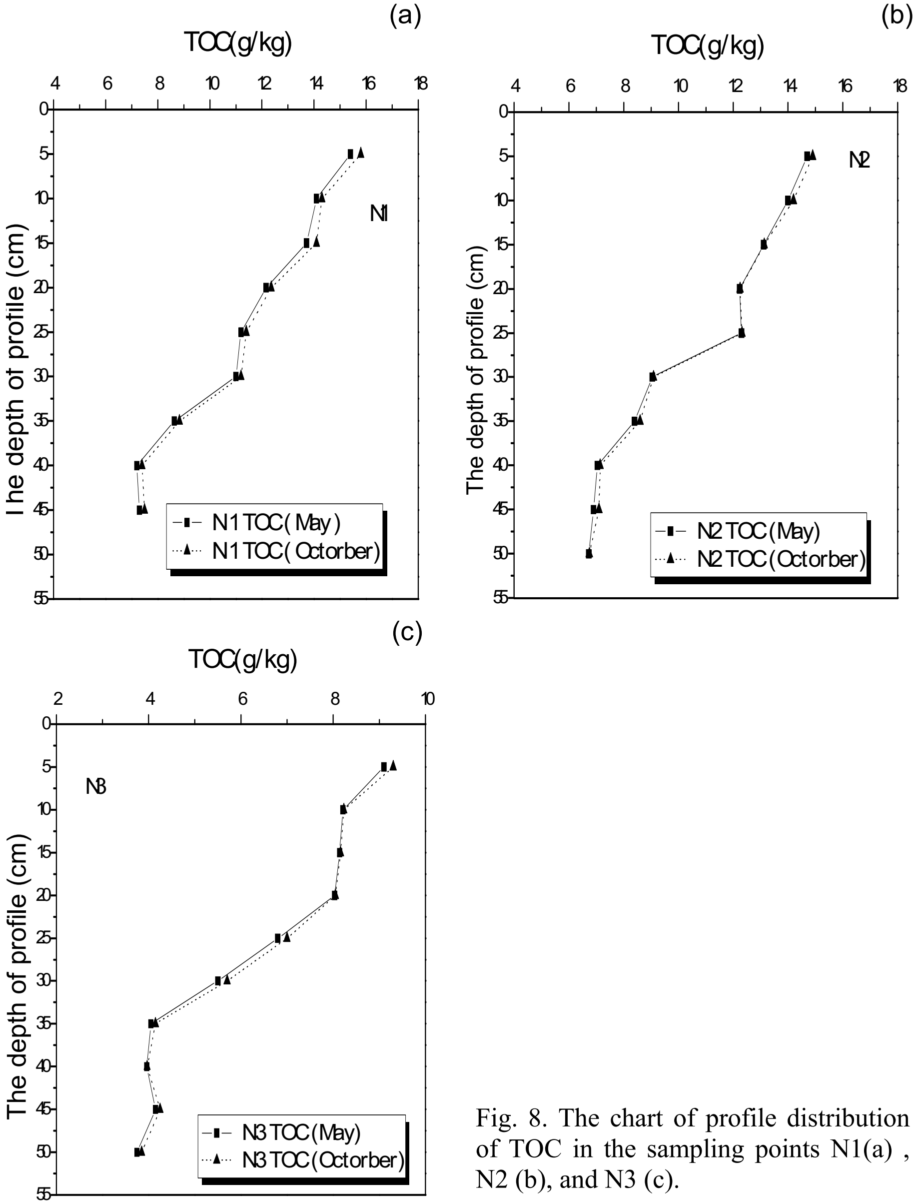


Fig. 8. The chart of profile distribution of TOC in the sampling points N1(a) , N2 (b), and N3 (c).

2009). However, eutrophication still should be a great concern for the Longbao Lake, based on N and P inputs in the sediment. In the three sampling years, from 2011 to 2013, the interannual significant difference was not observed (Figs. 5-7), but the slightly continuous increases on TOC, TN, and TP with the sediments suggested that more and more nutrients could be input

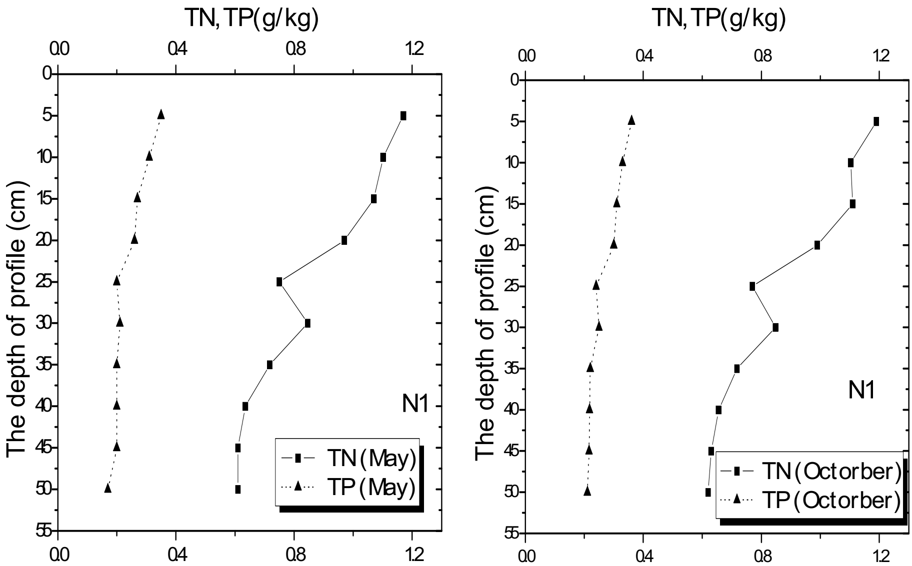


Fig. 9. The chart of profile distribution of TN and TP in the sampling points N1.

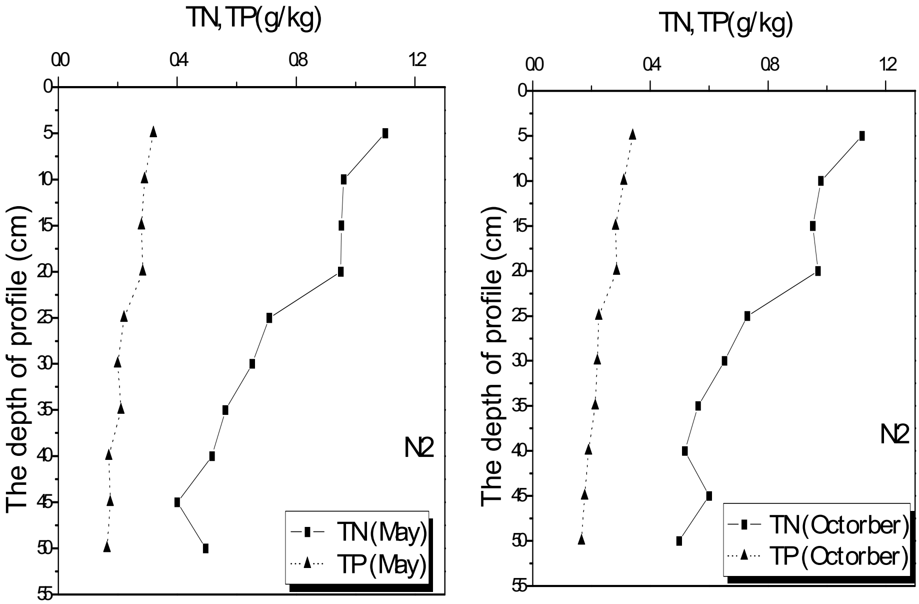


Fig. 10. The chart of profile distribution of TN and TP in the sampling points N2.

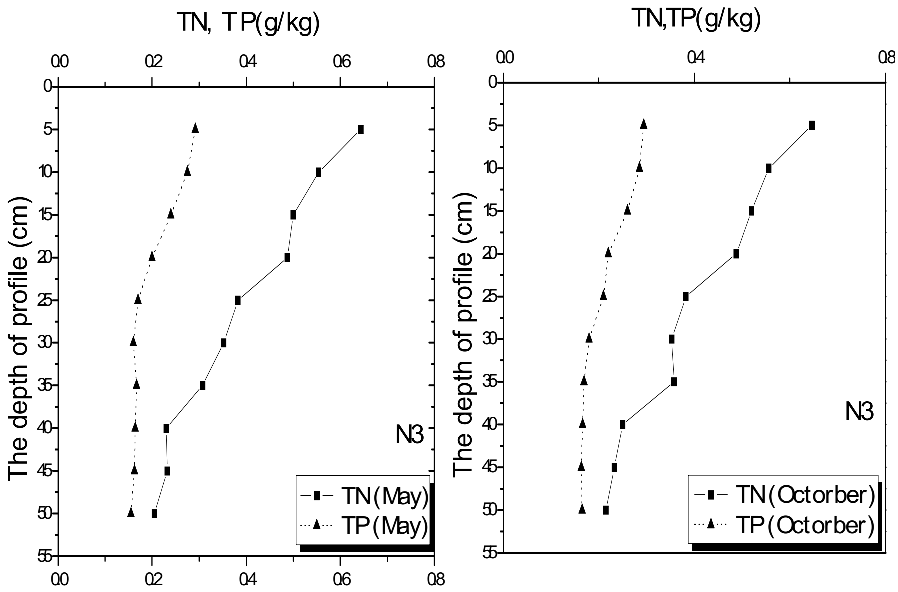


Fig. 11. The chart of profile distribution of TN and TP in the sampling points N3.

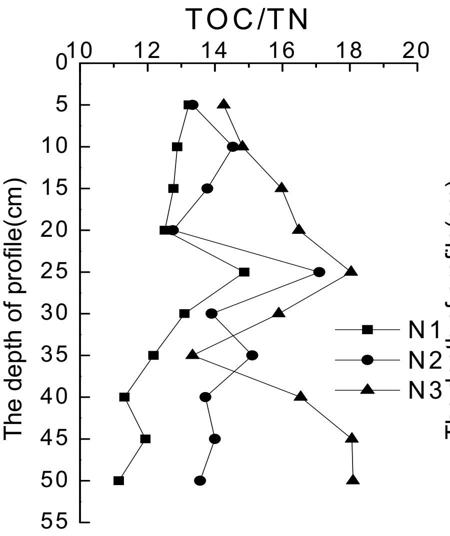


Fig. 12. The vertical distribution of TOC/TN of each sampling point.

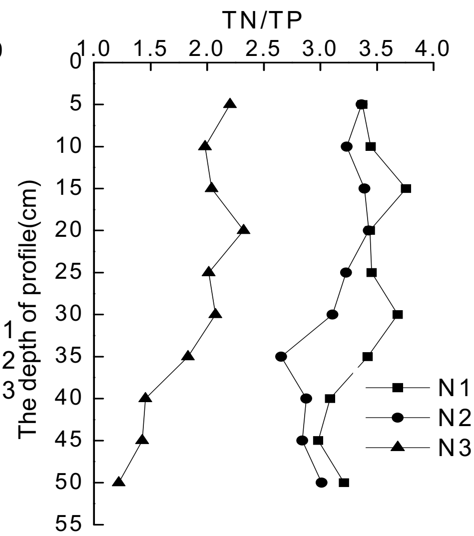


Fig. 13. The vertical distribution of TN/TP of each sampling point

into the lake, increasing the risk of eutrophication in the Longbao Lake. Therefore, the Longbao Lake had a high potential for eutrophication.

Our results showed that the concentrations of TOC, TP, and TN all reduced gradually with the increase of sediment depth (Figs. 8-11), attributed by the deposition. Lehtoranta *et al*. (1997) studied the accumulation of nutrients (N and P) in sediments in the Eastern Gulf of Finland, found that they had similar trends and showed that the mean TN and TP was approximately 40 and 50% higher in the sediment surface than in the deeper layer, respectively. Feng (2006) also reported a decrease of TN in the sediments with increasing depth, in Dianchi Lake in China.

In the three sampling sites, the TOC, TP, and TN, contents of N2 were higher than those of N1, and the N3 contents were the lowest (Figs. 2-4). The variance should result from the nutrients inputs surrounding the sampling sites. The sampling sites of N2 and N1 are, respectively, located in the west and east sides of the lake, closed to the lakeside with the wetlands. The great primary productivity of wetlands in summer could contribute to more nutrients with N, P and organic matter, which were input into the both sites. In addition, with grazing surrounding the lake, the nutrients with livestock manure could be transferred into the sites nearby, due to runoff. In the comparison of N1 and N2, the N3 site is located in the central part of the Lake; therefore, the productivity of wetlands and grazing had lesser effect on the nutrients inputs of the N3 site. Furthermore, the TOC, TP, and TN content in the three sites varied in different seasons (Figs. 2-4). Specifically, the TOC, TN, TP contents in October were higher than in May. During summer, from May to October, the wetlands and grasslands had a great productivity with more biomass, which input more nutrients into the water system, especially organic matter. The summer was not only the season for plants growing rapidly, but also for active growing and breeding period for animals and birds. On the one hand, some migratory birds, such as Black Stork (*Ciconia nigra*), Lark (*Alauda positos*), *etc*., would migrate to the Longbao Lake in May and leave in October. The birds lived in the wetland and produced bird manure into the wetland and water body. On the other hand, the grasslands surrounding lakes had a greatest grazing intensity, since the drinking water was available for livestock. The grazed livestock had a great activity on the grassland and wetland. And the increasing number of livestock from May to October would lead to more manure input into the ecosystem. And the greater rainfall in summer would cause more runoff that accelerated nutrients inputs into the lake. What's more, the greater number of tourists might affect the water quality. These reasons were combined to attribute to higher nutrients inputs in Longbao Lake in October.

4.2 Correlation of TN and TP, TOC and TN

In our results, a strong correlation was detected between TN and TP in the wetland sediments of Longbao Lake ($R^2 = 0.968$). The correlation between TN and TP testified that N and P in the sediment had the same source (Li and Song 2005). Xiang *et al*. (2014) did research on the N and P pollution in the eastern Bay of Tai Lake area, and indicated that TP and TN showed a significant positive correlation in the surface of sediments, so that the N and P in sediments result from similar sources. This conclusion is consistent with the study on Chao Lake (Li 2010).

Some studies showed that C/N ratio of sediments can effectively indicate the source of organic matter (Chen and Wan 2000). So the sediments in the lake had the two sources of organic matter, including endogenesis and exogenous. With the endogenesis, the catabolite of aquatic organisms contained more protein, and the TOC/TN value was generally less than 7. With exogenous, terrigenous higher plants were rich in lignin but less protein, with the TOC/TN values 20~30. In our research, the TOC/TN ratio in the sediment varied from 11.14 to 18.14 in the three sites in Longbao Lake (Fig. 12). It is stated that exogenous inputs of nutrients from grassland had a major effect on the sediment in Longbao Lake. Therefore, exogenous inputs of nutrients or pollution might be the major factor for the high potential for eutrophication in the Longbao Lake. The efficient solutions, for avoiding eutrophication and protecting the wetland in the Longbao Lake, should be reducing grazing intensity, and lowing the pollution and damage caused by human activities.

5. CONCLUSION

With the profile depths, the content of TOC, TP, TN with sediments in the Longbao Lake appeared in a decreasing trend. The higher nutrients were input into Longbao Lake in October, compared to May, due to the greater plant growth, and higher intensity on animals, birds and grazing, and so on. The significant correlation between TN and TP indicated that N and P in the sediment had the same source. Although the Longbao Lake had a great water quality, the slightly continuous increases on TOC, TN, and TP with the sediments suggested a high potential for eutrophication. Since the TOC/TN ratio in the sediment varied from 11.14 to 18.14 in the three sites in Longbao Lake, exogenous inputs of nutrients or pollution might be the major factor for the high potential for eutrophication in the Longbao Lake. Therefore, the human activities surrounding the Longbao Lake, especially grazing, should be the great concern for environmental protection in the Longbao Lake.

Acknowledgements. This research was financially supported by The National Natural Science Funds Fund (No. 31260128), international cooperation project of Qinghai Province Science and Technology Agency (2013-H-806), and China New Zealand plateau grassland nutrient flow and sustainable production research (2015DFG31870) international cooperation project of the Ministry of science and technology; this study is finished in the laboratory of Eco-environmental Engineering College of Qinghai University, thanks for providing the testing facility and accessories in the laboratory measurements.

References

Barbier, E.B., M.C. Acreman, and D. Knowler (1997), *Economic Valuation of Wetlands: A Guide for Policy Makers and Planners*, Ramsar Convention Bureau, Gland, Switzerland.

Catalan, J., M.G. Barbieri, F. Bartumeus, P. Bitusik, I. Botev, A. Brancelj, D. Cogalniceanu, M. Manca, A. Marchetto, N. Ognjanova-Rumenova, S. Pla, M. Rieradevall, S. Sorvari, E. Stefkova, E. Stuchlik, and M. Ventura (2009), Ecological thresholds in European alpine lakes, *Freshwater Biol.* **54**, 12, 2494-2517, DOI: 10.1111/j.1365-2427.2009.02286.x.

Chen, J.A., and G.J. Wan (2000), Study on the environmental record of modern sediments in the Yunnan Cheng Sea, *J. Mineral.* **20**, 2, 112-116.

Chen, Y.F., H. Liu, W.T. Zou, and H.Q. Zhang (2012), Quantitative study on the drive factors of wetland change in three river's source area, *Forest Sci.* **25**, 5, 545-550.

Ding, S.M., C. Han, Y.P. Wang, L. Yao, Y. Wang, D. Xu, Q. Sun, N.W. Paul, and C.S. Zhang (2015), In situ, high-resolution imaging of labile phosphorus in sediments of a large eutrophic lake, *Sci. Direct.* **74**, 100-109, DOI: 10.1016/j.watres.2015.02.008.

Feng, F. (2006), Study on the Vertical Distribution of Sedimentary Microbial Biomass, Species of Carbon, Nitrogen, Phosphorus and their Correlation, The Chinese Academy of Sciences Institute of Hydrobiology, Wuhan.

Guo, Z.Y. (2007), Distribution and Transformation of Phosphorus Fractionations in the Sediments of Three Typical Urban Shallow Lakes-Xuanwu Lake, Damning Lake and Mochou Lake, Hehai University, Nanjing.

ISSCAS (1987), *Physical and Chemical Analysis of Soil*, Science and Technology Press, Shanghai, Institute of Soil Science, Chinese Academy of Sciences, Nanjing, China, 189-194.

Lehtoranta, J., H. Pitkanen, and O. Sandman (1997), Sediment accumulation of nutrients(N, P) in the Eastern Gulf of Finland (Baltic Sea), *Water Air Soil Poll.* **99**, 477-486.

Li, L. (2010), Studies on the Distribution of Organic Carbon, Nitrogen, Phosphorous and Their Correlation in ChaoHu Wetland Sediments, Anhui Normal University, Hefei.

Li, R.X., M.Y. Zhu, S. Chen, R.H. Lu, and B.H. Li (2001), Responses of phytoplankton on phosphate enrichment in mesocosms, *Acta Ecol. Sin.* **21**, 4, 603-607.

Li, X.G., and J.M. Song (2005), Source and biogeochemical characteristics of nitrogen and phosphorus in Jiaozhou Bay sediments, *Oceanol. Limnol. Sin.* **36**, 6, 562-571.

Lin, J.C. (2003), Biological diversity conservation in Yushulong National nature reserve of Qinghai province, *Jilin Agr. Sci. Technol.* **32**, 4, 37-40.

Lin, Q. (2008), Longbao Lake: home of black-necked crane, *Hum. Nature* **4**, 7, 56-57.

Lu, S., H. Shi, P. Li, and Y.L. Yang (2009), Assessment on the currently environmental impact of surface water in Yangtze River of the source of three rivers, *Environ. Health* **26**, 7, 604-605.

Ma, Z.T., F.X. Li, P. Li, and J.S. Xiao (2009), The study of dynamic change of ecological environment of Longbao area in Qinghai Yushu, *Pratacultur. Sci.* **26**, 7, 6-11.

Mitch, W.J., and J.G. Gosselink (2000), *Wetland*, 3rd ed., John Wiley & Sons, New York.

Qing, D.H. (2013), The Source Region of Ecological Protection and Sustainable Development in Sanjiang, Science Publishing House, Beijing, 29-30.

Shi, H.X., S.J. Lu, P. Li, and Y.L. Yang (2009), Assessment on the currently environmental impact of surface water in Yellow River of the source of three rivers, *Anhui Agr. Sci.* **17**, 1134-1136.

Søndergaard, M., E. Jeppesen, T.L. Lauridsen, C. Skov, N.E.H. Van, R. Roijackers, E. Lammens, and R. Portielje (2007), Lake restoration:successes, failures and long-term effects, *J Appl. Ecol.* **44**, 6, 1095-1105, DOI: 10.1111/ j.1365-2664.2007.01363.x.

Sun, Z.G., and J.S. Liu (2006), The actuality, problems and sustainable utilization countermeasures of wetland resources in China, *J. Arid Land Resour. Environ.* **20**, 2, 83-88.

Xiang, S.L., M.Y. Zhu, G.W. Zhu, and H. Xu (2014), Pollution charateristics of nitrogen and phosphorus in the sediment of the growth area of the aquatic plants in the east of Taihu Lake, *Acta Sediment. Sin.* **32**, 6, 1083-1088.

Yang, S.Z., H.J. Jin, Y.J. Ji, Z. Wei, and R.X. He (2008), Revegetation in Permafrost Regions along a Linear Project, *J. Glaciol. Geocryol.* **30**, 5, 875-822.

Zheng, J. (2011), *Study of Nature Protection Area of Qinghai*, Qinghai People's Publishing House, Xining.

Received 28 May 2015
Received in revised form 16 February 2016
Accepted 13 April 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2487-2509
DOI: 10.1515/acgeo-2016-0095

Experimental and Numerical Analysis of Air Trapping in a Porous Medium with Coarse Textured Inclusions

Paulina SZYMAŃSKA[1], Witold TISLER[1], Cindi SCHÜTZ[2], Adam SZYMKIEWICZ[1], Insa NEUWEILER[2], and Rainer HELMIG[3]

[1]Gdańsk University of Technology, Faculty of Civil and Environmental Engineering, Department of Geotechnics, Geology and Marine Civil Engineering, Gdańsk, Poland; e-mail: adams@pg.gda.pl (corresponding author)

[2]Institute of Fluid Mechanics and Environmental Physics in Civil Engineering, Leibniz University Hannover, Hannover, Germany

[3]Institute for Modelling of Hydraulic and Environmental Systems, University of Stuttgart, Stuttgart, Germany

Abstract

The paper presents a 2D upward infiltration experiment performed on a model porous medium consisting of fine sand background with two inclusions made of coarser sands. The purpose of the experiment was to investigate the effects of structural air trapping, which occurs during infiltration as a result of heterogeneous material structure. The experiment shows that a significant amount of air becomes trapped in each of the inclusions. Numerical simulations were carried out using the two-phase water-air flow model and the Richards equation. The experimental results can be reproduced with good accuracy only using a two-phase flow model, which accounts for both structural and pore-scale trapping. On the other hand, the Richards equation was not able to represent the structural trapping caused by material heterogeneity.

Key words: vadose zone, Richards equation, heterogeneous soils, air trapping.

Ownership: Institute of Geophysics, Polish Academy of Sciences;

1. INTRODUCTION

The pore space of soils and rocks in the vadose zone is filled partly with water and partly with air. In many practical problems, the focus is primarily on the flow of water and the air flow is neglected. This approach is justified by assuming that the pore air is continuous, connected to the atmospheric air and much more mobile than the pore water. This implies that any perturbation of the air pressure instantly equilibrates with the atmosphere and the pore air remains essentially at constant atmospheric pressure. In such a case the water flow can be described by a single equation, known as the Richards equation (RE) (Richards 1931). In contrast, if the above assumptions are not satisfied, one should use a full two phase model (2PH), which accounts explicitly for the flow of both water and air, caused by potential gradients within each phase (*e.g.*, Helmig 1997).

While the RE is commonly used in hydrological and geotechnical modeling, a number of authors showed its limitations caused by various factors. First, the mobility ratio between air and water is not infinite and even at medium water saturation levels there are some discrepancies in the results obtained with RE and 2PH approaches, *e.g.*, Tegnander (2001). Close to full water saturation, the smaller viscosity of air is offset by its smaller relative permeability, which makes the mobilities of the two fluids similar. In such a case, important differences between the models can be observed (Forsyth 1988). Other well known situations where RE is not applicable concern infiltration in laboratory columns with sealed bottom (Vachaud *et al.* 1973, Touma *et al.* 1984, Touma and Vauclin 1986), ponded infiltration over large area (Hammecker *et al.* 2003, Delfs *et al.* 2013) or overtopping of earth dikes by flood wave (Bogacz *et al.* 2006, Leśniewska *et al.* 2008).

Attention has been also directed to the impact of heterogeneous soil structure on the patterns of air and water flow. In this context the importance of the so-called air-entry barriers has been emphasized (Silliman *et al.* 2002). They are regions of fine textured material, remaining fully or nearly fully water-saturated even at significant negative water pressures, until their characteristic value of the air entry pressure is exceeded. Experiments in a two-dimensional flume described in (Silliman *et al.* 2002) showed that thin layers or inclusions of coarse material present in the capillary fringe did not drain due to the presence of quasi saturated fine material around them, and provided important paths for horizontal tracer migration. Other experiments (Kuang *et al.* 2011) showed that drainage from a laboratory column filled with coarse sand is significantly hampered by the presence of a thin layer of fine sand at the upper surface of sand. The fine layer remained quasi-saturated with water and consequently had very low permeability to air, causing retarded drainage, while the suction in the coarse sand remained higher than it would result from the water retention curve. Several contributions focused on heterogeneity patterns consisting of coarse textured inclusions embedded in fine textured background (Dunn and Silliman 2003, Dunn 2005, Vasin *et al.* 2008, Szymkiewicz *et al.* 2012, 2014; Haberer *et al.* 2015). If the inclusions are disconnected from each other and from the atmospheric air, two types of interesting phenomena can be observed. First, if the heterogeneous medium is subject to drainage starting from full water saturation, the inclusions remain saturated until the suction exceeds the entry pressure of the fine background material, which can be much larger than their own entry pressure. Second, during imbibition the background material becomes fully or quasi fully saturated at relatively large negative pressure, corresponding to a much lower saturation in the coarse material. However, once the pores of the background material are completely filled with water, the air from coarse inclusions cannot escape. Even if the water pressure in the system becomes positive (larger than atmospheric), the inclusions remain unsaturated. Both types of behavior were confirmed by laboratory experiments and numerical analyses (Dunn and Silliman 2003, Dunn 2005, Vasin *et al.* 2008). Such effects were also investigated for other types of two phase flow, *e.g.*, in water-NAPL or water-CO2 systems (Dunn 2005, Szymkiewicz *et al.* 2011, Saadatpoor *et al.* 2009). If the heterogeneous medium contains a large number of inclusions, it is possible to derive an upscaled model with effective parameters, which is characterized by hysteresis of the water retention function and the phase permeability functions (Mikelic *et al.* 2002, van Duijn *et al.* 2007, Schweizer 2008, Szymkiewicz *et al.* 2011, 2012; Szymkiewicz 2013). One should also note that the effects occurring on coarse-fine soil interfaces are used in capillary barriers protecting waste storage site from infiltration or in the drainage layers placed under lawns. These structures are composed of a fine-textured soil layer supporting vegetation with underlying coarse grained material. In unsaturated conditions the coarse material has very low permeability, which effectively limits the downward infiltration. Only if the water pressure at the interface reaches values close to zero, a significant amount of water can enter the coarse material and is rapidly transmitted by this layer. Capillary barriers were a subject of numerous modeling studies (*e.g.*, Oldenburg and Pruess 1993, Aubertin *et al.* 2009, Prédélus *et al.* 2015). Assuming that the air in coarse layer can move unobstructed downward and laterally, the Richards' equation can be used with sufficient accuracy for such analyses (Webb 1998).

The entrapment of air in coarse inclusions during imbibition, which will be called here structural trapping, is a phenomenon qualitatively distinct from the pore-scale entrapment of isolated air bubbles. The latter process is well recognized in the literature and has been a subject of several studies, *e.g.*, Marinas *et al.* (2013), McLeod *et al.* (2015). Air bubbles in pores typically arise during infiltration, but may be also products of various biological and chemical processes in porous medium. In order to account for their presence, the concept of "field-saturated" conditions is often used in vadose zone modeling. The maximum attainable value of the volumetric water content θ_{sw} is smaller than the porosity ϕ. The residual air saturation varies in a wide range, from 4 to 54% of the soil porosity (Marinas *et al.* 2013). Consequently, the field-saturated hydraulic conductivity is also smaller than the "true" saturated conductivity, even up to 10 times (Marinas *et al.* 2013).

In natural soils we can expect overlapped effects of structural and pore-scale air trapping, occurring especially in the zone of water table fluctuations and the adjacent capillary fringe. This transition zone, termed "partially saturated fringe" (Berkowitz *et al.* 2004) is increasingly recognized as crucial for contaminant transport and various biochemical processes in soils (Berkowitz *et al.* 2004, Silliman *et al.* 2002, Yakirevich *et al.* 2010), since the entrapped air determines oxygen availability (Haberer *et al*. 2015). More research is needed to better understand the air and water dynamics in the vicinity of the water table. This paper presents the results of a laboratory experiment on a model porous medium with separated coarse inclusions, subject to upward infiltration. We focus on the combined effects of structural and pore scale trapping. Moreover, we compare numerical solutions obtained with the Richards equation and the two phase flow model. To the best of our knowledge, such a comparison of RE and 2PH models with experimental results for heterogeneous media has not been undertaken before, since in previous contributions only one of those two approaches was used, *e.g.*, 2PH model in (Dunn 2005) or RE in (Vasin *et al.* 2008).

2. EXPERIMENTS

Experiments were carried out in the laboratory of the Institute of Fluid Mechanics and Environmental Physics in Civil Engineering, Leibniz University, Hannover (Szymańska 2012). The main part of the experimental setup was a glass flume having dimensions 55 cm by 29 cm by 1.2 cm, which allowed to impose conditions of two-dimensional flow. Three types of well sorted sand were used, which will be denoted here as fine, medium and coarse (see Table 1). Fine sand was used as the background material, filling most of the flume, while medium and coarse sands formed two inclusions placed symmetrically (Fig. 1). The purpose of the experiment was to investigate the air trapping in inclusions during upward infiltration. Water entered the flume from two inlets located at the bottom of the device.

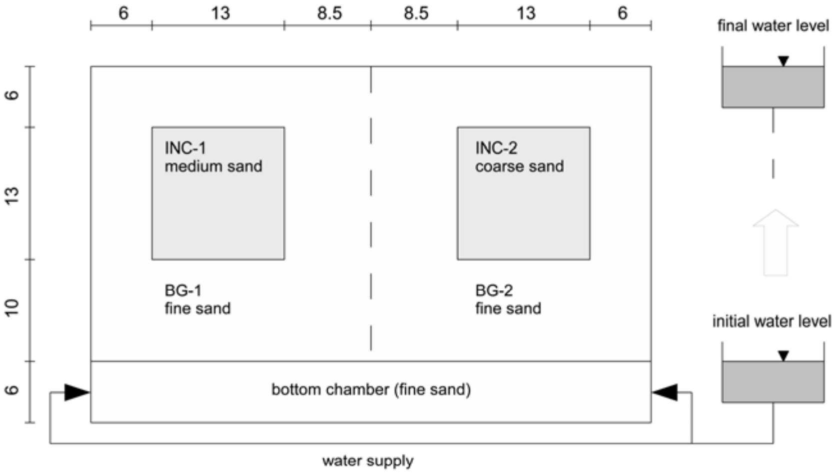
In order to facilitate observation of the changing saturation field, water was dyed to a blue colour. During the experiment, pictures of the flume were taken by a camera placed in front of it. The saturation field was obtained by

Table 1

Parameters of three sands used in the experiment

Parameter	Unit	Coarse	Medium	Fine
d	[mm]	0.7 to 1.2	0.4 to 0.8	0.1 to 0.3
ϕ	[–]	0.540	0.450	0.370
MVG – measurements				
θ_{sw}	[–]	–	0.360	0.280
θ_{rw}	[–]	0.00	0.00	0.02
$p_{g;d}$	[Pa]	1143	1990	4004
$p_{g;i}$	[Pa]	565	–	–
n	[–]	6.07	4.86	3.35
k_s	[m^2]	$1.57 \cdot 10^{-11}$	$8.66 \cdot 10^{-12}$	$4.48 \cdot 10^{-12}$
BCB0 – first estimation				
θ_{sw}	[–]	0.540	0.450	0.370
θ_{rw}	[–]	0.00	0.00	0.00
$p_{e;i}$	[Pa]	388	746	1582
λ	[–]	2.86	2.25	1.47
k_s	[m^2]	$1.57 \cdot 10^{-11}$	$8.66 \cdot 10^{-12}$	$4.48 \cdot 10^{-12}$
BCB1 – best fit				
θ_{sw}	[–]	0.443	0.370	0.355
θ_{rw}	[–]	0.00	0.00	0.00
$p_{e;i}$	[Pa]	300	500	1100
λ	[–]	4.0	4.0	2.8
k_s	[m^2]	$2.56 \cdot 10^{-11}$	$1.86 \cdot 10^{-11}$	$1.44 \cdot 10^{-11}$

image analysis techniques. A digital camera (Nikon D90) was placed in front of the experimental device. Still pictures were taken in 10 s time intervals. The resolution of the pictures was 4288 by 2848 pixels. During image processing, the number of pixels representing dyed water was used to obtain the average saturation in each of the three materials. The reference level of saturation was obtained from an independent experiment, where the flume was filled with sands forming three vertical strips parallel to each other, and having the same porosity as in the main experiment. In contrast to the principal experiment, the device was flushed with CO_2 and slowly saturated by upward water infiltration. In such conditions, no structural trapping could occur and the pore trapping was negligibly small, so it was assumed that the image parameters correspond to full saturation in each material. The total volume of water in the flume was monitored with a balance.



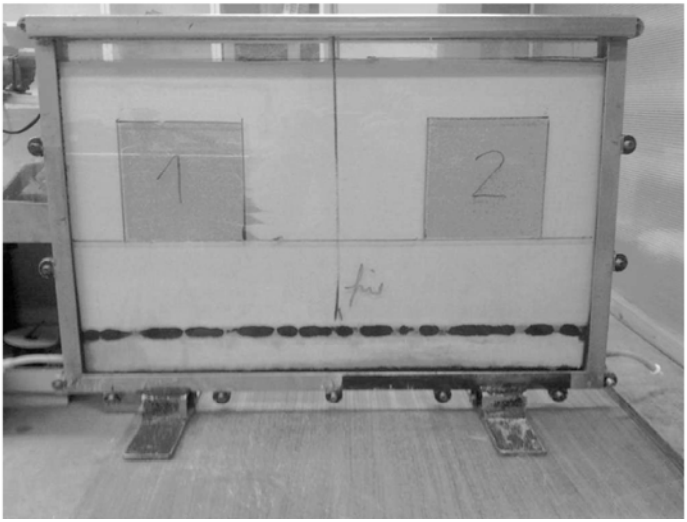


Fig. 1. Scheme of the experimental setup (top) and picture of the flume after packing (bottom). Dimensions in cm.

Initially all three sands were in air dry conditions. After filling the flume, the inlets were connected to a water-filled reservoir with water level at the same elevation as the bottom of the flume. The water level in reservoir was kept constant for 12 hours, in order to approach the state of capillary equilibrium. During this time, capillary forces caused significant increase of saturation in the fine material; however, the two inclusions remained virtually dry (Fig. 2, $t = 0$).

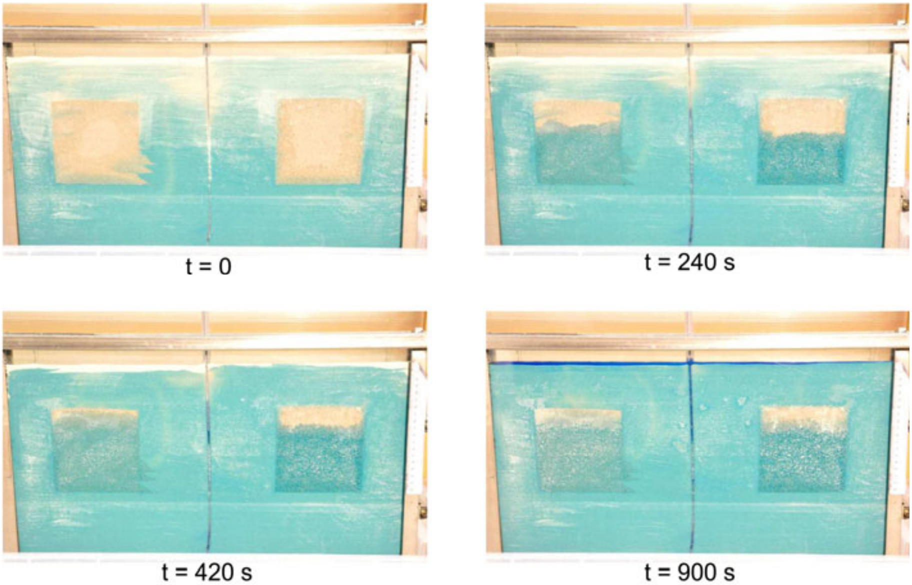


Fig. 2. Distribution of dyed water in the flume at four stages of the experiment.

The actual experiment was started by rising the reservoir with water in such a way that the water level corresponded to the upper surface of the fine sand in the flume. The rising took about 10 s. The resulting increase of positive water pressure at the bottom of the flume caused upward infiltration. Water distribution at various stages of infiltration is shown in Fig. 2 ($t = 240$, 420, and 900 s). A stable configuration of water distribution was achieved about 600 s after the rising of the water table started. The observations were continued for another 25 min, without any visible changes in the saturation distribution inside the sand. At about 900 s a thin layer of free water appeared at the surface of fine sand, indicating that equilibrium with supplying water reservoir had been reached.

From Fig. 2 it can be clearly seen that the upward movement of water occurs principally in fine sand, where capillary forces are relatively strong. The saturation process in the inclusions starts from the bottom, where the water pressure is high enough to ensure sufficient hydraulic conductivity to enter the coarse porous medium. A clear horizontal boundary can be seen between the saturated and dry zone in the inclusions. There is apparently no flow into the inclusions from the upper part of their perimeter, even though they contact fine sand at relatively high saturation. Only in the later stages of infiltration, the boundary between dry and wet zone becomes more diffuse. The inclusions remain only partly saturated till the end of the experiment.

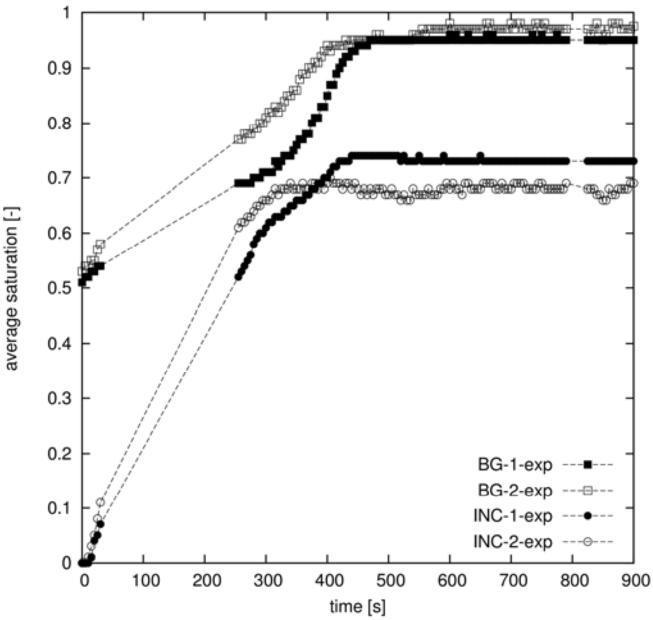


Fig. 3. Evolution of average water saturation in four regions of the flume during the experiment, as obtained by image analysis.

One can clearly see an unsaturated region at the top of each inclusion, and also many small air bubbles in the lower parts of inclusions. This indicates the occurrence of both structural trapping and pore scale trapping.

Figure 3 presents plots of the average saturation evolution in time, as obtained from the image analysis. There are four average values, corresponding to four regions of the plume. The fine sand background was divided into two regions, the left-hand one, BG1, around the medium sand inclusion INC1, and the right-hand one, BG2, around the coarse sand inclusion INC2. These plots confirm visual observations. It can be seen that at $t = 0$, when the medium is in a state close to hydrostatic equilibrium, the average saturation in the fine sand is close to 0.5, while the saturation in inclusions is practically zero. After the rising of the water table saturation in all regions increases, with higher increase rate in inclusions than in the background material. Neither the background nor inclusions become fully saturated during the experiment, however the background material becomes very close to full saturation, with the maximum saturation value about 0.97. On the other hand, the final values of saturation in inclusions are 0.73 for the medium sand and 0.69 for the coarse sand. These are average values, which account for both structural and pore-scale trapping. Starting from about 600 s, a stable state of saturation is reached, which lasts till the end of the experiment.

3. NUMERICAL MODELING

The two phase flow model used in this paper is based on the following assumptions: the fluids are immiscible (due to the short time of the experiments the dissolution of air in water is not modelled); the flow is isothermal; each porous material is rigid and has isotropic permeability. Under these assumptions the governing equations for either water or air, representing the mass balance principle for each fluid, can be written as follows (*e.g.*, Helmig 1997):

$$\frac{\partial}{\partial t}(\rho_\alpha \phi S_\alpha) + \nabla \cdot (\rho_\alpha \mathbf{v}_\alpha) = 0 \ , \tag{1}$$

where α is the fluid phase index (a = air, w = water), ρ_α the fluid density, ϕ the porosity of the medium, S_α the fluid saturation and $\mathbf{v}_\alpha$ the volumetric fluid flux (or seepage velocity). According to the Darcy's law for isotropic medium the flux can be defined as:

$$\mathbf{v}_\alpha = -\frac{k_s \, k_{r\alpha}}{\mu_\alpha}(\nabla p_\alpha - \rho_\alpha \mathbf{g}) \ , \tag{2}$$

where k_s is the intrinsic permeability, $k_{r\alpha}$ the relative permeability for phase α, μ_α fluid viscosity, p_α fluid pressure (or, more generally, pressure potential), and $\mathbf{g}$ the gravitational acceleration vector.

The governing equations have to be completed by additional relationships. The difference in pressure potentials between air and water is known as the capillary pressure or suction and is related to the water saturation via the water retention function. Moreover, the relative permeabilities depend on the fluid saturations. These interdependencies are commonly described using the Mualem−van Genuchten (MVG) model (Mualem 1976, van Genuchten 1980):

$$S_e = \left[1 + \left(\frac{p_c}{p_g}\right)^n\right]^{-m} \tag{3}$$

$$k_{rw} = \sqrt{S_e}\left[1 - \left(1 - S_e^{1/m}\right)^m\right]^2 \tag{4}$$

$$k_{ra} = \sqrt{1 - S_e}\left(1 - S_e^{1/m}\right)^{2m} \tag{5}$$

where S_e is the effective water saturation, p_c the capillary pressure or suction ($p_c = p_a - p_w$), p_g the pressure scaling parameter, and n, m are shape parameters ($m = 1 - 1/n$). The effective water saturation represents values of saturation normalized between residual states of water and air:

$$S_e = \frac{S_w - S_{rw}}{1 - S_{rw} - S_{ra}} = \frac{\theta_w - \theta_{rw}}{\theta_{sw} - \theta_{rw}} \tag{6}$$

where S_{rw} and S_{ra} are residual saturations of water and air, respectively, θ_w is the volumetric water content, θ_{rw} the residual water saturation, and θ_{sw} is the water content at apparent saturation, with air phase at residual state.

Another widely used set of functions is the Brooks−Corey−Burdine (BCB) model (Brooks and Corey 1964, Burdine 1953):

$$S_e = \begin{cases} \left(\dfrac{p_c}{p_e}\right)^{\lambda} & \text{if} \quad p_c > p_e \\ 1 & \text{if} \quad p_c \le p_e \end{cases} \tag{7}$$

$$k_{rw} = S_e^{3+2/\lambda} \tag{8}$$

$$k_{ra} = \left(1 - S_e\right)^2 \left(1 - S_e^{1+2/\lambda}\right) \tag{9}$$

where p_e is the air-entry pressure, and λ is the shape parameter. In the BCB model the air entry effect is explicitly represented, *i.e.*, the porous material remains saturated with water until a specific value of the capillary pressure is exceeded. In the MVG model the air entry pressure is not accounted for, although for large values of n the effective water saturation remains close to 1 for a range of positive capillary pressure values. The following relationships between parameters of the two models have been suggested (Lenhard *et al.* 1989):

$$\lambda = \left(n-1\right)\left(1 - 0.5^{n/(n-1)}\right) \tag{10}$$

$$p_e = p_g\, \varsigma^{1/\lambda} \left(\varsigma^{n/(n-1)} - 1\right)^{1/n} \tag{11}$$

$$\varsigma = 0.72 - 0.35 \exp\left(-n^4\right) \tag{12}$$

The Richards' equation (Richards 1931) represents a special case of two-phase model, obtained for the assumption of constant air pressure. In such a case the capillary pressure is uniquely defined by the water pressure and the relative permeability and water saturation depend only on the water pressure. Mass balance is considered only for the water phase.

In order to facilitate comparisons, all numerical simulations described in this paper and the following sections were carried out using the same approach to discretize 2PH and RE mathematical models. The algorithm is based on vertex centered finite volume discretization in space and fully implicit first-order finite difference discretization in time. For each mathematical model the discretization results in a system of nonlinear algebraic equations, which must be solved with respect to the primary unknowns. These unknowns represent the nodal values of the water pressure, and additionally for 2PH also nodal values of the water saturation. The grid consisted of uniform square cells, 1 cm by 1 cm. The computational nodes correspond to grid vertices, while the material properties are assigned to grid cells since the material properties are assigned to grid cells and not to vertices (nodes), the water saturation was discontinuous at vertices positioned on inclusion boundaries. For these nodes, as the primary variable we chose the saturation from the material with the lowest value of the entry pressure and we calculate the saturation in the other material from the extended capillary equilibrium condition (de Neef and Molenaar 1997). Further details of the algorithm can be found in Szymkiewicz (2013).

In all simulations the same initial and boundary conditions were used. As the initial condition we assumed a uniform value of air pressure (p_a atmospheric) and hydrostatic distribution of water pressure, with p_w at the initial position of the water table (Fig. 1). The boundary condition at the bottom was given in terms of the water pressure calculated from the position of the water table in the supplying reservoir. The value increased from the initial one to the final one within 10 s and then was kept constant till the end of the simulations. The water saturation at the bottom was assumed to be $S_w = 1$ for the whole duration of the experiment. These values of water pressure and saturation were applied along the whole bottom edge of the flume, even though in reality the water was supplied from two inlets located at the sides of flume just above the bottom. However, visual examination shows a rather uniform upward movement of the infiltration front in fine sand (Fig. 2), which justifies the application of a uniform boundary condition along the bottom of the flume. At the top of the flume a no-flow condition was specified for water, while air was kept at a constant atmospheric pressure. The vertical sides of the flume were modeled as impermeable to both water and air.

4. INITIAL ESTIMATION OF MATERIAL PARAMETERS

In order to obtain hydraulic characteristics of the three sands, experiments were carried out at the Helmholtz Centre for Environmental Research − UFZ. The retention curve for each sand was obtained from multistep outflow experiments, starting from fully saturated conditions. Additionally, the saturated hydraulic conductivity of each sand was also measured. The parameters of MVG model were fitted to the results of the outflow experiment. The obtained values are listed in Table 1 (denoted by MVG model – measurements).

The parameters obtained in drainage (outflow) experiments cannot be used directly in infiltration experiments, due to hysteresis of the water retention curve (*e.g.*, Luckner *et al.* 1989). Also, the samples of sand used in the independent experiments were more compacted than the sand in 2D flume, as indicated by the comparison of θ_{sw} for MVG with porosity ϕ in Table 1 (the outflow experiments started with full saturation, so θ_{sw} corresponds to the porosity in this case). It is well known that the porosity and packing method significantly affects hydraulic parameters of soils.

For our preliminary numerical simulations we derived a set of hydraulic parameters based on the independent experiments described above and the following assumptions:

- The MVG water retention functions for imbibition are characterized by the same values of the parameter n as for drainage, while the values of the pressure scaling parameter p_g are two times smaller. The latter assumption is supported by the literature (Luckner *et al.* 1989, Likos *et al.* 2014) and by measurements of the water retention curve in imbibition conditions for the coarse sand, which yielded $p_{g;i} = 565$ Pa, *i.e.*, nearly exactly $p_{g;d}/2$.
- In order to account explicitly for the air entry pressure, the water retention functions described by MVG model are converted to functions described by BCB model, according to Eqs. 10-12. Preliminary numerical experiments (Szymańska 2012) showed that the BCB model was more suitable than the MVG model to represent the macroscopic air trapping caused by material heterogeneities. Note that due to looser packing of sand in the 2D flume experiments, resulting in larger pores, the real values of the entry pressure can be expected to be smaller than the estimated ones.
- The residual saturations of both water and air are set to zero, which means that the value of the saturated water content was equal to the porosity.
- The relative permeability functions for water and air in each type of sand are assumed to follow the BCB model, Eqs. 8-9.
- The intrinsic permeability of each medium corresponds to the value obtained in the independent experiments (again, due to larger pores, one can expect the real values to be higher).

The parameters based on the above assumptions are listed in Table 1 and denoted as BCB0 – first estimation. The results of the simulations with 2PH model for BCB0 parameters are shown in Figs. 4 and 5. Comparing saturation distribution in Fig. 4 with the one observed in the experiments one can note a qualitative agreement regarding the effect of structural trapping, *i.e.*, the upper parts of inclusions remain at low water saturation. However, in the

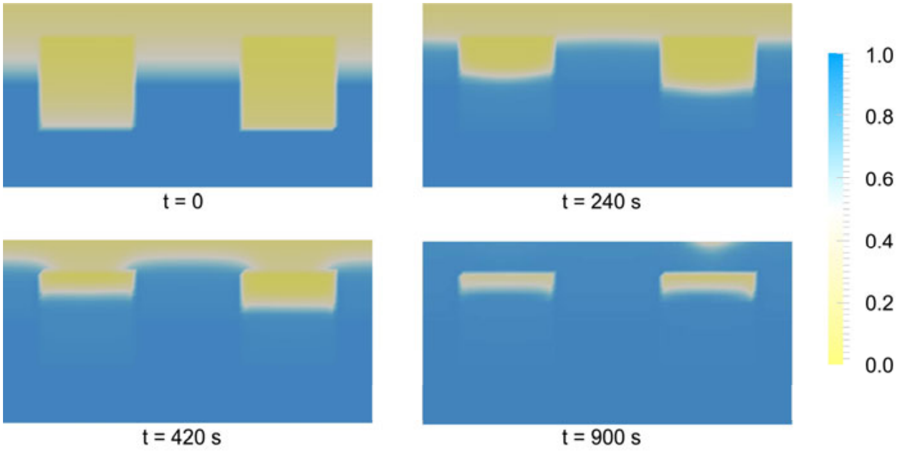


Fig. 4. Distribution of water saturation in the flume at four stages of the experiment, simulations with 2PH model with BCB0 set of parameters.

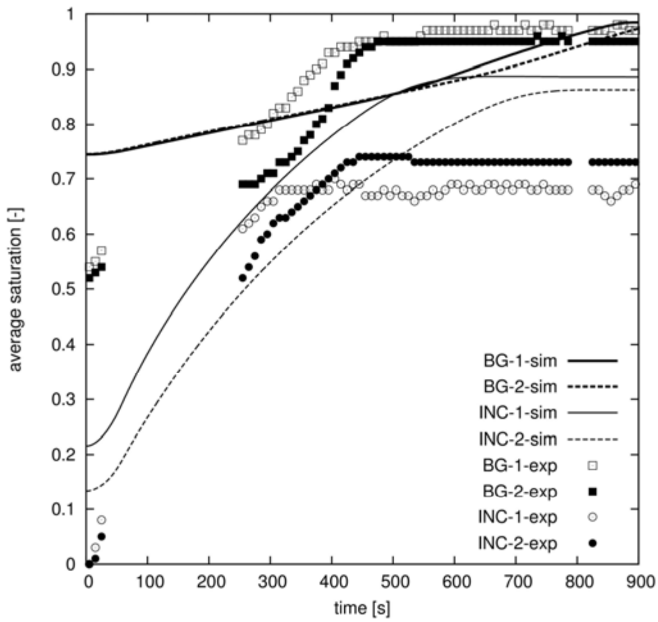


Fig. 5. Evolution of water saturation in four regions of the flume, simulations with 2PH model with BCB0 set of parameters.

numerical simulations steady state is not reached even for $t = 900$s, as there is still some air flowing out of the background material above the coarse inclusion. Comparison of the evolution of average saturation in each material

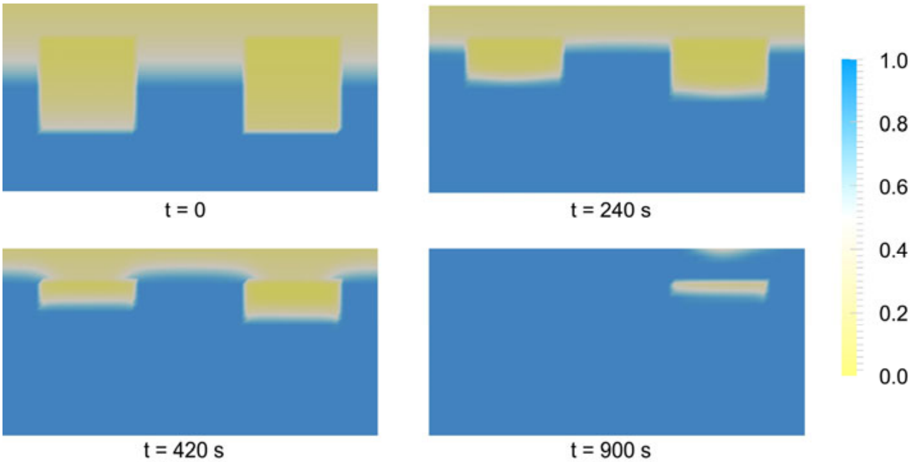


Fig. 6. Distribution of water saturation in the flume at four stages of the experiment, simulations with RE model with BCB0 set of parameters.

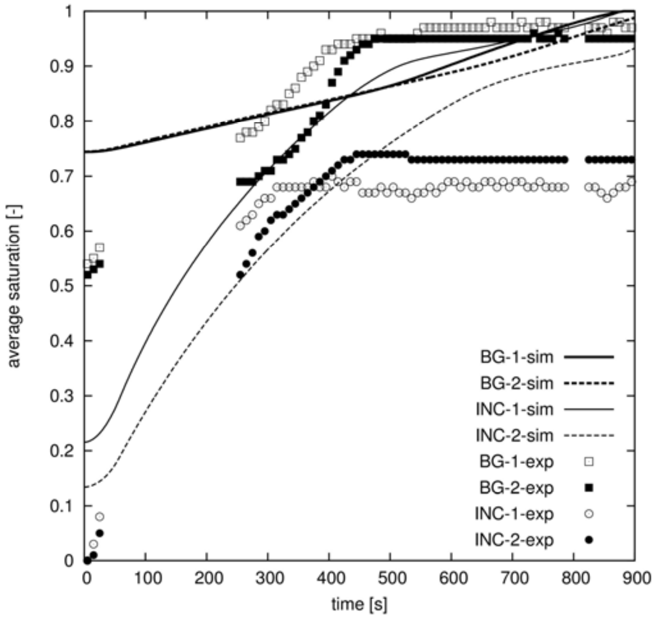


Fig. 7. Evolution of average water saturation in four regions of the flume, simulations with RE model with BCB0 set of parameters.

region (Fig. 5) also shows that the numerical simulations predict changes in the background material saturations long after the development of steady state in the experiments. On the other hand, the saturation in the inclusions becomes stable before the end of the simulation, but later than in the experiments. The amount of air trapped in inclusions is smaller according to the simulations, as compared to observations (*ca*. 12% of the pore volume in the simulations *versus ca*. 30% in the experiment). Also, BCB0 parameters lead to higher initial saturations in all four regions (by about 15 to 20% of pore volume). In the fine sand, the simulations predict much slower increase in saturation, compared to the experiment in the initial stage of infiltration. In the medium and coarse sand, the rate of saturation change is similar according to the simulations and experiment.

For the same parameter set BCB0 we also carried out a simulation using RE and the results are shown in Figs. 6 and 7. For 900 s the medium sand inclusion is fully saturated, while the coarse sand inclusion is still unsaturated in its upper part, but it is clear that the saturation in the coarse inclusion steadily increases (Fig. 7). For a longer time of about 1200 s both inclusions become fully saturated (results not shown here). RE does not account for the fact that water saturation can be increased only if there is a possibility for pore air to flow out. Consequently, saturation increases everywhere, even if the cells are fully surrounded with water. However, for the initial phase of infiltration, before the structural air trapping effects occur (at time of about 400 s), the results obtained with 2PH and RE models are very close to each other.

The major findings from these preliminary simulations were as follows. First, it was clear that the parameters from different experiment do not provide accurate results. This could, however, be expected because of different sand porosity. Second, the Richards equation could not qualitatively correctly capture the trapping effect, while the two phase model was able to do so.

5. PARAMETER FITTING

In order to improve the fit between the two-phase model and the experimental results we adjusted the hydraulic parameters by a trial and error procedure, based on the following assumptions:

- due to larger porosity in each material the permeability should be increased, while the air entry pressure should be decreased in each material;
- parameter λ can be increased, since all sands have rather uniform, well sorted grains, which implies a sharp decrease in saturation just above the air entry value;
- the residual air saturation in all types of sand should be non-zero, *i.e.*, the saturated water content θ_{sw} should be smaller than the porosity ϕ; the residual air saturation should be larger in the inclusions than in the background material.

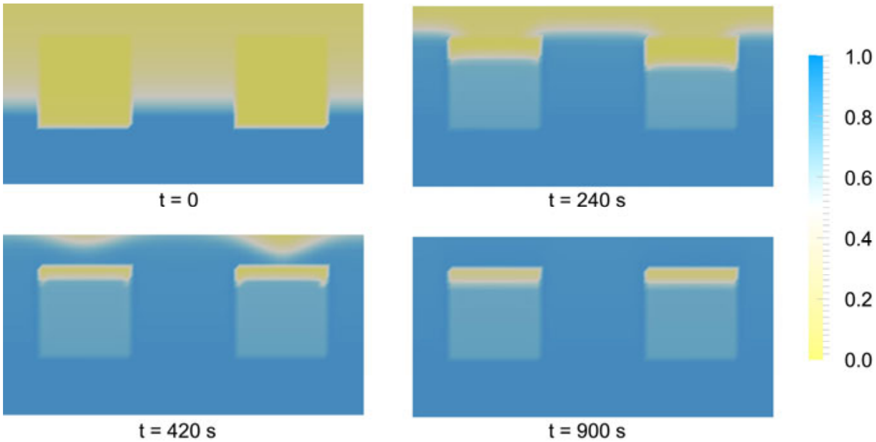


Fig. 8. Distribution of water saturation in the flume at four stages of the experiment, simulations with 2PH model with best-fitted hydraulic parameters BCB1, including pore-scale trapping.

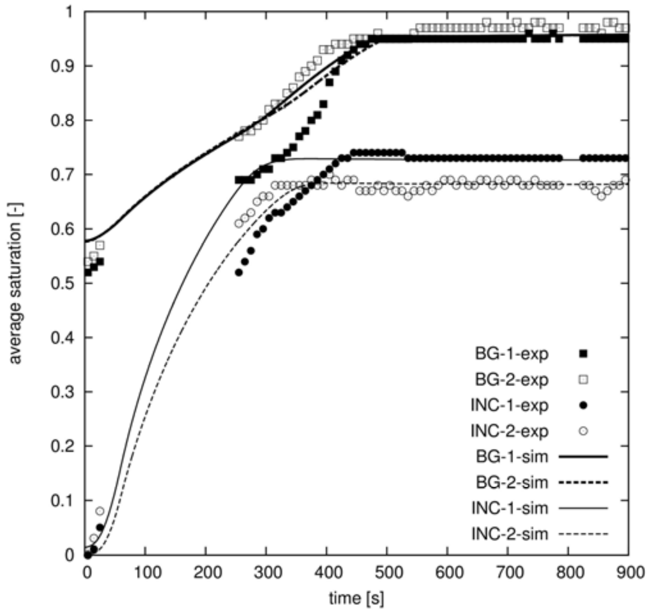


Fig. 9. Evolution of average water saturation in four regions of the flume, simulations with 2PH model with fitted hydraulic parameters BCB1, including pore-scale trapping.

By performing a large number of numerical tests we arrived at the set of parameters shown in Table 1 as BCB1 – best fit). While these parameters are clearly different from the initial estimation, the discrepancy seems accepta-

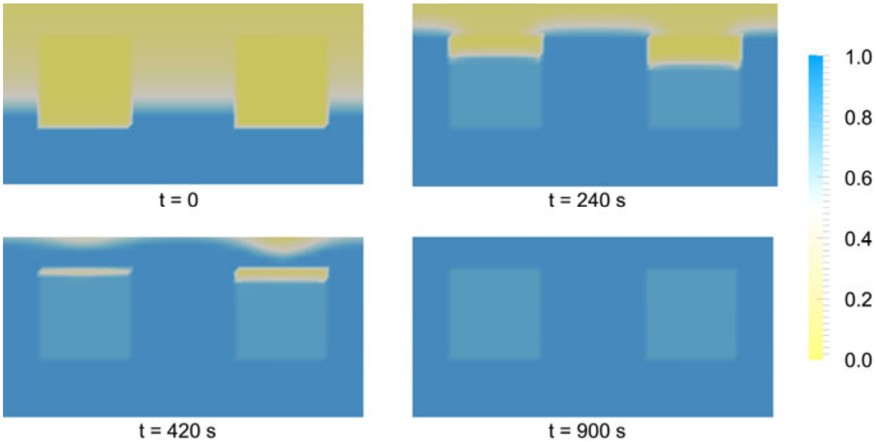


Fig. 10. Distribution of water saturation in the flume at four stages of the experiment, simulations with RE model with best-fitted hydraulic parameters BCB1, including pore-scale trapping.

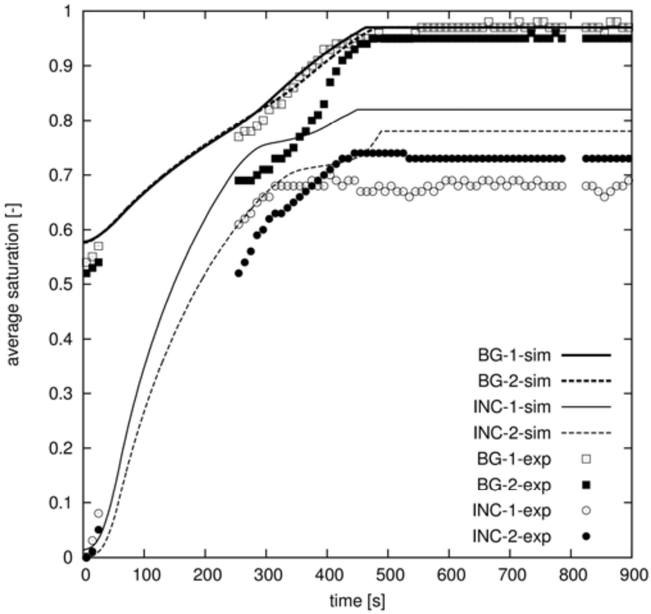


Fig. 11. Evolution of average water saturation in four regions of the flume, simulations with RE model with best-fitted hydraulic parameters BCB1, including pore-scale trapping.

ble in view of the assumptions listed above. The intrinsic permeability increased about 4 times for the fine sand and about 2 times for the other sands, while the entry pressure decreased by about 30% for each of the materials.

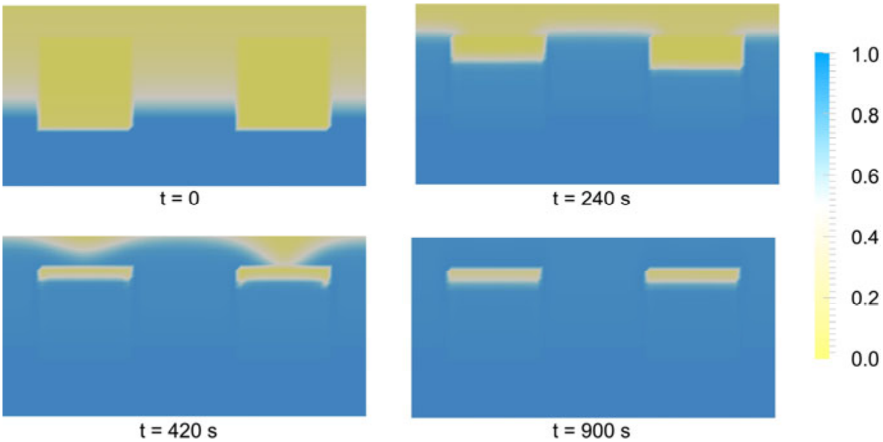


Fig. 12. Distribution of water saturation in the flume at four stages of the experiment, simulations with 2PH model with best-fitted hydraulic parameters, assuming no pore-scale air trapping.

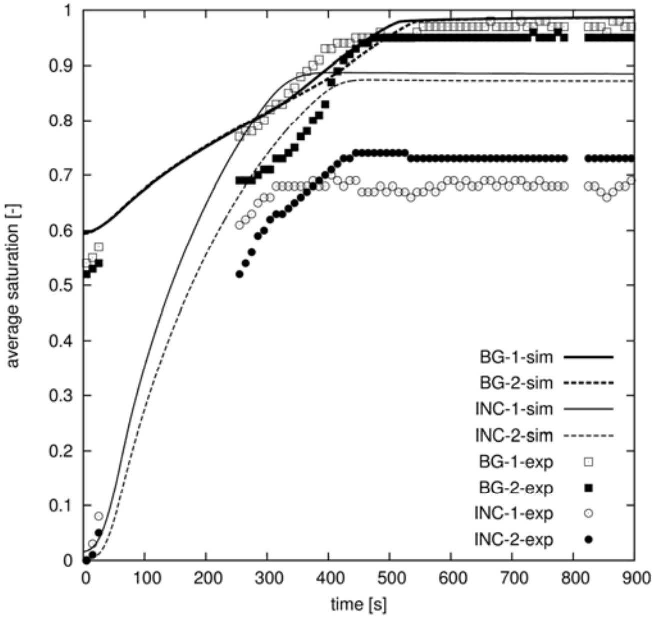


Fig. 13. Evolution of average water saturation in four regions of the flume, simulations with 2PH model with best-fitted hydraulic parameters, assuming no pore-scale trapping.

The fitted residual air content $\theta_{ra} = \phi - \theta_{sw}$ ranges from 0.015 in the fine material to 0.097 in the coarse material.

The results of the simulations performed with fitted BCB1 parameters are present in Figs. 8 and 9. They are in a good agreement with the results obtained from experiments. The initial saturation, saturation change rate, time to steady state and steady state saturation are generally well represented. Some discrepancies can be observed for the medium sand inclusion (INC1) and the surrounding fine sand (BG1). In the experiments, BG1 region saturation was significantly lower than BG2 saturation for time between 200 and 300 s. However, we were not able to represent this effect in any of the tested combinations of parameters. It could be caused by local heterogeneities within the background sand or by non-uniform water supply from the bottom.

The same set of parameters was used in a simulation with RE. The results are presented in Figs. 10 and 11. In this case there is no structural trapping effect and only pore scale air trapping occurs. At the final steady state, each inclusion has a uniform saturation, corresponding to its θ_{sw} value, which does not agree with the non-uniform distribution of water saturation observed in the experiments. On the other hand, the evolution of average saturation is in a qualitatively good agreement with the experiment, but the final values of water saturation in the inclusions are higher than observed. A better agreement of the saturation curves can be obtained for RE if the residual air saturation for coarse and medium sand is increased (results not shown here). However, it does not affect the final distribution of saturation in inclusions, which is still uniform, in contrast to the experimental results.

In order to further investigate the interplay between structural and pore scale air trapping we performed an additional simulation using 2PH model with a modified set of parameters BCB1 with residual air saturations set to 0 for each material. The results are shown in Figs. 12 and 13. This set of parameters produced a better fit than BCB0, but worse than BCB1. It can be seen that the effect of structural trapping is well represented, with unsaturated zones in the upper parts of inclusions, but the total amount of trapped air is smaller than in the experiments, due to the neglecting of pore scale air trapping. The structural trapping accounts for about 10% of pore volume in inclusions.

6. CONCLUSIONS

The experiment described in this paper provided additional evidence for the importance of local-scale heterogeneities for water and air flow in porous materials. During infiltration a significant amount of air was trapped in the coarse textures inclusions, both in the form of isolated bubbles in pores and larger unsaturated regions in the upper part of each inclusions. These results are in agreement with earlier experiments (Dunn and Silliman 2003, Dunn 2005) and with theoretical and numerical analyses (Szymkiewicz *et al.* 2012, 2014). The experimental results were reproduced with numerical simulations based on a two phase flow model with fitted parameters. The values of parameters obtained by the fitting procedure were in the range expected for the material used in the experiments, although they were different from the values obtained from independent measurements. The discrepancy can be explained by different porosity of sand and flow conditions (imbibition *versus* drainage). A satisfactory agreement with experimental results could be achieved only if both structural air trapping and pore scale air trapping were included in the model. In contrast, the results obtained with the Richards equation did not match the experiment. In RE only the pore scale trapping effect can be included, by reducing the apparent saturated water content with regard to porosity. However, the resulting final distribution of saturation in inclusions was uniform and the presence of unsaturated regions in the upper part of inclusions could not be reproduced, because RE cannot describe air entry barriers well. These macroscopic trapping effects could be mimicked by an apparent residual air saturation in the RE. However, this modeling approach would be quite unsatisfactory, as the parameters are specific to the initial and boundary conditions in the scenario considered and would need to be changed for different conditions. The 2PH model is able to predict the macroscopic trapping with unique model parameters. These finding seem to be important for modeling of flow and mass transport in the capillary fringe and water table fluctuation zone.

References

Aubertin, M., E. Cifuentes, S.A. Apithy, B. Bussière, J. Molson, and R.P. Chapuis (2009), Analyses of water diversion along inclined covers with capillary barrier effects, *Can. Geotech. J.* **46**, 10, 1146-1164, DOI: 10.1139/T09-050.

Berkowitz, B., S.E. Silliman, and A.M. Dunn (2004), Impact of the capillary fringe on local flow, chemical migration and microbiology, *Vadose Zone J.* **3**, 2, 534-548.

Bogacz, P., J. Kaczmarek, and D. Leśniewska (2006), Influence of air entrapment on flood embankment failure mechanics – model tests, *Technol. Sci.* **11**, 188-201.

Brooks, R.H., and A.T. Corey (1964), Hydraulic properties of porous media, Technical Report, Hydrology Paper 3, Colorado State University, Fort Collins, Colorado, USA.

Burdine, N.T. (1953), Relative permeability calculations from pore size distribution data, *J. Petrol. Technol.* **5**, 03, 71-78.

de Neef, M.J., and J. Molenaar (1997), Analysis of DNAPL infiltration in a medium with a low permeable lens, *Comput. Geosci.* **1**, 2, 191-214, DOI: 10.1023/A:1011569329088.

Delfs, J.O., W. Wang, T. Kalbacher, A.K. Singh, and O. Kolditz (2013), A coupled surface/subsurface flow model accounting for air entrapment and air pressure counterflow, *Environ. Earth Sci.* **69**, 2, 395-414, DOI: 10.1007/s12665-013-2420-1.

Dunn, A.M. (2005), Air and LNAPL entrapment in the partially saturated fringe: Laboratory and numerical investigations, Ph.D. Thesis, University of Notre Dame, Indiana, USA.

Dunn, A.M., and S.E. Silliman (2003), Air and water entrapment in the vicinity of the water table, *Ground Water* **41**, 729-734.

Forsyth, P.A. (1988), Comparison of the single-phase and two-phase numerical model formulation for saturated-unsaturated groundwater flow, *Comput. Meth. Appl. Mech. Eng.* **69**, 2, 243-259, DOI: 10.1016/0045-7825(88)90190-9.

Haberer, C.M., M. Rolle, O.A. Cirpka, and P. Grathwohl (2015), Impact of heterogeneity on oxygen transfer in a fluctuating capillary fringe, *Ground Water* **53**, 1, 57-70, DOI: 10.1111/gwat.12149.

Hammecker, C., A.C.D. Antonino, J.L. Maeght, and P. Boivin (2003), Experimental and numerical study of water flow in soil under irrigation in northern senegal: evidence of air entrapment, *Europ. J. Soil Sci.* **54**, 3, 491-503, DOI: 10.1046/j.1365-2389.2003.00482.x.

Helmig, R. (1997), *Multiphase Flow and Transport Processes in the Subsurface: A Contribution to the Modeling of Hydrosystems*, Springer.

Kuang, X., J.J. Jiao, L. Wan, X. Wang, and D. Mao (2011), Air and water flows in a vertical sand column, *Water Resour. Res.* **47** 4, W04506, DOI: 10.1029/2009WR009030.

Lenhard, R.J., J.C. Parker, and S. Mishra (1989), On the correspondence between Brooks-Corey and van Genuchten models, *J. Irrig. Drain. Eng. ASCE* **115**, 4, 744-751, DOI: 10.1061/(ASCE)0733-9437(1989)115:4(744).

Leśniewska, D., H. Zaradny, P. Bogacz, and J. Kaczmarek (2008), Study of flood embankment behaviour induced by air entrapment. **In:** P. Samuels, S. Huntington, W. Allsop, and J. Harrop (eds.), *Flood risk Management: Research and Practice*, Taylor & Francis, London, 655-665.

Likos, W.J., N. Lu, and J.W. Godt (2014), Hysteresis and uncertainty in soil water-retention curve parameters, *J. Geotech. Geoenviron. Eng.* **140**, 4, 04013050, DOI: 10.1061/(ASCE)GT.1943-5606.0001071.

Luckner, L., M.Th. van Genuchten, and D.R. Nielsen (1989), A consistent set of parametric models for the two-phase flow of immiscible fluids in the subsurface, *Water Resour. Res.* **25**, 10, 2113-2124, DOI: 10.1029/WR025i010p02187.

Marinas, M., J.W. Roy, and J.E. Smith (2013), Changes in entrapped gas content and hydraulic conductivity with pressure, *Ground Water* **51**, 1, 41-50, DOI: 10.1111/j.1745-6584.2012.00915.x.

McLeod, H.C., J.W. Roy, and J.E. Smith (2015), Patterns of entrapped air dissolution in a two-dimensional pilot-scale synthetic aquifer, *Ground Water* **53**, 2, 271-281, DOI: 10.1111/gwat.12203.

Mikelic, A., C.J. van Duijn, and I.S. Pop (2002), Effective equations for two-phase flow with trapping on the micro scale, *SIAM J. Appl. Math.* **62**, 5, 1531-1568, DOI: 10.1137/S0036139901385564.

Mualem, Y. (1976), A new model for predicting the hydraulic conductivity of unsaturated porous media, *Water Resour. Res.* **12**, 3, 513-522, DOI: 10.1029/WR012i003p00513.

Oldenburg, C.M., and K. Pruess (1993), On numerical modeling of capillary barriers, *Water Resour. Res.* **29**, 4, 1045-1056, DOI: 10.1029/92WR02875.

Prédélus, D., A.P. Coutinho, L. Lassabatere, B. Bien Le, T. Winiarski, and R. Angulo-Jaramillo (2015), Combined effect of capillary barrier and layered slope on water, solute and nanoparticle transfer in an unsaturated soil at lysimeter scale, *J. Contam. Hydrol.* **181**, 69-81, DOI: 10.1016/j.jconhyd.2015.06.008.

Richards, L.A. (1931), Capillary conduction of liquids through porous medium, *J. Appl. Physics.* **1**, 318-333, DOI: 10.1063/1.1745010.

Saadatpoor, E., S.L. Bryant, and K. Sepehrnoori (2009), Effect of capillary heterogeneity on buoyant plumes: a new local mechanism, *Energy Procedia* **1**, 1, 3299-3306, DOI: 10.1016/j.egypro.2009.02.116.

Schweizer, B. (2008), Homogenization of degenerate two-phase flow equations with oil-trapping, *SIAM J. Math. Anal.* **39**, 1740-1763, DOI: 10.1137/060675472.

Silliman, S.E., B. Berkowitz, J. Simunek, and M.Th. Van Genuchten (2002), Fluid flow and solute migration within the capillary fringe, *Ground Water* **40**, 1, 76-84, DOI: 10.1111/j.1745-6584.2002.tb02493.x.

Szymańska, P. (2012), Flow in unsaturated porous media: Numerical and experimental evaluation of the two-phase model and the richards equation, M.Sc. Thesis, Gdańsk University of Technology, Gdańsk, Poland.

Szymkiewicz, A. (2013), *Modelling Water Flow in Unsaturated Porous Media: Accounting for Nonlinear Permeability and Material Heterogeneity*, GeoPlanet: Earth and Planetary Sciences, Springer.

Szymkiewicz, A., R. Helmig, and H. Kuhnke (2011), Two-phase flow in heterogeneous porous media with non-wetting phase trapping, *Transport Porous Med.* **86**, 1, 27-47, DOI: 10.1007/s11242-010-9604.

Szymkiewicz, A., R. Helmig, and I. Neuweiler (2012), Upscaling unsaturated flow in binary porous media with air entry pressure effects, *Water Resour. Res.* **48**, 4, W04522, DOI: 10.1029/2011WR010893.

Szymkiewicz, A., I. Neuweiler, and R. Helmig (2014), Influence of heterogeneous air entry pressure on large scale unsaturated flow in porous media, *Acta Geophys*. **62**, 5, 1179-1191, DOI: 10.2478/s11600-014-0224-7.

Tegnander, C. (2001), Models for groundwater flow: A numerical comparison between Richards model and fractional flow model, *Transport Porous Med*. **43**, 2, 213-224, DOI: 10.1023/A:1010749708294.

Touma, J., and M. Vauclin (1986), Experimental and numerical analysis of two-phase infiltration in a partially saturated soil, *Transport Porous Med*, **1**, 1, 27-55, DOI: 10.1007/BF01036524.

Touma, J., G. Vachaud, and J.-Y. Parlange (1984), Air and water flow in a sealed, ponded vertical soil column: experiment and model, *Soil Science* **137**, 3, 181-187.

Vachaud, G., M. Vauclin, D. Khanji, and M. Wakil (1973), Effects of air pressure on water flow in an unsaturated stratified vertical column of sand, *Water Resour. Res*. **9**, 1, 160-173, DOI: 10.1029/WR009i001p00160.

van Duijn, C.J., H. Eichel, R. Helmig, and I.S. Pop (2007), Effective equations for two-phase flow in porous media: the effect of trapping at the micro-scale, *Transport Porous Med*. **69**, 3, 411-428, DOI: 10.1007/s11242-006-9089-9.

van Genuchten, M.Th. (1980), A closed form equation for predicting the hydraulic conductivity of unsaturated soils, *Soil Sci. Soc. Am. J.* **44**, 5, 892-898, DOI: 10.2136/sssaj1980.03615995004400050002x.

Vasin, M., P. Lehmann, A. Kaestner, R. Hassanein, W. Nowak, R. Helmig, and I. Neuweiler (2008), Drainage in heterogeneous sand columns with different geometric structures, *Adv. Water Resour*. **31**, 9, 1205-1220, DOI: 10.1016/j.advwatres.2008.01.004.

Webb, S.W. (1998), Using TOUGH2 to model capillary barriers. **In:** *Proc. TOUGH Workshop 1998, 4-6 May 1998, Lawrence Berkeley National Laboratory, Berkeley, California, USA*.

Yakirevich, A., T.J. Gish, J. Simunek, M.Th. Van Genuchten, Y.A. Pachepsky, T.J. Nicholson, and R.E. Cady (2010), Potential impact of a seepage face on solute transport to a pumping well, *Vadose Zone J.* **9**, 3, 686-696, DOI: 10.2136/vzj2009.0054.

Received 23 February 2016
Accepted 11 May 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2510-2529
DOI: 10.1515/acgeo-2016-0099

Oil Formation Volume Factor Determination Through a Fused Intelligence

Amin GHOLAMI

Reservoir Engineering Division, Iranian Offshore Oil Company, Tehran, Iran
e-mails: amingholamiput@yahoo.com, AGHolami@iooc.co.ir

Abstract

Volume change of oil between reservoir condition and standard surface condition is called oil formation volume factor (FVF), which is very time, cost and labor intensive to determine. This study proposes an accurate, rapid and cost-effective approach for determining FVF from reservoir temperature, dissolved gas oil ratio, and specific gravity of both oil and dissolved gas. Firstly, structural risk minimization (SRM) principle of support vector regression (SVR) was employed to construct a robust model for estimating FVF from the aforementioned inputs. Subsequently, an alternating conditional expectation (ACE) was used for approximating optimal transformations of input/output data to a higher correlated data and consequently developing a sophisticated model between transformed data. Eventually, a committee machine with SVR and ACE was constructed through the use of hybrid genetic algorithm-pattern search (GA-PS). Committee machine integrates ACE and SVR models in an optimal linear combination such that makes benefit of both methods. A group of 342 data points was used for model development and a group of 219 data points was used for blind testing the constructed model. Results indicated that the committee machine performed better than individual models.

Key words: PVT, oil formation volume factor (FVF), alternating conditional expectation (ACE), support vector regression (SVR).

Ownership: Institute of Geophysics, Polish Academy of Sciences;

1. INTRODUCTION

Oil formation volume factor (FVF) is defined as the ratio of the volume of oil (plus the gas in solution) at the prevailing reservoir temperature and pressure to the volume of oil at standard conditions (Ahmed 2000). FVF value has significance in calculating various parameters such as the depletion rate, oil in place, predicting the future of the reservoir, optimizing the rate of production, designing of production operation and facilities (Bagheripour *et al.* 2013). Rigorous depiction of FVF through differential vaporization test on bottom-hole or recombined surface samples is very time, cost and labor intensive (Zargar *et al.* 2014). Furthermore, sampling for experiments is limited to early producing life of reservoir which imposes a restraint to laboratory measurements (Dake 1988). Early attempts of researchers for presenting a practical, cheap and accurate way of determining FVF from available PVT data led to several empirical correlations (Katz 1942, Knopp and Ramsey 1960, Vazquez and Beggs 1970, Glaso 1980, Al-Marhoun 1988, Farshad *et al.* 1996, Petrosky and Farshad 1993, Omar and Todd 1993, Al-mehaideb 1997, Al-Shammasi 1999, Dindoruk and Christman 2001, El-Banbi *et al.* 2006, Hemmati and Kharrat 2007, Elmabrouk *et al.* 2010). A striving competition between intelligent systems and empirical correlations *versus* exactness and generalization has been done to show superiority of intelligent systems (Asoodeh and Kazemi 2013, Kazemi *et al.* 2013, Asoodeh and Bagheripour 2012a, 2013a; Bagheripour and Asoodeh 2013, 2014; Gholami *et al.* 2014a, b; Afshar *et al.* 2014). Hitherto, some scientists utilized intelligent systems for formulating oil FVF to available PVT data (Gharbi and Elsharkawy 1996, Elsharkawy 1998, Elsharkawy and Gharbi 2000, Al-Marhoun and Osman 2002, Dutta and Gupta 2010). The quest for higher accuracy forced researchers not to satisfy themselves with individual intelligent systems but to develop integrated models such as committee machines for enhancing precision of final prediction (Asoodeh and Bagheripour 2012b, Asoodeh 2013, Asoodeh *et al.* 2014a, b; Gholami *et al.* 2014c, d; Bagheripour *et al.* 2014). In this study, two sophisticated models, including support vector regression (SVR) and alternating conditional expectation (ACE) were employed to construct a strong formulation between PVT data and oil FVF. Several researches showed high performance of ACE and superiority of SVR to traditional networks (Shokir 2007, Al-Anazi and Gates 2010, Rafiee-Taghanaki *et al.* 2013, Na'imi *et al.* 2014, Asoodeh and Bagheripour 2013b, Asoodeh *et al.* 2014b, Gholami *et al.* 2014c, Fattahi *et al.* 2014, Bagheripour *et al.* 2015). At next stage, results of SVR and ACE models were combined by means of genetic algorithm-pattern search technique in an optimal linear combination of committee machine. This strategy was successfully applied to open source oil samples. Results indicated that the committee machine significantly enhanced accuracy of final prediction compared with individual ACE and SVR models.

2. THEORY: COMMITTEE MACHINE WITH SVR AND ACE

Committee machine is a parallel framework, as shown in Fig. 1, which gathers outputs of different models and combines them in an optimal linear structure by means of hybrid genetic algorithm-pattern search (GA-PS) tool. The GA-PS assigns a weight factor showing involvement of each model in overall estimation of target. In this study, PVT data, including reservoir temperature, dissolved gas oil ratio, and specific gravity of both oil and dissolved gas are introduced to support vector regression (SVR) and alternating conditional expectation (ACE) models for estimating oil formation volume factor (FVF). Outputs of SVR and ACE models are then input in committee machine. The GA-PS subsequently extracts involvement weights of each model such that mean square error (MSE) of prediction reaches its global minimum. This process consequently enhances the accuracy of final prediction. A brief introduction of SVR, ACE, and GA-PS is brought in the following paragraphs.

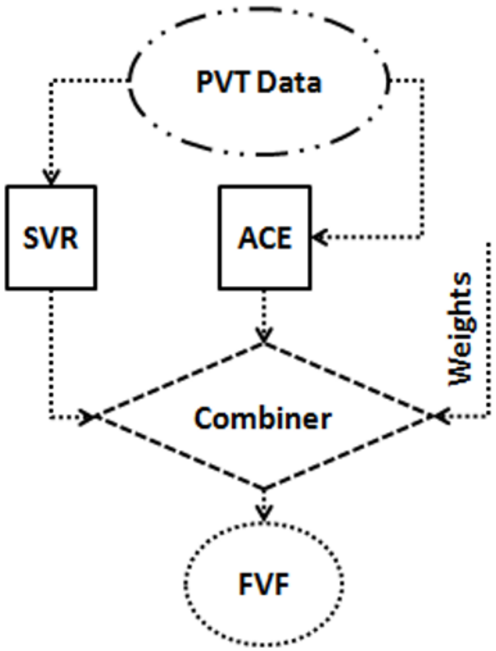


Fig. 1. Schematic diagram of committee machine used in this study.

2.1 Alternating conditional expectation

Alternating conditional expectation (ACE) is a nonparametric regression algorithm invented by Breiman and Friedman (1985). It is widely used for situations where underlying dependency between input/output data space is inexact or functional form between them is unidentified. This method transforms input/output data space to a higher correlated data space and develops the quantitative formulation between them in that space. It approximates optimal transformation of input/output data through minimizing error variance between transformed output data and sum of transformed input data in an alternative minimization process. When optimal transformations are achieved, a simple curve fitting can demonstrate optimal quantitative formulation between input/output data. In current part, a concise description about ACE formulation is given. More details about ACE are brought in a work by Breiman and Friedman (1985). General form of linear regression for formulation between p independent variables, $X_1, X_2, \ldots, X_p,$ and a response variable Y is

$$Y = \beta_0 + \sum_{i=1}^{p} \beta_i X_i + \varepsilon \ . \tag{1}$$

Here, $(\beta_i,\ i = 0 - p)$ are the regression coefficients which must be determined accurately, and ε is an error term. In ACE, transformations of Y and $X_1, X_2, \ldots, X_p$ are substituted in regression equation for making formulation between independent variables and response variable. Indeed, independent variables and response variable are firstly transformed into higher correlated data space and then, in this space, the underlying dependency between those is computed.

Based on aforementioned regression form, the non-parametric ACE algorithm is defined as the following equation:

$$\theta(Y) = \alpha + \sum_{i=1}^{p} \phi(X_i) + \varepsilon \ , \tag{2}$$

where $\theta(Y), \phi_1(X_1), \ldots, \phi_p(X_p)$ are the arbitrary measurable mean-zero functions of $Y, X_1, X_2, \ldots, X_p$, respectively. Hence, the main objective in ACE is to find the optimal transformation $\phi_i^*(X_i),\ i = 1, \ldots, p$ and $\theta^*(Y)$ which concluded the maximum correlation between transformed dependent variable and sum of transformed predicted variables. Breiman and Friedman (1985) suggested that the aforementioned objective is achieved through minimizing the value of the error variance (ε^2). The value of the error variance (ε^2) of a linear regression of the transformed dependent variable on the sum of transformed independent variables (under the constraint, $E[\theta^2(Y)] = 1$) is given by the following equation:

$$\varepsilon^2(\theta, \phi_i, \ldots, \phi_p) = E\left(\left[\theta(Y) - \sum_{i=1}^{p} \phi(X_i)\right]\right)^2 \Big/ E\theta^2(Y) \ . \tag{3}$$

By implementing minimization of the value of ε^2 with respect to $\theta(Y)$ and $\phi_k(X_k)(i = 1, 2, \ldots, k)$ through a series of single-function minimizations, the following equation for response variable and predictor variable is computed, respectively:

$$\theta(Y) = E\left[\sum_{i=1}^{p} \phi(X_i) \mid Y\right] \Big/ \left\| E\left[\sum_{i=1}^{p} \phi(X_i) \mid Y\right] \right\| , \tag{4}$$

$$\phi_{j,1}(X_j) = E\left[\theta(Y) - \sum_{i=1}^{p} \phi_i(X_i) \mid X_k\right] . \tag{5}$$

Performing of iteration process of minimizing ε^2 leads to determining the real-valued measurable zero-mean functions $\phi_i(X_i)$, $i = 1, \ldots, p$ and $\theta(Y)$, which is equivalent to optimal transformation $\phi_i^*(X_i)$, $i = 1, \ldots, p$ and $\theta^*(Y)$. In the transformed space, the response and predictor variables are related as following.

$$\theta^*(Y) = \sum_{i=1}^{p} \phi_i^*(X_i) + e^* \ , \tag{6}$$

where e^* is the error not captured by the use of the ACE transformations and is assumed to have a normal distribution with zero mean.

2.2 Support vector regression

Support vector regression, invented by Vapnik (1995), is a supervised model inspired by statistical learning theory. SVR utilizes structural risk minimization (SRM) in conjunction with empirical risk minimization (ERM) that fortifies it with the highest generalization owing to constructing a structural model that would be as smooth as possible. SVR nonlinearly maps input/output data to a higher-dimensional feature space by means of kernel functions such that a linear relationship between input/output data exists in feature space. Linear hyperplane in feature space produces a nonlinear regression hypersurface in original data space. Therefore, nonlinear relationship between input/output data is extracted. Here brief description of SVR method is given. For more study about SVR, refer to Al-Anazi and Gates (2010). The primary aim of SVR regression is to discover linear relation between n-dimensional input vectors $x \in R^n$ and output variables $y \in R$ as follows:

$$f(x) = w^T x + b \ , \tag{7}$$

where w and b are the slope and offset of the regression line, respectively. To find the relation between n-dimensional input vectors, the values of regression parameters (w and b) must be determined. For attaining the aforementioned objective, minimizing of the following equation is essential

$$R = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}\left|y_i - f\left(x_i\right)\right|_\varepsilon \quad . \tag{8}$$

The loss function used in this strategy is ε-insensitive. This loss function, introduced by Vapnik (1995), is expressed by the following equation:

$$\left|y_i - f\left(x_i\right)\right|_\varepsilon = \begin{cases} 0 & \text{if} \quad \left|y_i - f(x_i)\right| \le \varepsilon \\ \left|y_i - f(x_i)\right| - \varepsilon & \text{otherwise} \end{cases} \tag{9}$$

This problem can be reformulated in a dual space by

Maximize

$$L_p\left(\alpha_i, \alpha_i^*\right) = -\frac{1}{2}\sum_{i,j=1}^{l}\left(\alpha_i - \alpha_i^*\right)\left(\alpha_j - \alpha_j^*\right)x_i^T x_j - \varepsilon\sum_{i=1}^{l}\left(\alpha_i + \alpha_i^*\right) + \sum_{i=1}^{l}\left(\alpha_i - \alpha_i^*\right)y_i$$

subject to

$$\begin{cases} \sum_{i=1}^{l}\left(\alpha_i - \alpha_i^*\right) = 0 \\ 0 \le \alpha_i \le C, \quad i = 1, \ldots, l \\ 0 \le \alpha_i^* \le C, \quad i = 1, \ldots, l \end{cases} \tag{10}$$

After calculation of Lagrange multipliers, α_i and α_i^*, training data points from which those meeting the conditions $\alpha_i - \alpha_i^* \ne 0$ will be employed to construct the decision function. The total number of the points with prior criteria will be considered as the number of support vectors. Hence, the best linear hyper surface regression is given by:

$$f(x) = w_o{}^T x + b = \sum_{i=1}^{l} \left(\alpha_i - \alpha_i^*\right) x_i^T x + b \tag{11}$$

in which the desired weight vector of the regression hyper plane is given by:

$$w_o = \sum_{i=1}^{l} \left(\alpha_i - \alpha_i^*\right) x_i \ . \tag{12}$$

In the case of the nonlinear regression, learning problem is again formulated in the same way as in the linear case. The only difference between linear and nonlinear regression is the implanting of kernel function in regression function. Hence the nonlinear hyperplane regression function becomes:

$$f(x) = \sum_{i=1}^{l} \left(\alpha_i - \alpha_i^*\right) K\left(x_i, x\right) + b \ . \tag{13}$$

In above equation, $K(x_i, x)$ is kernel function which is defined as follows:

$$k\left(x_i, x_j\right) = \Phi^T\left(x_i\right)\Phi\left(x_j\right) \qquad i, j = 1, \ldots, l \ , \tag{14}$$

where, $\Phi(x_i)$ and $\Phi(x_j)$ are projection of the x_i and x_j in feature space, respectively.

2.3 Hybrid genetic algorithm-pattern search technique

Genetic algorithm (GA) is an optimization approach which starts with a random population of chromosome-like solutions and evolves to better solutions by applying genetic operations. Genetic algorithm discovers the global minimum of fitness function (function which its global minimum is desired). Therefore, a function meant to be solved should be rearranged such that the global minimum of the rearranged function and the desired point of original function are the same. Evaluation of each chromosome (solution) produces the corresponding fitness score which in turn is used for selection procedure and forming the succeeding population after applying genetic operations. This process continues until the desired chromosome is achieved. For better performance of genetic algorithm, a pattern search technique is integrated with GA. This means that, after each generation, all chromosomes are enhanced by means of pattern search technique. In the pattern search technique, the algorithm searches a set of points, called a mesh, around the current chromosome. The mesh is formed by adding the current chromosome to a scalar multiple of a set of vectors called a pattern. After assessment of all points according to fitness function, the best solution in the mesh is replaced by current chromosome. Before the run of GA-PS, the number of regulation parameters must be adjusted. These parameters include population type, population size, initial range, scaling function, selection function, elite preservation, crossover fraction, mutation function, crossover function, hybrid function, generations, stall generations, fitness tolerance, and time limit. Population type specifies the data type of the input to the fitness function. Population size determines the number of individuals which are in each generation. Initial range limits the range of the points in the initial population through setting the lower and upper bounds. Scaling function changes the raw fitness determined by virtue of the fitness function to values in a range of that is fit for the selection function. Selection function specifies how the genetic algorithm chooses parents for the next generation. Elite preservation specifies the number of individuals that are guaranteed to survive to the next generation. Crossover fraction specifies the fraction of the next generation. Mutation function specifies how the genetic algorithm makes small random changes in the individuals in the population to create mutation children. Crossover function specifies how the genetic algorithm combines two individuals, or parents, to form a crossover child for the next generation. Hybrid function is another minimization function that runs after the genetic algorithm terminates. Generations specify the maximum number of iterations for the genetic algorithm to perform. This algorithm stops if the weighted average change in the fitness function value over stall generations is less than function tolerance. The algorithm runs until the cumulative change in the fitness function value over stall generations is less than or equal to function tolerance. The algorithm stops if there is no improvement in the best fitness value for an interval of time in seconds specified by time limit. More details about GA-PS tool are available in MATLAB user's guide (MATLAB User's Guide 2011), Mohaghegh (2000), Asoodeh and Bagheripour (2012b, 2013c), Asoodeh *et al.* (2014c).

3. INPUT/OUTPUT DATA SPACE

Generalization of intelligence based model is mainly a function of range of dataset employed for its construction. Moreover, owing to infeasibility to incorporate *a priori* knowledge into this group of models, its performance depends on the reliability of data employed for its construction. Hence, gathering of data is an important step in model development using intelligence based model. Dataset which employed in current study for building predictive model for estimation of formation volume factor of crude oil from production data is borrowed from papers available in literature (Al-Marhoun 1988, Bello *et al.* 2008, Dokla and Osman 1990, Mahmood and Al-Marhoun 1996, Moghadam *et al.* 2011, Obomanu and Okpobiri 1987, Omar and Todd 1993). The dataset consists of production data (reservoir temperature, solution gas oil ratio, reservoir oil gravity (API), and dissolved gas relative density) and the corresponding value of FVF. Out of 561 data points, 342 data points are used for training of model and 219 data points are employed for evaluating the reliability of constructed model. Statistical description of employed dataset is given in Table 1. As seen in the table, crude oils studied herein have a wide range of reservoir and production conditions.

Table 1

Statistical description of employed dataset

Parameter	Minimum	Maximum	Average
Solution gas oil ratio (SCF/STB)	169.53	1608.26	701.43
Dissolved gas relative density [%]	0.91	1.71	1.19
Reservoir oil gravity [^{o}API]	19.30	43.58	27.92
Reservoir temperature [^{o}F]	90.00	260.00	181.37
Formation volume factor [bbl/STB]	1.15	2.01	1.42

4. RESULTS AND DISCUSSION

4.1 ACE model

At the first stage of this study, an alternating conditional expectation algorithm is employed to construct a model meant to estimate oil formation volume factor from available PVT data. ACE transforms input/output data such that error variance between transformed output (FVF) and sum of transformed input data (PVT data) is minimized. After a nonparametric transformation of each input/output dataset is evaluated, a functional form is approximated to each transformation by use of a simple curve fitting tool. Optimal transformation of PVT data (inputs) and FVF (output) is depicted in Fig. 2. In the next step, sum of transformed input data is evaluated. Once again, a simple curve fitting between output and sum of transformations produces functional form for estimating formation volume factor from PVT data. To evaluate performance of constructed ACE model, unseen test data were input to it and FVF was estimated. Figure 3 shows crossplot between actual FVF and ACE predicted FVF along with residual of predictions. This figure indicated that ACE was successful in estimation of FVF.

4.2 SVR model

At the next stage of this study, an epsilon support vector regression algorithm was used for formulating available PVT data to FVF. Firstly, all available data were scaled in range of [–1 1] and subsequently all data were transformed to feature space using radial basis kernel function (RBF) owing

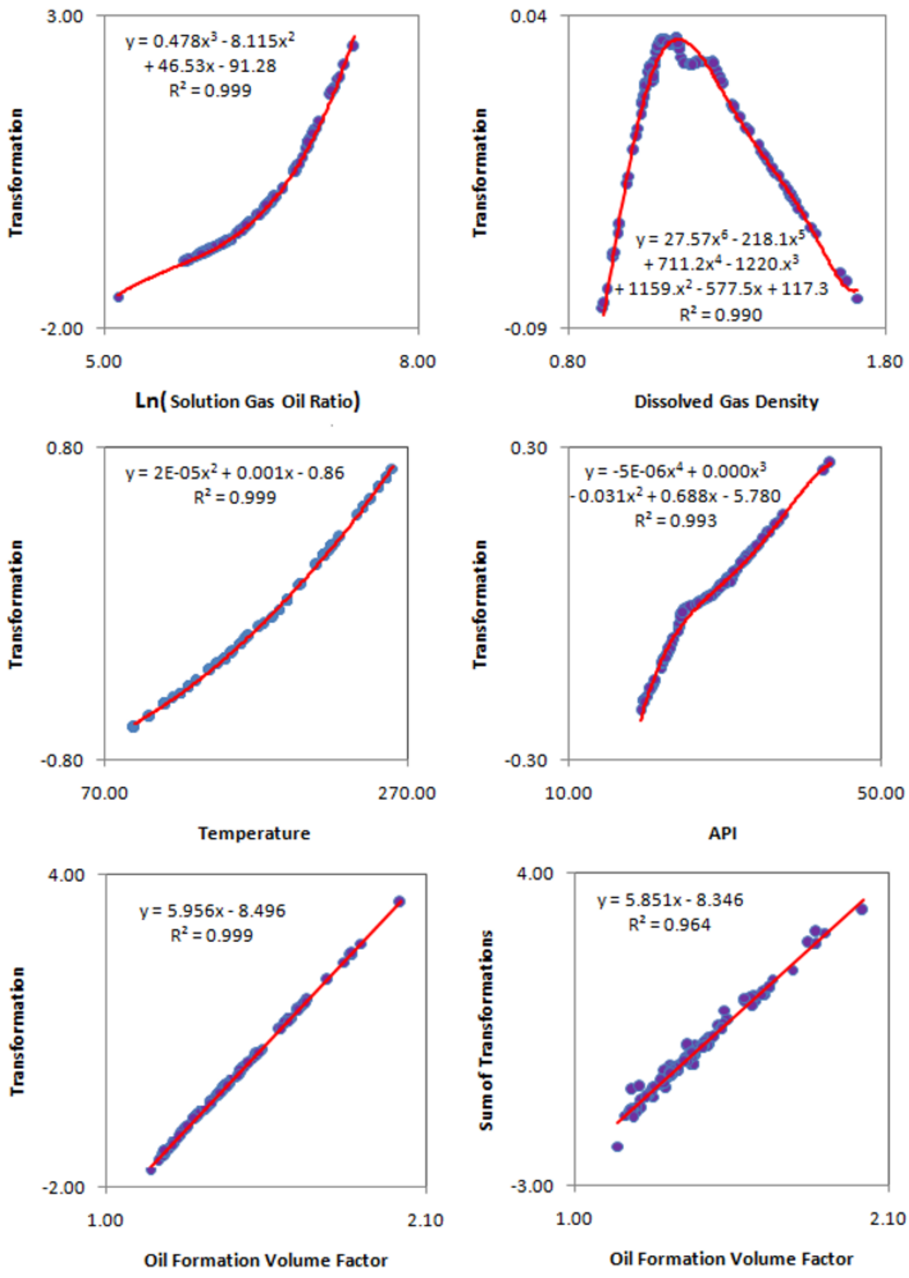


Fig. 2. Optimal transformation of input/output data and crossplot of sum of transformations vs. oil FVF.

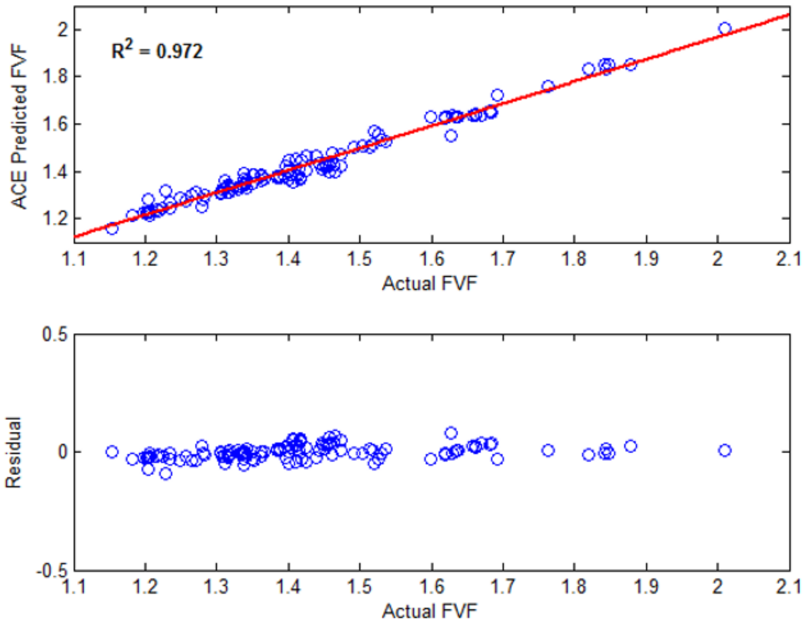


Fig. 3. Graph evaluating the performance of ACE model *versus* correlation coefficient and residuals of prediction.

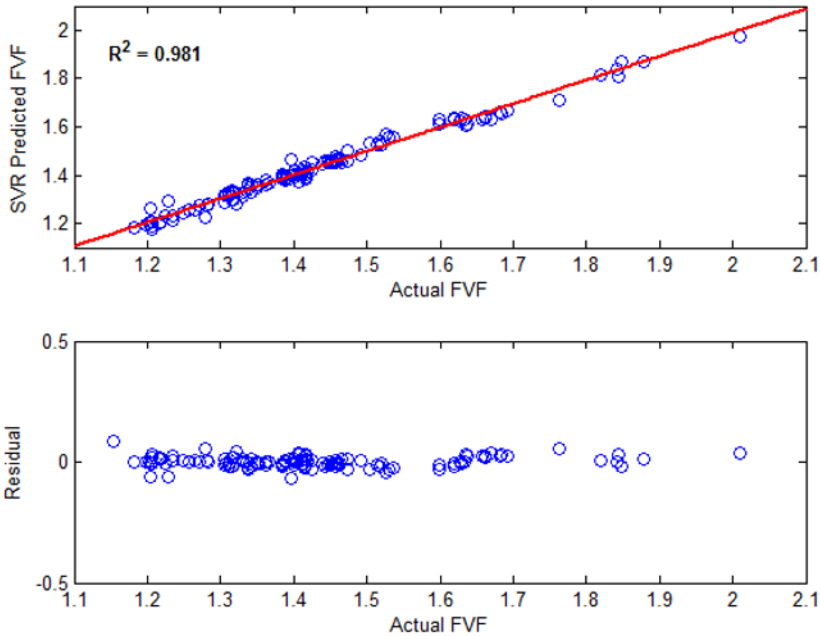


Fig. 4. Graph evaluating the performance of SVR model *versus* correlation coefficient and residuals of prediction.

Table 2

The parameter values corresponding to the optimum ε-SVR model

	C	Gamma	Epsilon
Search ranges	0.01-1000000	0.000001-20	0.0001-100
Optimum value for ε-SVR model	2456.650274	0.0567989	0.0023314

to fewer parameters to be tuned and low computational cost (Keerthi and Lin 2003). Parameters involved in SVR and kernel function (*i.e.*, epsilon, Gamma, and C) were determined through a thorough surveying using combination of grid search and pattern search techniques, as You *et al.* (2014) suggested. Optimum values of these parameters are shown in Table 2. In the present model, 342 Lagrange multiplier pairs were employed where 311 support vectors among them were used for model construction. After the SVR model was built, unseen test data were used for blind testing performance of SVR model. Figure 4 shows assessment of SVR model using concepts of correlation coefficient and residuals. This figure proves there is a satisfying match between predicted and actual FVF.

4.3 Committee machine with ACE and SVR

In the latter stage of present study, a committee machine with ACE and SVR models was constructed to combine their outputs in an optimal linear structure. Committee machine reaps the benefits of both ACE and SVR models through assigning a weight of contribution to each model such that accuracy of final prediction is enhanced. Therefore, estimated FVF from committee machine will be simply of the following form:

$$\mathrm{FVF_{CM}} = w_1 \times \mathrm{SVR} + w_2 \times \mathrm{ACE}\ . \tag{15}$$

To extract the optimal weight of contribution of each model (w_1 and w_2), MSE function of committee machine was introduced to hybrid genetic algorithm-pattern search technique. The GA-PS tool starts with a population of randomly generated pairs of probable solutions (pairs of w_1 and w_2) in a chromosome-like structure. By applying different genetic operators over each population, a new generation of enhanced chromosomes (solutions) is obtained. This process continues until the desired pair of (w_1 w_2) is achieved. This process is shown in Fig. 5. Regulations of genetic algorithm before running are shown in Table 3. Performance of committee machine with ACE and SVR is assessed using concepts of correlation coefficient, relative error and residual analysis of prediction (Fig. 6). High value of correlation coeffi-

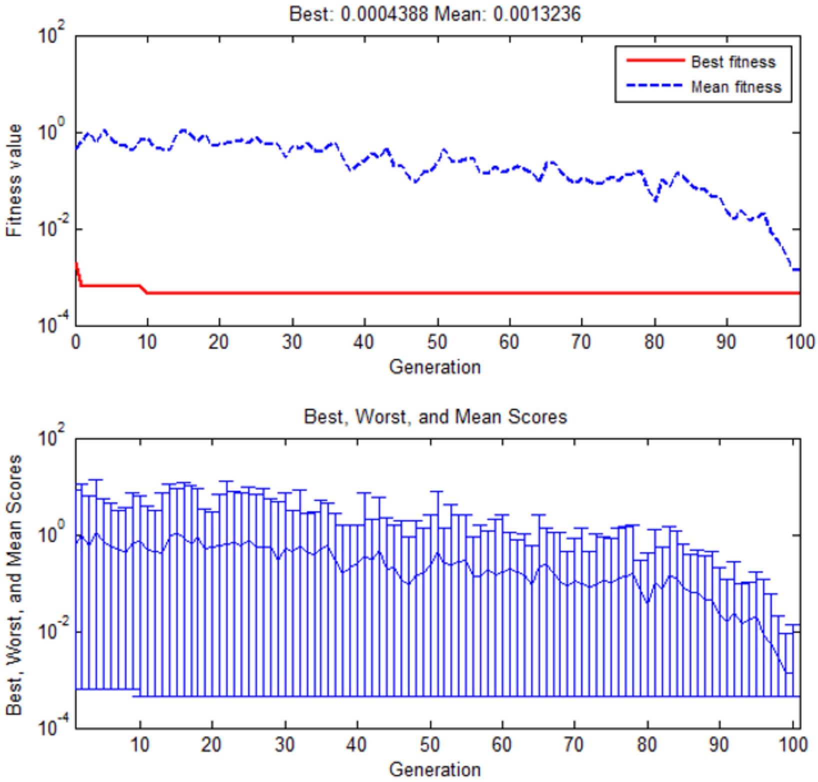


Fig. 5. Graph showing mean, best and worst fitness scores of FVF fitness function during 100 generations.

Table 3

Regulations done before the run of genetic algorithm

Parameter/setting	Type/value	Parameter/setting	Type/value
Population type	Double vector	Mutation function	Gaussian
Population size	20 chromosomes	Crossover function	Scattered
Initial range	[-1 1]	Hybrid function	Pattern search
Scaling function	Proportional	Generations	100
Selection function	Roulette	Stall generations	100
Elite preservation	2	Fitness tolerance	1.0 E -6
Crossover fraction	0.85	Time limit	Infinity

cient, low value of relative errors, low value of residuals, and concentration of residuals for most samples in close proximity of zero are evidences of supreme performance of committee machine.

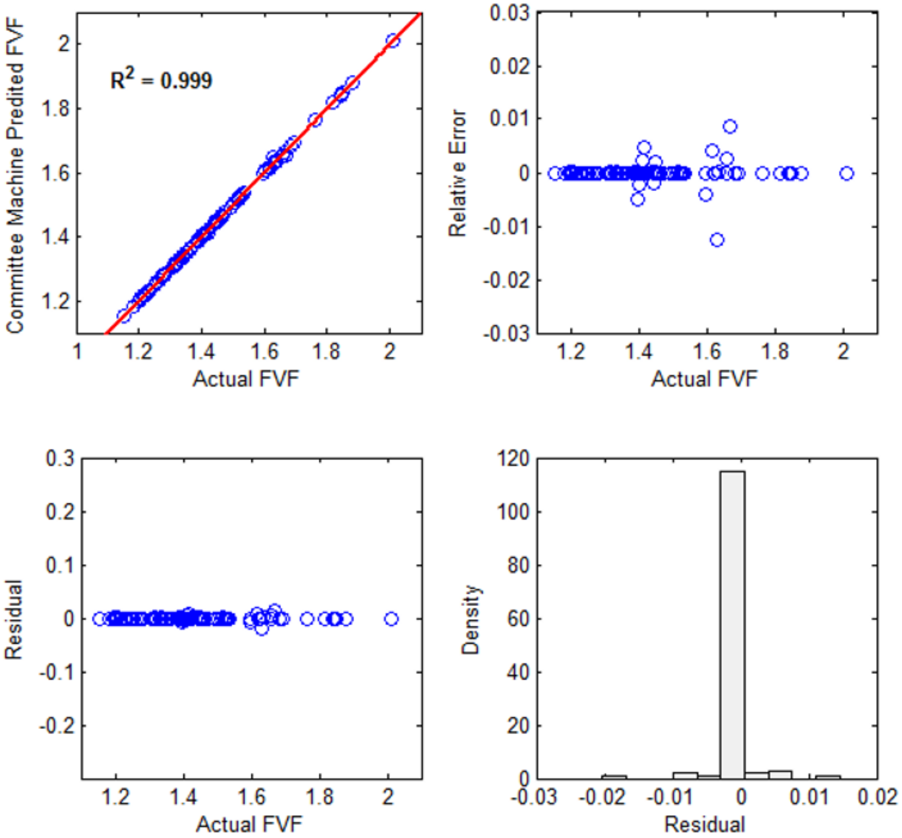


Fig. 6. Graph evaluating the performance of committee machine model using concepts of correlation coefficient, relative error, and residual analysis of prediction.

4.4 Comparison of models

Zargar *et al.* (2014) proposed a fuzzy logic model for estimating FVF from available PVT data. Performing an error distribution analysis, they concluded that the fuzzy model is an accurate model for FVF estimation. Bagheripour *et al.* (2013) made a comparison between traditional and stochastically optimized neural network for estimation of FVF. They showed that the use of genetic algorithm instead of back-propagation algorithm enhances the precision of overall FVF prediction. In this section, results of the mentioned works are compared with models presented in current study. Table 4 provides an opportunity to compare MSE and R-square factors of different models. Results show that committee machine surpasses other methods and provides more reliable results relative to Zargar *et al.* (2014), Bagheripour *et al.* (2013), and individual SVR and ACE models.

Table 4

Comparing different models for estimation of FVF *versus* *R*-square and MSE

Method	Results	*R*-square	MSE
Present study	Committee machine	0.999	0.0004389
	SVR	0.981	0.00051017
	ACE	0.972	0.00081161
Zargar *et al.*	Fuzzy Logic	0.992	0.0004767
Bagheripour *et al.*	Optimized neural network	0.982	0.0004822
	Neural network	0.931	0.0022669

5. CONCLUSIONS

This study proposed an accurate, cheap and rapid way for estimating oil formation volume factor from available PVT data. In situations where sampling is not applicable owing to the fact that the producing life of reservoir is too long or in situations where sampling are not desired on account of costs and time-consumption, the proposed strategy is an appealing alternative. ACE and SVR are featured to exactly extract underlying dependency between oil formation volume factor and available PVT data. Genetic algorithm is a sophisticated approach for taking part as combiner of committee machine. It is capable of finding optimal linear combination of SVR and ACE models for enhancing accuracy of final prediction. In situations where multiple options are available for solving a problem, committee machine is a great idea for enhancing accuracy of final prediction by little additional computation. Comparing current study with previous ones revealed superiority of committee machine using concepts of *R*-square and MSE.

Acknowledgments. A. Gholami would like to acknowledge the departments of research and technology of the National Iranian Oil Company and Iranian Offshore Oil Company for support throughout this research.

A. Gholami is grateful to Mr. M. Asoodeh and Mr. M. Vaezzadeh-Asadi for their kind help during the preparation of the manuscript.

References

Afshar, M., A. Gholami, and M. Asoodeh (2014), Genetic optimization of neural network and fuzzy logic for oil bubble point pressure modeling, *Korean. J. Chem. Eng.* **31**, 3, 496-502, DOI: 10.1007/s11814-013-0248-8.

Ahmed, T. (2000), *Reservoir Engineering Handbook*, 2th ed., Gulf Professional Publishing, Burlington, 863 pp.

Al-Anazi, A.F., and I.D. Gates (2010), Support vector regression for porosity prediction in a heterogeneous reservoir: A comparative study, *Comput. Geosci.* **36**, 12, 1494-1503, DOI: 10.1016/j.cageo.2010.03.022.

Al-Marhoun, M.A. (1988), PVT correlations for Middle East crude oils, *J. Petrol. Technol.* **40**, 5, 650-666, DOI: 10.2118/13718-PA.

Al-Marhoun, M.A., and E.A. Osman (2002), Using artificial neural networks to develop new PVT correlations for Saudi crude oils, SPE Paper 78592, DOI: 10.2118/78592-MS.

Almehaideb, R.A. (1997), Improved PVT correlations for UAE crude oils, SPE Paper 26644, DOI: 10.2118/37691-MS.

Al-Shammasi, A.A. (1999), Bubble Point pressure and oil formation volume factor correlations, SPE Paper 53185.

Asoodeh, M. (2013), Prediction of Poisson's ratio from conventional well log data: A committee machine with intelligent systems approach, *Energy Sources A* **35**, 10, 962-975, DOI: 10.1080/15567036.2011.557693.

Asoodeh, M., and P. Bagheripour (2012a), Estimation of bubble point pressure from PVT data using a power-law committee with intelligent systems, *J. Pet. Sci. Eng.* **90-91**, 1-11, DOI: 10.1016/j.petrol.2012.04.021.

Asoodeh, M., and P. Bagheripour (2012b), Prediction of compressional, shear, and stoneley wave velocities from conventional well log data using a committee machine with intelligent systems, *Rock. Mech. Rock. Eng.* **4**5, 1, 45-63, DOI: 10.1007/s00603-011-0181-2.

Asoodeh, M., and P. Bagheripour (2013a), Neuro-fuzzy reaping of shear wave velocity correlations derived by hybrid genetic algorithm-pattern search technique, *Open Geosci.* **5**, 2, 272-284, DOI: 10.2478/s13533-012-0129-4.

Asoodeh, M., and P. Bagheripour (2013b), Fuzzy classifier based support vector regression framework for Poisson's ratio determination, *J. Appl. Geophys*. **96**, 7-10, DOI: 10.1016/j.jappgeo.2013.06.006.

Asoodeh, M., and P. Bagheripour (2013c), Core porosity estimation through different training approaches for neural network: Back-propagation learning vs. genetic algorithm, *Int. J. Comput. Appl.* **63**, 5, 11-15, DOI: 10.5120/10461-5172.

Asoodeh, M., and K. Kazemi (2013), Estimation of bubble point pressure using a genetic integration of empirical formulas, *Energy Sources A* **35**, 12, 1102-1109, DOI: 10.1080/15567036.2011.574195.

Asoodeh, M., A. Gholami, and P. Bagheripour (2014a), Asphaltene precipitation of titration data modeling through committee machine with stochastically optimized fuzzy logic and optimized neural network, *Fluid. Phase Equilibr.* **364**, 67-74, DOI: 10.1016/j.fluid.2013.12.016.

Asoodeh, M., A. Gholami, and P. Bagheripour (2014b), Oil-CO_2 MMP determination in competition of Neural Network, support vector regression and committee machine, *J. Disper. Sci. Technol.* **35**, 4, 564-571, DOI: 10.1080/01932691.2013.803255.

Asoodeh, M., A. Gholami, and P. Bagheripour (2014c), Renovating scaling equation through hybrid genetic algorithm-pattern search tool for asphaltene precipitation modeling, *J. Disper. Sci. Technol.* **35**, 4, 607-611, DOI: 10.1080/01932691.2013.825209.

Bagheripour, P., and M. Asoodeh (2013), Fuzzy ruling between core porosity and conventional well logs: subtractive clustering vs. genetic algorithm-pattern search, *J. Appl. Geophys*. **99**, 35-41, DOI: 10.1016/j.jappgeo.2013.09.014.

Bagheripour, P., and M. Asoodeh (2014), Genetic implanted fuzzy model for water saturation determination, *J. Appl. Geophys*. **103**, 232-236, DOI: 10.1016/j.jappgeo.2014.02.002.

Bagheripour, P., M. Asoodeh, and A. Asoodeh (2013), Oil formation volume factor modeling: Traditional NN vs. stochastically optimized neural network, *Open Geosci*. **5**, 4, 508-513, DOI: 10.2478/s13533-012-0154-3.

Bagheripour, P., A. Gholami, and M. Asoodeh (2014), Support vector regression between PVT data and bubble point pressure, *J. Petrol. Explor. Prod. Technol*. **5**, 3, 227-231, DOI: 10.1007/s13202-014-0128-8.

Bagheripour, P., A. Gholami, M. Asoodeh, and M. Vaezzadeh-Asadi (2015), Support vector regression based determination of shear wave velocity, *J. Petrol. Sci. Eng*. **125**, 95-99, DOI: 10.1016/j.petrol.2014.11.025.

Bello, O., K. Reinicke, and P. Patil (2008), Comparison of the performance of empirical models used for the prediction of the PVT properties of crude oils of the Niger delta, *Petrol. Sci. Technol.* **26**, 5, 593-609, DOI: 10.1080/10916460701204685.

Breiman, L., and J.H. Friedman (1985), Estimating optimal transformations for multiple regression and correlation, *J. Am. Stat. Assoc*. **80**, 391, 580-598, DOI: 10.1080/01621459.1985.10478157.

Dake, L.P. (1988), *Fundamental of Reservoir Engineering*, 17th ed., Elsevier Science.

Dindoruk, B., and P.G. Christman (2001), PVT properties and viscosity correlations for Gulf of Mexico oils, SPE Paper 26644, DOI: 10.2118/71633-MS.

Dokla, M.E., and M.E Osman (1990), Correlation of PVT properties for U.A.E. crudes.

Dutta, S., and J.P. Gupta (2010), PVT correlations for Indian crude using artificial neural networks, *J. Petrol. Sci. Eng*. **72**, 1-2, 93-109, DOI: 10.1016/j.petrol.2010.03.007.

El-Banbi, A.H., K.A. Fattah, and M.H. Sayyouh (2006), New modified black-oil correlations for gas condensate and volatile oil fluids, SPE Paper 26644, DOI: 10.2118/102240-MS.

Elmabrouk, S., A. Zekri, and E. Shirif (2010), Prediction of bubble point pressure and bubble point oil formation volume factor in the absence of PVT analysis, SPE Paper 26644.

Elsharkawy, A.M. (1998), Modeling the properties of crude oil and gas system using RBF network, SPE Paper 26644, DOI: 10.2118/49961-MS.

Elsharkawy, A.M., and R.B.C. Gharbi (2000), Comparing classical and neural regression techniques in modeling crude oil viscosity, *Adv. Eng. Softw.* **32**, 3, 215-224, DOI: 10.1016/S0965-9978(00)00083-1.

Farshad, F.F., J.L. Leblance, J.D. Garber, and J.G. Osorio (1996), A new correlation for bubble point pressure according to the separator conditions, SPE Paper 26644.

Fattahi, H., A. Gholami, S. Amiribakhtiar, and S. Moradi (2014), Estimation of asphaltene precipitation from titration data: a hybrid support vector regression with harmony search, *Neural Comput. Appl.* **26**, 4, 789-798, DOI: 10.1007/s00521-014-1766-y.

Gharbi, R.B., and A.M. Elsharkawy (1996), Neural network model for estimating the PVT properties of Middle East crude oils, SPE Paper 56850.

Gholami, A., M. Asoodeh, and P. Bagheripour (2014a), Fuzzy assessment of asphaltene stability in crude oils, *J. Dispersion. Sci. Technol.* **35**, 4, 556-563, DOI: 10.1080/01932691.2013.800457.

Gholami, A., M. Asoodeh, and P. Bagheripour (2014b), Smart determination of difference index for asphaltene stability evaluation, *J. Dispersion. Sci. Technol.* **35**, 4, 572-576, DOI: 10.1080/01932691.2013.805654.

Gholami, A., M. Asoodeh, and P. Bagheripour (2014c), How committee machine with SVR and ACE estimates bubble point pressure of crudes, *Fluid Phase Equilibr.* **382**, 139-149, DOI: 10.1016/j.fluid.2014.08.033.

Gholami, A., S. Moradi, M. Asoodeh, P. Bagheripour, and M. Vaezzadeh-Asadi (2014d), Asphaltene precipitation modeling through ACE reaping of scaling equations, *Sci. Chin. Chem.* **57**, 12, 1774-1780, DOI: 10.1007/s11426-014-5253-1.

Glaso, O. (1980), Generalized pressure-volume-temperature correlations, *J. Petrol. Technol.* **34**, 85-95, DOI: 10.2118/8016-PA.

Hemmati, M.N., and R. Kharrat (2007), A correlation approach for prediction of crude oil PVT properties, SPE Paper 26644, DOI: 10.2118/104543-MS.

Katz, D.L. (1942), Prediction of the shrinkage of crude oils, SPE Paper 26644.

Kazemi, K., S. Moradi, and M. Asoodeh (2013), A neural network based model for prediction of saturation pressure from molecular components of crude oil, *Energy Source A* **35**, 11, 1039-1045, DOI: 10.1080/15567036.2011.584127.

Keerthi, S.S., and C.J. Lin (2003), Asymptotic behavior of support vector machines with Gaussian kernel, *Neural Comput.* **15**, 7, 1667-1689, DOI: 10.1162/089976603321891855.

Knopp, C.R., and L.A. Ramsey (1960), Correlation of oil formation volume factor and solution gas-oil ratio, *J. Petrol. Technol.* **12**, 8, 27-29, DOI: 10.2118/1433-G.

Mahmood, M.A., and M.A. Al-Marhoun (1996), Evaluation of empirically derived PVT properties for Pakistani crude oils, *J. Petrol. Sci. Eng.* **16**, 4, 275-290, DOI: 10.1016/S0920-4105(96)00042-3.

MATLAB User's Guide (2011), Fuzzy logic, neural network & GA and direct search toolboxes, MATLAB CD-rom, Mathworks, Inc.

Moghadam, J.N., K. Salahshoor, and R. Kharrat (2011), Introducing a new method for predicting PVT properties of Iranian crude oils by applying artificial neural networks, *Petrol Sci. Technol.* **29**, 10, 1066-1079, DOI: 10.1080/10916460903551040.

Mohaghegh, S. (2000), Virtual-intelligence applications in petroleum engineering: Part 2 – evolutionary computing, *J. Petrol. Technol.* **52**, 10, 40-46, DOI: 10.2118/61925-JPT.

Na'imi, S.R., A. Gholami, and M. Asoodeh (2014), Prediction of crude oil asphaltene precipitation using support vector regression. *J. Disper. Sci. Technol.* **35**, 4, 518-525, DOI: 10.1080/01932691.2013.798585.

Obomanu, D.A., and G.A. Okpobiri (1987), Correlating the PVT properties of Nigerian crudes, *J. Energy Resour. Technol.* **109**, 4, 214-217, DOI: 10.1115/1.3231349.

Omar, M.I., and A.C. Todd (1993), Development of new modified black oil correlations for Malaysian crudes, SPE-25338-MS, Asia Pacific Oil and Gas Conference, 8-10 February, Singapore, DOI: 10.2118/25338-MS.

Petrosky, J., and F. Farshad (1993), Pressure volume temperature correlation for the Gulf of Mexico, SPE, Annual Technical Conference and Exhibition, 3-6 October, Houston, USA, SPE-26644-MS, DOI: 10.2118/26644-MS.

Rafiee-Taghanaki, S., M. Arabloo, A. Chamkalani, M. Amani, M.H. Zargari, and M.R. Adelzadeh (2013), Implementation of SVM framework to estimate PVT properties of reservoir oil, *Fluid Phase Equilibr.* **346**, 25-32, DOI: 10.1016/j.fluid.2013.02.012.

Shokir, E.M. (2007), CO_2-oil minimum miscibility pressure model for impure and pure CO_2 streams, *J. Petrol. Sci. Eng.* **58**, 1-2, 173-185, DOI: 10.1016/j.petrol.2006.12.001.

Vapnik, V. (1995), *The Nature of Statistical Learning Theory*, Springer, New York.

Vazquez, M., and H.D. Beggs (1970), Correlation for fluid physical property prediction, *J. Petrol. Technol.* **32**, 06, 103-107, DOI: 10.2118/6719-PA.

You, Z., Z. Yin, K. Han, D. Huang, and X. Zhou (2014), A semi-supervised learning approach to predict synthetic genetic interactions by combining functional and topological properties of functional gene network, *BMC Bioinformatics* **11**, 343-355, DOI: 10.1186/1471-2105-11-343.

Zargar, G., P. Bagheripour, and M. Asoodeh (2014), Fuzzy modeling of volume reduction of oil due to dissolved gas run off and pressure release, *J. Petrol. Explor. Prod. Technol*. **4**, 4, 439-442, DOI: 10.1007/s13202-014-0099-9.

Received 26 August 2015
Received in revised form 19 April 2016
Accepted 15 June 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2530-2549
DOI: 10.1515/acgeo-2016-0110

Changes in Drought Conditions in Poland over the Past 60 Years Evaluated by the Standardized Precipitation-Evapotranspiration Index

Urszula SOMOROWSKA

University of Warsaw, Faculty of Geography and Regional Studies,
Warsaw, Poland; e-mail: usomorow@uw.edu.pl

Abstract

This paper investigates the variability of drought conditions in Poland in the years 1956-2015 with the use of the Standardized Precipitation-Evapotranspiration Index (SPEI). The study provides a new insight into the phenomenon of the past expansion of the drought-affected area as well as evidence of drying trends in a spatiotemporal context. 3-month, 6-month, and 12-month SPEI were considered, representing drought conditions relevant to agriculture and hydrology. The analysis demonstrates that the spatial extent of droughts shows a broad variability. The annual mean of the percentage of the area under drought has witnessed an increase for all three SPEI timescales. This also pertains to the mean area affected by drought over the growing season (April-September). A decreasing trend in the SPEI values indicates an increase in the severity of droughts over the 60-year period in question in an area extending from the south-west to the central part of Poland.

Key words: drought conditions, SPEI, changes, Poland.

1. INTRODUCTION

In recent decades an increase in the frequency and severity of summer droughts is reported to be an emerging issue globally (Kundzewicz 2008). This also concerns Poland. Prolonged dry and hot periods in the summer

---

Ownership: Institute of Geophysics, Polish Academy of Sciences;

lead to reduced recharges of soil moisture and groundwater, resulting as a consequence in long-lasting low flows and a deficit in the water balance (Tokarczyk 2013, Kędziora *et al.* 2014). Since the 1980s, Poland has experienced significant summer droughts. The period from 1982 to 2006 was marked by multiple years of extreme heat and precipitation shortfalls, resulting in widespread droughts (Łabędzki 2007, Lorenc *et al.* 2008). In a warming climate in Poland, an increase in the number of extremely warm days in a year and an increase in the maximum number of consecutive hot days have been observed for the period 1951-2010 (Graczyk and Kundzewicz 2014). Summer deficit does not show any statistically significant trend (Wibig 2012). As the summer precipitation deficit is projected to increase considerably in the future, Poland might face a high risk of water shortages in the next decades (Szwed *et al.* 2010).

Noteworthy is that an extremely hot and dry summer occurred in Poland in 2015. Significant dry conditions occurred in August across the whole country, from the Silesian Lowland, through the Wielkopolska Lowland and Mazovian Lowland to the Lublin Upland and Podlasie Lowland (IUNG-PIB 2015). Over the majority of the country, rainfall in August ranged from 10 to 30% of the long-term norm. As a consequence, discharge from the Vistula basin in August constituted only 42.5% of the long-term mean, and from the Oder Basin 35% (IMGW 2015). The River Vistula reached a new low record which was the lowest stage since records began in the eighteenth century. Simultaneously, in many rivers the stages (and discharges) fell to the lowest values, reaching the absolute minimum registered since the 1950s, especially in August and September (IMGW 2015).

As there is evidence of several drought events in Poland during the last decades, a new assessment of drought trends over the past 60 years seems to be a challenging issue. It might give new insights into an expansion of the area affected by drought in the past, and into the evidence for drought trends in a spatiotemporal context. A range of different single or combined indicators is already used to detect and monitor droughts (*e.g.*, Zargar *et al.* 2011, Łabędzki and Bąk 2014, Ziese *et al.* 2014). The most commonly used is the Standardized Precipitation Index (WMO 2012), also used in drought studies in Poland (*e.g.*, Osuch *et al.* 2015, Radzka 2015). In addition to the Standardized Precipitation Index (SPI), its newer variant, called the Standardized Climatic Water Balance (SCWB), was introduced in Poland by Łabędzki and Bąk (2004) and used for an assessment of regional droughts. The difference between precipitation and Penman–Monteith reference evapotranspiration was utilized. More recently proposed is the Standardized Precipitation-Evapotranspiration Index (SPEI) which is based on the same concept, using the difference between precipitation and potential evapotranspiration (Vicente-Serrano *et al.* 2010). It was designed as an improved drought index for studies of the effect of warming on drought severity (Begueria *et al.* 2014). The advantage of the SPEI (alternatively called the SCWB) over the SPI is that it is based not only on precipitation, but includes the component of potential evapotranspiration (PET). It normalizes anomalies in accumulated climatic water balance, calculated as the difference between precipitation and potential evapotranspiration. Different evapotranspiration equations might be applied in the SPEI calculation (Stagge *et al.* 2014), among which there is a Thornthwaite equation (Thornthwaite 1948), based on air temperature with an adjustment being made for the number of daylight hours. This method, requiring only limited data, was applied in the original SPEI methodology proposed by Vicente-Serrano *et al.* (2010) and is used in the SPEI Global Drought Monitor (Begueria *et al.* 2010). The choice of a more sophisticated PET method is limited by higher input requirements. Previous studies proved that the largest differences between SPEI calculated using different PET equations occur during the winter and spring, whereas the best agreement occurs during the summer (Stagge *et al.* 2014). Thus it might justify the choice of SPEI data from the Global Drought Monitor as first guess data, to consider and investigate, especially the summer droughts. Another SPEI data, the SPEIbase, is based on the FAO-56 Penman–Monteith estimation but at the moment it covers the temporal range up to December 2014 only. In Poland, the SPEI based on the Thornthwaite equation was investigated by Wibig, using data from 18 synoptic stations for the years 1951-2006 (Wibig 2012). In the context of the recent drought that occurred in 2015, further research on an expanded temporal window might give new evidence of drought severity trends, proving or contradicting previous findings.

This study analyzes the changes in the areas under drought in Poland over the past sixty years and gives an insight into drying trends evaluated by the SPEI over the long-term period, chosen here as 1956-2015.

2. DATA AND METHODS

The SPEI data used in this study were acquired from the Global Drought Monitor database, in which the PET is calculated by the Thornthwaite equation (Begueria *et al.* 2010). Climate data used for the SPEI calculation include air temperature data from the station observation-based global land monthly mean surface air temperature dataset at 0.5° spatial resolution, developed at the Climate Prediction Center, National Centers for Environmental Prediction in the US (Fan and van den Dool 2008). Additionally, monthly precipitation sums data were acquired for the SPEI calculation from the Global Precipitation Climatology Center (GPCC). The "first guess" monthly land-surface precipitation product at 1.0° spatial resolution (Ziese *et al.* 2011), interpolated to a resolution of 0.5°, is applied. The SPEI time series over Poland have been retrieved at 196 grid cells (Fig. 1) for the period

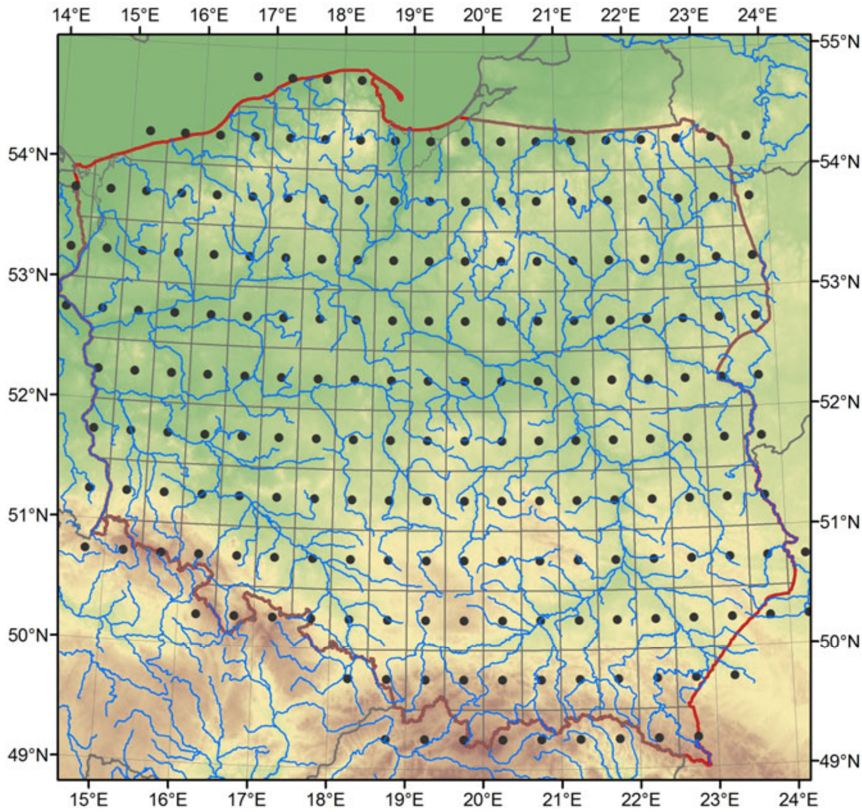


Fig. 1. Spatial distribution of grid cells of the SPEI Global Drought Monitor covering the territory of Poland.

January 1956 up to December 2015. Data were downloaded from online resources (http://sac.csic.es/spei). The dataset at a 0.5° spatial resolution includes different time-scales between 1 and 48 months. In this study, three time scales, 3, 6, and 12 months, have been selected, representing dryness/ wetness conditions relevant to agriculture and hydrology (WMO 2012). The SPEI-3 represents cumulative moisture conditions for the 3-month period. For example, a 3-month SPEI at the end of June represents cumulative moisture conditions for April–May–June. Similarly, the SPEI-6 and the SPEI-12 represent cumulative wetness conditions for the 6-month and 12-month periods. Positive values of the SPEI indicate wetness conditions wetter than average, whilst negative values indicate conditions drier than average. A drought is considered to occur when the SPEI value is less than or equal to −1. Three different drought categories were distinguished according to the SPEI value: moderate (D1), severe (D2), and extreme (D3) events (Table 1).

Table 1

Dryness/wetness categories according to the SPEI values

SPEI value	Dryness/wetness category
≤ −2.00	Extreme drought (D3)
−1.99 to −1.50	Severe drought (D2)
−1.49 to −1.00	Moderate drought (D1)
−0.99 to 0.99	Near normal
1.00 to 1.49	Moderately wet
1.50 to 1.99	Severely wet
≥ 2.00	Extremely wet

Area affected by 3-month and 12-month droughts of three different categories at a country level was determined by summing up the area of grid cells (Fig. 1). Based on that, the most widespread drought events were detected over the years 1956-2015. Averaging monthly values of percent of area under drought in each year, the annual mean was calculated and checked for a trend or tendency. Similarly, seasonal means for the winter half (November−April) of the year and for the summer half (May−October), and for the growing season (April−September), were calculated and tested for any changes.

Following this, time series of SPEI-3 and SPEI-12 at each grid cell were checked for each month, whether or not there is a long-term trend. Independently, a long-term trend analysis was conducted for the SPEI-3 and SPEI-12 averaged over the growing season (April−September). Additionally, a long-term trend analysis was conducted for the SPEI-6, based on the time series at each grid cell for September, ending the 6-month growing season. The non-parametric rank-based Mann–Kendall test was applied to detect drying or wetting trends of the SPEI. It is one of the most widely used methods for hydro-meteorological time series trend detection (Radziejewski and Kundzewicz 2004a, Machiwal and Jha 2012) applied formerly, among others, in the trend analysis of drought indices (*e.g.*, Wibig 2012, Damberg and AghaKouchak 2014, Potop *et al.* 2014). The HYDROSPECT software (Radziejewski and Kundzewicz 2004b) was used to calculate the Mann–Kendall test statistic (Z), and the statistical significance. Negative values of Z indicate decreasing trends in the SPEI (drying trend) whilst positive Z values characterize increasing trends (wetting trends). Trends were tested at the threshold values of significance level. Significance levels of 99.9, 99, 95, and 90% correspond to $|Z|$ values of 3.290, 2.575, 1.960, and 1.645. The Kendall–Theil robust line was used to quantify the magnitude of the identified trends (Theil 1950, Helsel and Hirsch 2002). The Kendall–Theil method was chosen as an alternative to simple linear regression because it requires no assumption of the data distribution and is less sensitive to outliers. It has been applied in many hydrological and environmental studies (*e.g.*, Wang *et al.* 2014, Zhang *et al.* 2015).

3. RESULTS AND DISCUSSION

3.1 Area under drought

Figures 2 and 3 provide an insight into the temporal evolution of the percentage of the country area under drought, evaluated respectively by SPEI-3 and SPEI-12. The most widespread 3-month extreme summer drought events (Fig. 2a) affected 46% of the country in April 1974, 43% in August 1992, and 47% in August 2015, whilst the most extensive winter drought covered 43% of the territory in March 1989, 65% in January 1997, and 39% in November 2011. Considering the area under extreme (D3) and severe (D2) droughts together (Fig. 2c), the most widespread summer events occurred in April 1974 (78%), August 1992 (87%), and August 2015 (70%), whilst in the winter half – they were in January 1997 (91%), March 1989 (88%), and November 2011 (73%). The percentage of area under drought conditions of D1, D2, and D3 together, exceeding 90% of the country territory, occurred in December 1957, March 1972, April 1974, November 1982, March 1989, August 1992, January−February 1997, November 2005, and August−September 2015. The year 1959 was also relatively dry, with the peak in May, when 89% of the country was in drought. In 47 months of the period from January 1956 until December 2015, the percentage of area under the 3-month drought (SPEI ≤ -1) was larger than 70%, comprising both droughts appearing in the summer and winter halves of the year.

The most widespread extreme 12-month drought (Fig. 3a) occurred in August 2015 (44%), September 2015 (41%), and October 2015 (28%). The occurrence of drought in a sequence of months shows its persistence. A relatively large area was detected in the sequential months from May till September 1983, covering an area of 12-25% of the country's territory. Considering the area under drought conditions D3 and D2 together (SPEI ≤ -1.5), such a sequential occurrence of dry months took place over the whole period of analysis (Fig. 3c). However, the widest drought occurred again in a sequence of months in 2015, increasing from April (13%) till August (83%), and then decreasing from September (75%), through October (63%) till November (36%). The area under drought conditions D1, D2, and D3 together (SPEI ≤ -1) was largest in 2015 (Fig. 3e). A sequence of dry months covering a large area of the country occurred already in September 2014 and lasted through the entire year 2015. The most widespread drought lasted from April 2015 (42%) till August and September 2015 (98%). Such sequences of

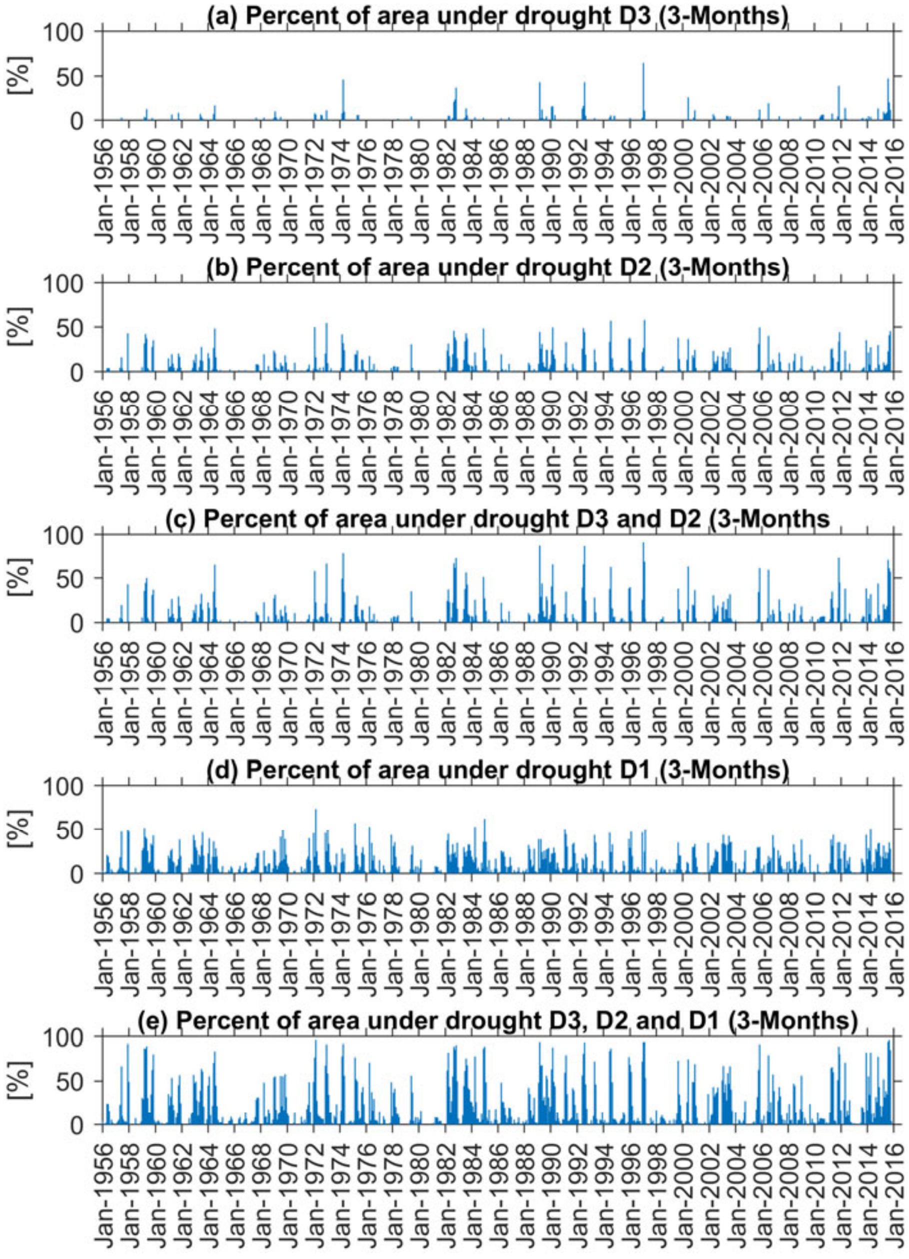


Fig. 2. Percent of area under the 3-month drought in the period 1956-2015: D3 (a), D2 (b), D3 and D2 (c), D1(d), and D3, D2 and D1 (e).

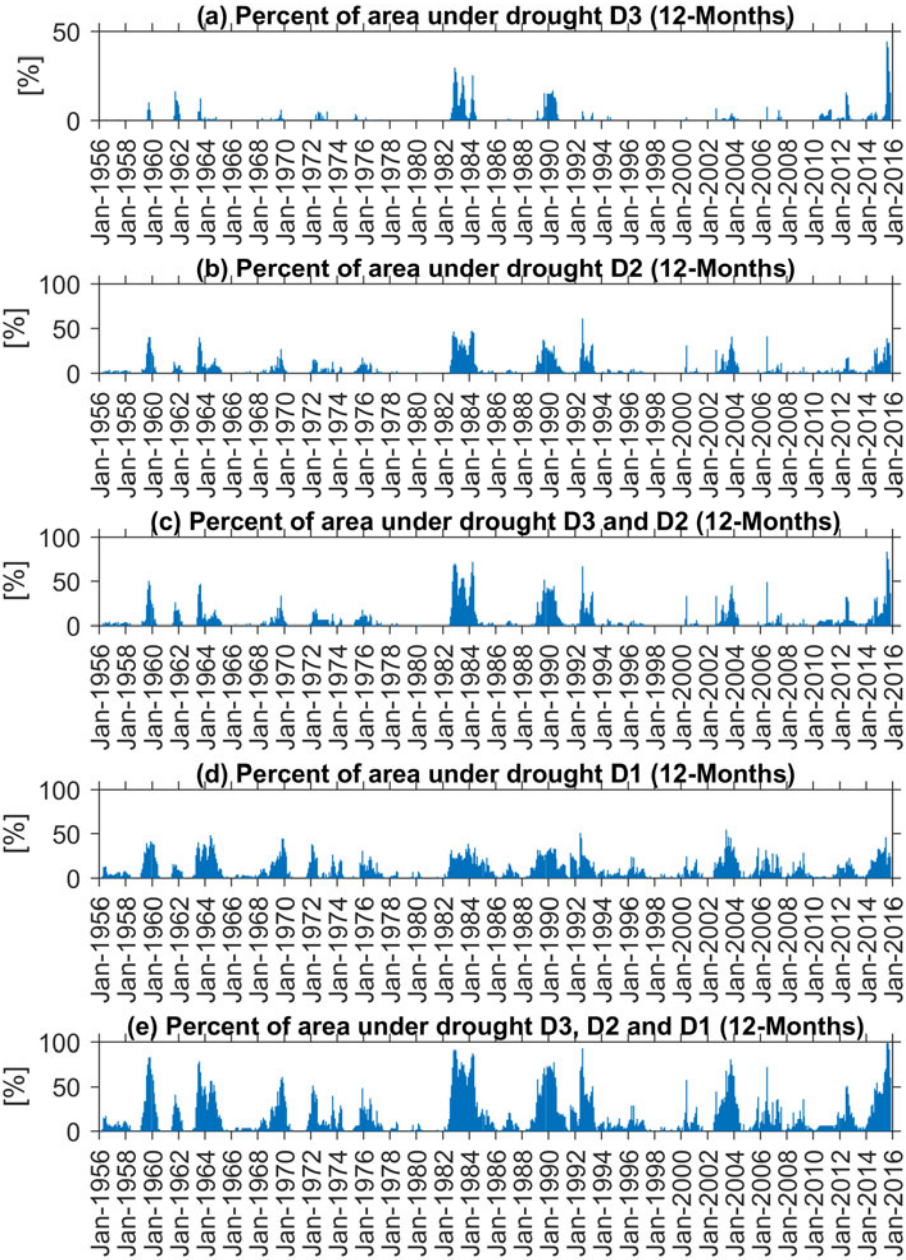


Fig. 3. Percent of area under the 12-month drought in the period 1956-2015: D3 (a), D2 (b), D3 and D2 (c), D1(d), and D3, D2 and D1 (e).

dry months (SPEI ≤ –1) occurred also in the past; the longest and the most widespread events took place in 1959-1960, 1963-1965, 1969, 1972-1974, 1976, 1982-1984, 1988-1990, 1992-1993, 2002-2003, and 2006.

The results show that large inter-annual variability in the area under drought exists. The annual mean of the percentage of area under drought, calculated by the Kendall–Theil robust line method and tested for significance by the Mann–Kendall test, increased in the years 1956-2015, with a change of 0.087%·yr^{-1} (270 $km^2·yr^{-1}$) for the 3-month droughts, and 0.052%·yr^{-1} (162 $km^2·a^{-1}$) for the 12-month droughts (Table 2). The long-term series of mean areas affected by drought over the growing season (April−September) show an increase with a rate of 0.105%·yr^{-1} (328 $km^2·yr^{-1}$) and 0.064%·yr^{-1} (200 $km^2·yr^{-1}$), respectively, for the 3-month and 12-month droughts. Much lower is an increase of the area under 6-month droughts appearing in the growing season, calculated both as a 6-month mean and for September only; it is within the range of 144-178 $km^2·yr^{-1}$ (Table 2). It is worth noting that the highest rate of increase concerns the area under the

Table 2

Summary statistics of changes of drought area in Poland in the years 1956-2015

Time series: Year Season Month	Mann–Kendall test statistics		Rate of change evaluated by Kendall–Theil robust line	
	Test statistic Z	Significance level [%]	Area percent [%· yr^{-1}]	Area [km^2· yr^{-1}]
3-month droughts (D3, D2, and D1)				
Year	0.969	67	0.087	270
November–April	0.625	47	0.035	110
May–October	1.352	82	0.117	367
April–September	1.276	80	0.105	328
6-month droughts (D3, D2, and D1)				
Year	0.944	65	0.077	241
November–April	1.225	78	0.087	272
May–October	0.680	51	0.043	136
April–September	0.561	43	0.046	144
M09	1.050	71	0.057	178
12-month droughts (D3, D2, and D1)				
Year	0.829	59	0.052	162
November–April	0.612	46	0.045	141
May–October	1.033	70	0.079	219
April–September	0.944	65	0.064	200

3-month drought in the summer season (May−October) and is 0.117%·yr^{-1} (367 $km^2 \cdot yr^{-1}$).

The drought conditions over Poland, detected in this study, refer to the most relevant European drought events evaluated by the combined indicators, reported recently by Spinoni *et al.* (2015). Among the list of 22 big European multi-region drought events that occurred from 1950 until 2011, ten of them concerned central and eastern Europe and were reflected also in Poland. The confirmed occurrence of such widespread droughts, marked also in Poland, concerns the years 1959, 1964, 1972-1974, 1976, 1983, 1992, 1996-1997, 2003, 2006, and 2011.

3.2 Changes in the SPEI over an entire year

In order to check if there is a trend in the SPEI values, 60-element series of singular SPEI-3 and SPEI-12 values were prepared for each grid cell, for each month. Then, the Mann–Kendall test was applied. Results are presented in Figs. 4 and 5. In the SPEI-3 monthly series (Fig. 4), a statistically significant trend occurred in many grid cells in the months April−October, covering 25% of the country's territory in April and 21% in October (Table 3).

Table 3

Percent of the country area with drying trend and drying signals

Month	Percent of area [%]			
	3-month SPEI		12-month SPEI	
	Drying trend	Drying signal	Drying trend	Drying signal
January	7	38	18	45
February	3	18	19	47
March	0	6	18	46
April	25	44	18	46
May	13	32	23	44
June	16	43	19	43
July	6	30	23	49
August	17	62	33	64
September	13	56	32	63
October	21	58	30	56
November	7	40	29	57
December	13	65	23	52

Explanations: Drying trend is assumed to occur for the test statistic of the Mann–Kendall test Z values ≤ -1.645. Drying signal is assumed to occur for the rate of change of the SPEI values ≤ -0.005 yr^{-1}, calculated as a slope of the Kendall–Theil robust line.

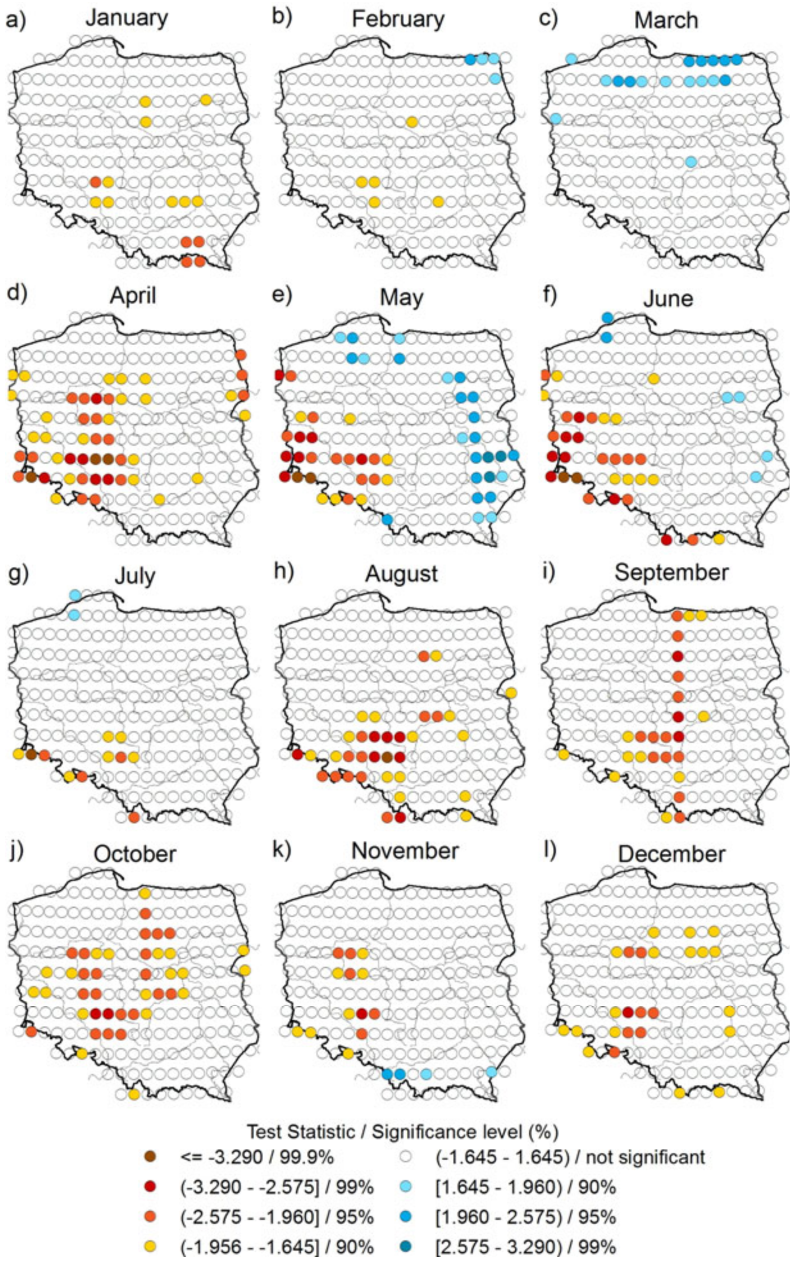


Fig. 4. Mann–Kendall test results of trend detection over Poland at the 90, 95, 99, and 99.9% significance levels for the SPEI-3 monthly series. Yellow, red and brown circles indicate decreasing trend in SPEI values, blue circles – increasing, whereas blank one – no trend detected.

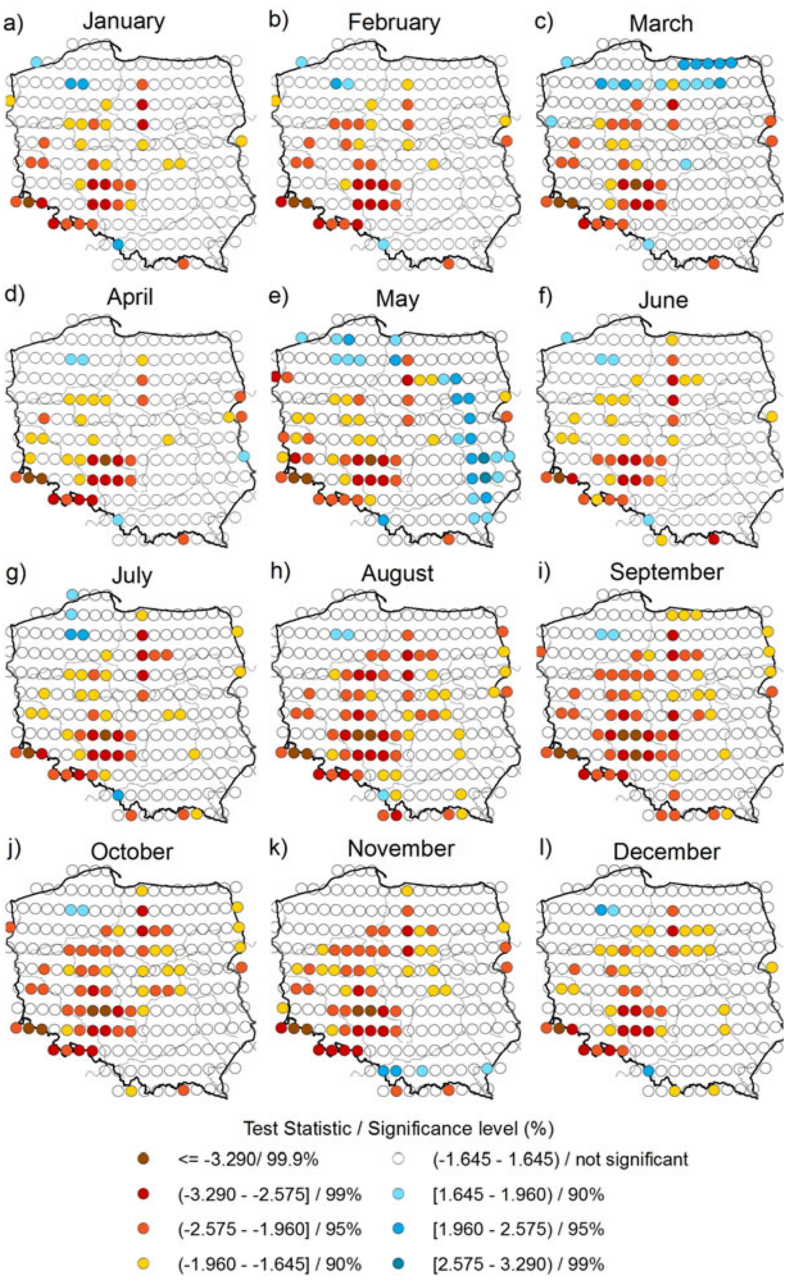


Fig. 5. Mann–Kendall test results of trend detection over Poland at the 90, 95, 99, and 99.9% significance levels for the SPEI-12 monthly series. Yellow, red and brown circles indicate decreasing trend in SPEI values, blue circles – increasing, whereas blank one – no trend detected.

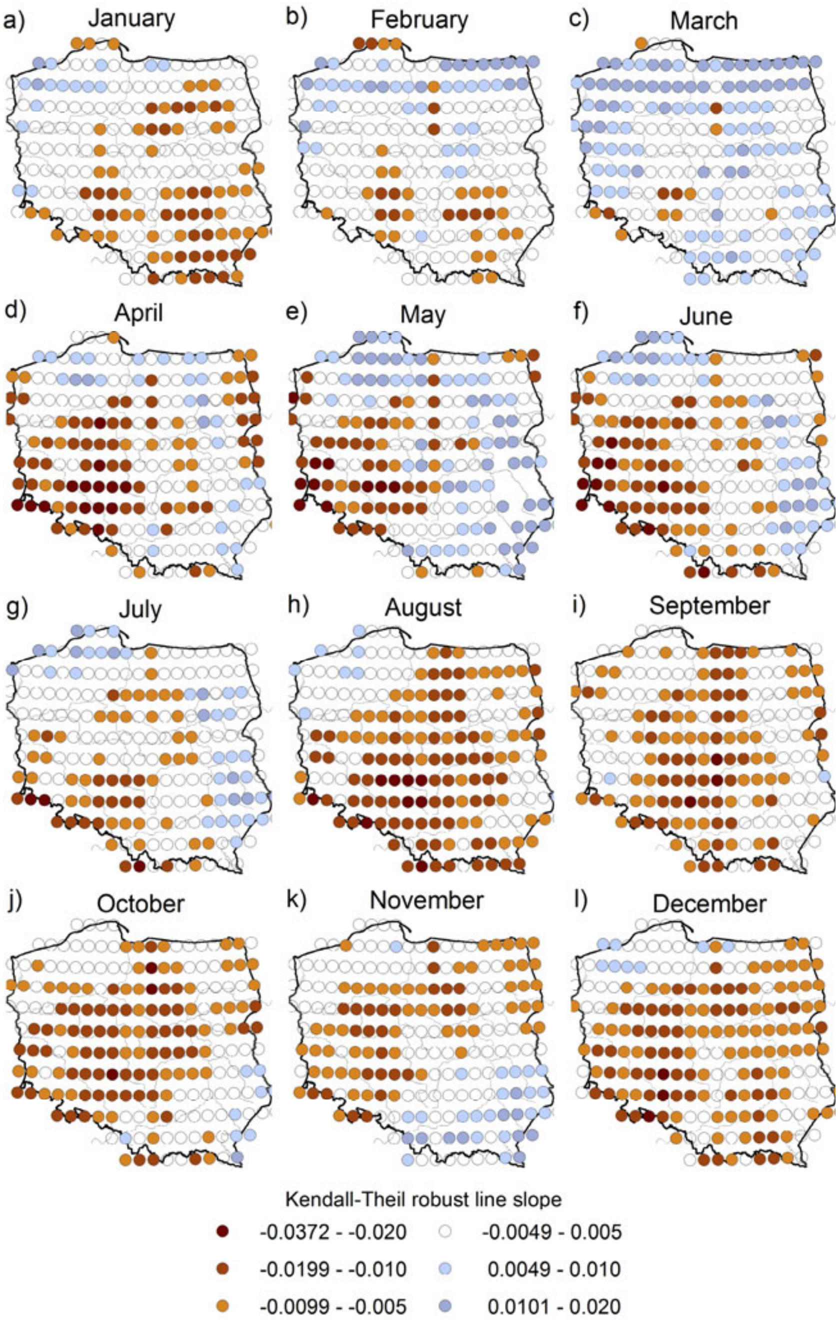


Fig. 6. Rate of change of the SPEI-3 values over Poland in the years 1956-2015, evaluated by the Kendall–Theil robust line method.

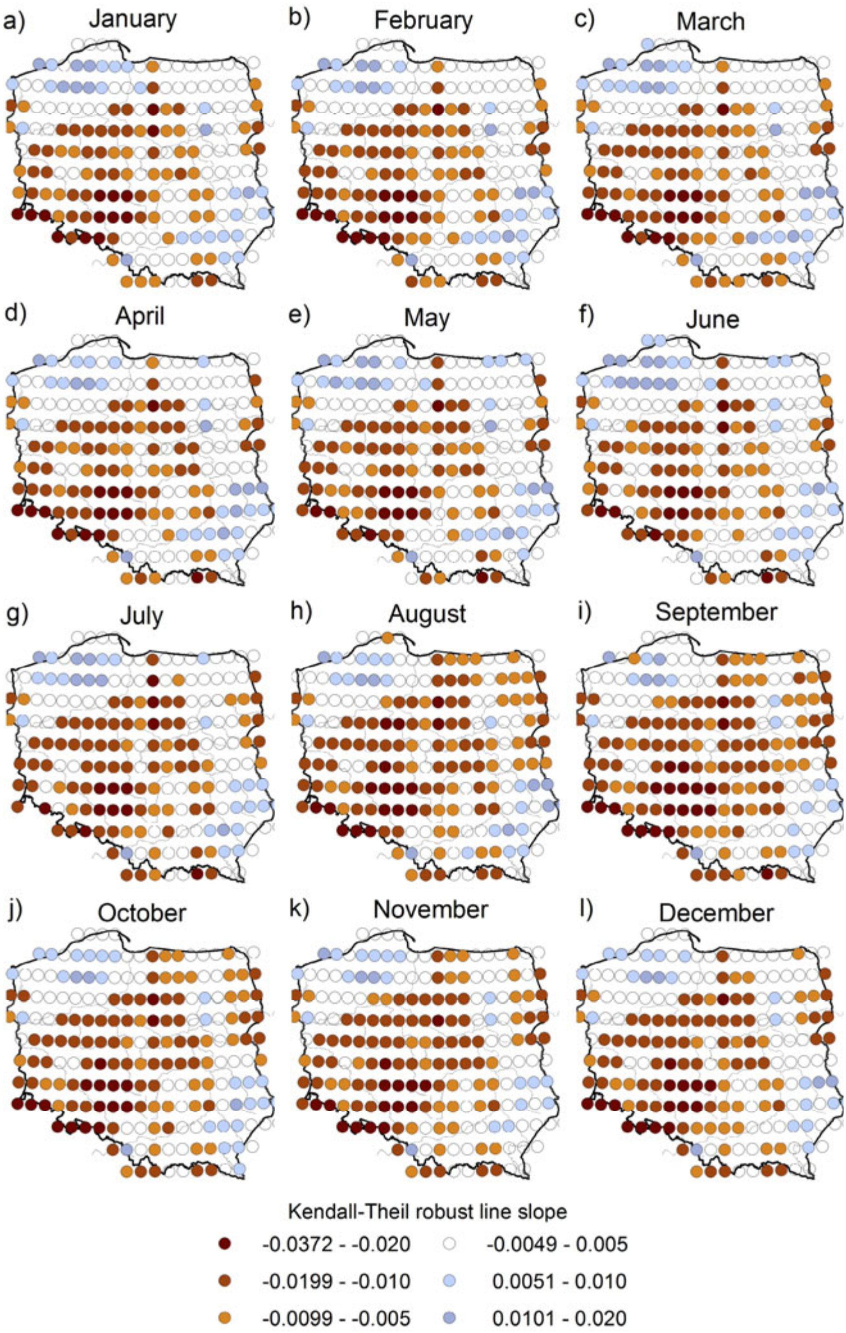


Fig. 7. Rate of change of the SPEI-12 values over Poland in the years 1956-2015, evaluated by the Kendall–Theil robust line method.

The spatial pattern of grid cells with decreasing (drying) trend extends from the south-west to the center of the country. It is especially visible in April (Fig. 4d), August (Fig. 4h), and October (Fig. 4j), whereas in other months, such a “compact” group of cells does not show up. It is worth mentioning that in May (Fig. 4e) and March (Fig. 4c) an increasing trend is marked, respectively, in the east and north. A much more consolidated group of grid cells with decreasing trends was detected in the monthly SPEI-12 series (Fig. 5). Here, the area affected by a decreasing trend extends from south-west to the central parts of the country, throughout all months. The largest spatial extent of a drying trend occurs in August, September, and October, covering approximately 30-33% of the territory (Fig. 5h-j, Table 3). An increasing trend concerns selected pixels only. Considering the rate of change of the SPEI over the 60-year period, evaluated by the Kendall–Theil robust line (Figs. 6 and 7), it is worth mentioning that a considerable part of Poland experiences drying signals (negative values of slope of the Kendall–Theil robust line). This concerns approximately 56-64% of the territory in August, September, and October, both for the 3-month and the 12-month SPEI (Table 3).

The obtained results concerning drying trends over Poland coincide with previously conducted studies, although for different long-term periods. Łabędzki *et al.* (2014) investigated reference evapotranspiration, based on 18 stations across Poland. It was proved that there is an increasing trend in the reference evapotranspiration in 1971-2010, explained by the trends of air temperature and sunshine as the main factors determining evapotranspiration. Simultaneously, it was detected that the reference evapotranspiration pattern across Poland showed differences, with the highest values recorded in central Poland, from west to east, where there is relatively low precipitation, causing significant meteorological and agricultural droughts. The region of the highest increase of reference evapotranspiration, although based only on 18 stations, coincides with the region of statistically significant drying trends of SPEI-3 and SPEI-12, detected in this study. However, the SPEI is considered to be a more complex index in detecting drying trends, including not only evapotranspiration but also precipitation.

3.3 Changes in the SPEI over the growing season

In order to check if there is a trend in the SPEI values over the growing season, the SPEI values were averaged over the 6-month period from April to September. Then, the 60-element series of the SPEI-3 and SPEI-12 were prepared for each grid point and the Mann–Kendall test was performed (Fig. 8a, c). For comparison, the SPEI-6 series for September was investigated, representing the cumulative moisture conditions for the six months of

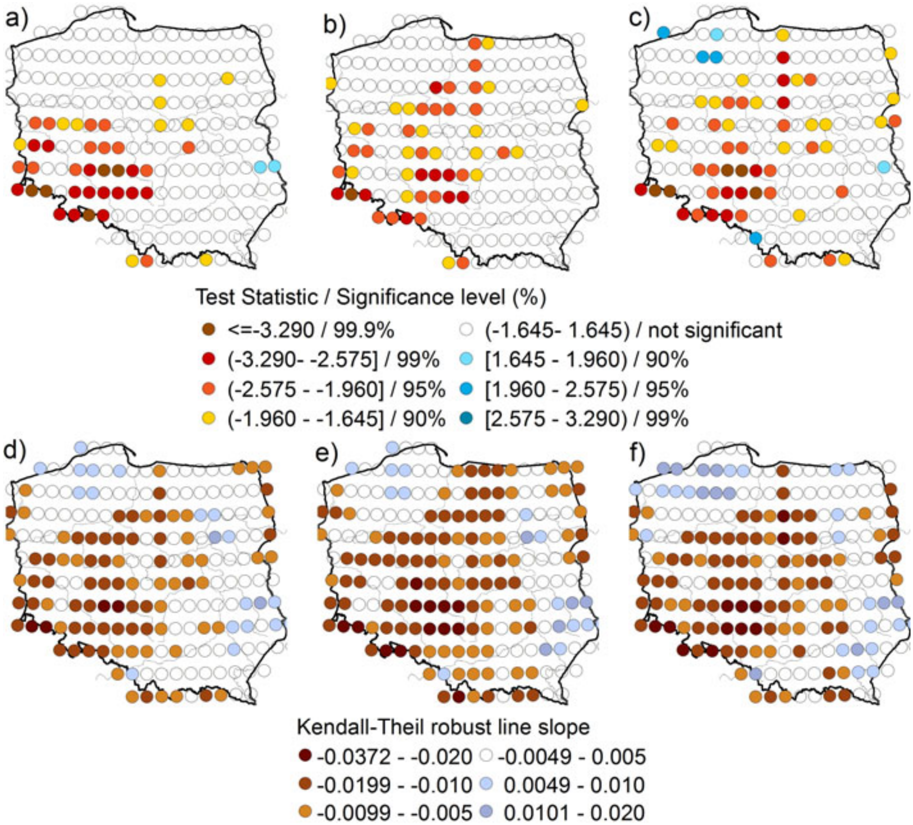


Fig. 8. Mann–Kendall test results of trend detection over the growing season (April–September) for the SPEI-3 (a), SPEI-6 for September (b), and SPEI-12 (c). Rate of change evaluated by the Kendall–Theil robust line slope in the time series of the SPEI-3 (d), SPEI-6 for September (e), and SPEI-12 (f).

the growing season (Fig. 8b). Additionally, the rate of the SPEI change was estimated by the Kendall–Theil robust slope for the respective SPEI time scales (Fig. 8d-f). Drying trends occurred in the region extending from the south-west to the center of the country and over single grid cells in the north-east. The changes of the SPEI values, reflecting the drying, range from –0.03 yr^{-1} to –0.005 yr^{-1}. It is worth mentioning that the spatial extent of a region with drying signals is much wider than that with a drying trend as it includes also these grid cells where there is no trend but the drying tendency is pronounced. The drying trend concerns approximately 21% of the country for the SPEI-3, 25% for the SPEI-6 (for September), and 27% for the SPEI-12 (Fig. 8a-c). The drying signals occur on 42% of the area in the case of SPEI-3, on the 58% for SPEI-6 (for September), and on the 47% for SPEI-12.

4. CONCLUDING REMARKS

This paper analyzes the extent of the drought-affected area in Poland over the 1956-2015 period and drying trends evaluated by the Standardized Precipitation-Evapotranspiration Index (SPEI). The analysis extends to the year 2015, which was marked as extremely hot and dry. The 3-month, 6-month, and 12-month SPEI values were investigated, representing drought conditions relevant to agriculture and hydrology.

The analysis shows a broad variability in the spatial extent of droughts on the territory of Poland. Drought events of differing severity occur in both the winter and summer halves of the year. The most widespread 3-month extreme summer drought occurred in August 2015, affecting 47% of the national land area. The most widespread extreme 12-month drought also occurred in August 2015, covering 44% of the aforesaid area. The percentage of the drought-affected area has witnessed an increase for all three SPEI timescales, albeit with different rates of change. This pertains to both the annual mean of the area under drought as well as to the drought-affected area separately in the summer and winter halves of the year.

A relatively large area of Poland exhibits significant drying trends, extending from the south-west towards the center part of the country and, in some cases, to the north-east. This especially pertains to months in the summer half of the year, as well as to the growing season considered as a whole. The drying trends during the growing season extend to approximately 21% of the total land area for SPEI-3, 25% for SPEI-6 (for September), and 27% for SPEI-12. However, drying signals occur over a much larger percentage of the total area, which amounts to over 40% in the case of SPEI-3 and SPEI-12, and as much as 58% in case of SPEI-6. In effect, the results reveal that drying trends or signals on the territory of Poland have affected a substantial area of the country. The year 2015 has witnessed considerably higher-than-normal air temperature and subnormal precipitation, which were reflected by the relatively long period of negative SPEI values.

In conclusion, it is worth noting that prolonged drying trends might have a subsequent negative impact on both the environment and society. As Poland has rather scarce water resources, the preparation of adaptation strategies to counteract climate change through appropriate water programs seems to be a challenging issue.

Acknowledgments. The author expresses her gratitude to the anonymous reviewers for their constructive and helpful remarks and suggestions.

References

Beguería, S., S.M. Vicente-Serrano, and M. Angulo-Martinez (2010), A multiscalar global drought dataset: the SPEIbase: a new gridded product for the analysis of drought variability and impacts, *Bull. Am. Meteorol. Soc.* **91**, 1351-1356, DOI: 10.1175/2010BAMS2988.1.

Beguería, S., S.M. Vicente-Serrano, F. Reig, and B. Latorre (2014), Standardized precipitation evapotranspiration index (SPEI) revisited: parameter fitting, evapotranspiration models, tools, datasets and drought monitoring, *Int. J. Climatol.* **34**, 10, 3001-3023, DOI: 10.1002/joc.3887.

Damberg, L., and A. AghaKouchak (2014), Global trends and patterns of drought from space, *Theor. Appl. Climatol.* **117**, 3, 441-448, DOI: 10.1007/s00704-013-1019-5.

Fan, Y., and H. van den Dool (2008), A global monthly land surface air temperature analysis for 1948–present, *J. Geophys. Res.* **113**, D1, D01103, DOI: 10.1029/2007JD008470.

Graczyk, D., and Z.W. Kundzewicz (2014), Changes in thermal extremes in Poland, *Acta Geophys.* **62**, 6, 1435-1449, DOI: 10.2478/s11600-014-0240-7.

Helsel, D.R., and R.M. Hirsch (2002), *Statistical Methods in Water Resources*, Techniques of Water Resources Investigations, Book 4, Chapter A3, U.S. Geological Survey, 395 pp.

IMGW (2015), Bulletin of the National Hydrological and Meteorological Service, Institute of Meteorology and Water Management (IMGW), State Research Institute, Warsaw, Poland, No. 13, 65 pp.

IUNG-PIB (2015), Communication report regarding the incidences of drought conditions in Poland, period: 13 (1.VIII - 30.IX), Institute of Soil Science and Plant Cultivation – State Research Institute (IUNG-PIB), Puławy, Poland, available from: http://www.susza.iung.pulawy.pl/en/arch15 (accessed: 4 September 2016).

Kędziora, A., M. Kępińska-Kasprzak, P. Kowalczak, Z.W. Kundzewicz, A.T. Miler, E. Pierzgalski, and T. Tokarczyk (2014), Risks resulting from water shortages, *Nauka* **1**, 149-172 (in Polish).

Kundzewicz, Z.W. (2008), Hydrological extremes in the changing world, *Folia Geograph. Ser. Geograph. Phys.* **39**, 37-52.

Lorenc, H., M. Mierkiewicz, and M. Sasim (2008), Drought in Poland with special regards to the year 2006, *Wiadomości IMGW* **2**, 1-2, 3-32 (in Polish).

Łabędzki, L. (2007), Estimation of local drought frequency in Central Poland using the standardized precipitation index SPI, *Irrig. Drainage* **56**, 1, 67-77, DOI: 10.1002/ird.285.

Łabędzki, L., and B. Bąk (2004), Standardized climatic water balance a drought index, *Acta Agrophys.* **3**, 1, 117-124 (in Polish).

Łabędzki, L., and B. Bąk (2014), Meteorological and agricultural drought indices used in drought monitoring in Poland: a review, *Meteorol. Hydrol. Water Manag*. **2**, 2, 3-14.

Łabędzki, L., B. Bąk, and K. Smarzyńska (2014), Spatio-temporal variability and trends of Penman–Monteith reference evapotranspiration (FAO-56) in 1971-2010 under climatic conditions of Poland, *Pol. J. Environ. Stud*. **23**, 6, 2083-2091.

Machiwal, D., and M.K. Jha (2012), *Hydrologic Time Series Analysis: Theory and Practice*, Springer, Dordrecht and Capital Publ. Co., New Delhi.

Osuch, M., R.J. Romanowicz, D. Lawrence, and W.K. Wong (2015), Assessment of the influence of bias correction on meteorological drought projections for Poland, *Hydrol. Earth Syst. Sci. Discuss*. **12**, 10331-10377, DOI: 10.5194/hessd-12-10331-2015.

Potop, V., C. Boroneant,, M. Mozny, P. Stepánek, and P. Skalák (2014), Observed spatiotemporal characteristics of drought on various time scales over the Czech Republic, *Theor. Appl. Climatol*. **115**, 3-4, 563-581, DOI: 10.1007/s00704-013-0908-y.

Radziejewski, M., and Z.W. Kundzewicz (2004a), Detectability of changes in hydrological records, *Hydrol. Sci. J.* **49**, 1, 39-51, DOI: 10.1623/hysj.49.1.39.54002.

Radziejewski, M., and Z.W. Kundzewicz (2004b), Development, use and application of the HYDROSPECT data analysis system for the detection of changes in hydrological time-series for use in WCP – Water and National Hydrological Services, WCASP-65, Hydrospect, Version 2.0, User's manual, WMO, Geneva, Switzerland.

Radzka, E. (2015), The assessment of atmoshperic drought during vegetation season (according to standardized precipitation index SPI) in central-eastern Poland, *J. Ecolog. Eng*. **16**, 1, 87-91.

Spinoni, J., G. Naumann, and J. Vogt (2015), Spatial patterns of European droughts under a moderate emission scenario, *Adv. Sci. Res*. **12**, 179-186, DOI: 10.5194/asr-12-179-2015.

Stagge, J.H., L.M. Tallaksen, C.-Y. Xu, and H.A.J. van Lanen (2014), Standardized precipitation-evapotranspiration index (SPEI): Sensitivity to potential evapotranspiration model and parameters. **In:** *Proc. FRIEND-Water 2014*, IAHS Red Book, 363, Montpellier, France.

Szwed, M., G. Karg, I. Pińskwar, M. Radziejewski, D. Graczyk, A. Kędziora, and Z.W. Kundzewicz (2010), Climate change and its effect on agriculture, water resources and human health sectors in Poland, *Nat. Hazards Earth Syst. Sci*. **10**, 8, 1725-1737, DOI: 10.5194/nhess-10-1725-2010.

Theil, H. (1950), A rank-invariant method of linear and polynomial regression analysis, 1, 2, and 3, *Proc. R. Neth. Acad.* Sci. **53**, 386-392, 521-525, 1397-1412, DOI: 10.1007/978-94-011-2546-8_20.

Thornthwaite, C.W. (1948), An approach towards rational classification of climate, *Geograph. Rev*. **38**, 1, 55-94, DOI: 10.2307/210739.

Tokarczyk, T. (2013), Classification of low flow and hydrological drought for a river basin, *Acta Geophys*. **61**, 2, 404-421, DOI: 10.2478/s11600-012-0082-0.

Vicente-Serrano, S.M., S. Begueria, and J.I. Lopez-Moreno (2010), A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index, *J. Climate* **23**, 7, 1696-1718, DOI: 10.1175/2009JCLI2909.1.

Wang, J., Y. Sheng, and T.S.D. Tong (2014), Monitoring decadal lake dynamics across the Yangtze Basin downstream of Three Gorges Dam, *Remote Sens. Environ*. **152**, 251-269, DOI: 10.1016/j.rse.2014.06.004.

Wibig, J. (2012), Moisture conditions in Poland in view of the SPEI index, *Woda-Środowisko-Obszary Wiejskie* **12**, 2, 38, 329-340 (in Polish).

WMO (2012), *Standardized Precipitation Index. User Guide*, WMO No. 01090, World Meteorological Organization, Geneva, Switzerland.

Zargar, A., R. Sadiq, B. Naser, and F.I. Khan (2011), A review of drought indices, *Environ. Rev.* **19**, NA, 333-349, DOI: 10.1139/a11-013.

Zhang, K., J.S. Kimball, R.R. Nemani, S.W. Running, Y. Hong, J.J. Gourley, and Z. B. Yu (2015), Vegetation greening and climate change promote multidecadal rises of global land evapotranspiration, *Nat. Sci. Rep*. **5**, 15956, DOI: 10.1038/srep15956.

Ziese, M., A. Becker, P. Finger, A. Meyer-Christoffer, B. Rudolf, and A. Schneider (2011), GPCC First Guess Product at 1.0°: Near real-time first guess monthly land-surface precipitation from rain-gauges based on SYNOP data, DOI: 10.5676/DWD_GPCC/FG_M_100.

Ziese, M., U. Schneider, A. Meyer-Christoffer, K. Schamm, J. Vido, P. Finger, P. Bissolli, S. Pietzsch, and A. Becker (2014), The GPCC Drought Index – a new, combined and gridded global drought index, *Earth Syst. Sci. Data* **6**, 2, 285-295, DOI: 10.5194/essd-6-285-2014.

Received 29 March 2016
Received in revised form 4 September 2016
Accepted 7 September 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2550-2590
DOI: 10.1515/acgeo-2016-0069

Modelling and Observation of Mineral Dust Optical Properties over Central Europe

Michał T. CHILINSKI[1,2,5], Krzysztof M. MARKOWICZ[1], Olga ZAWADZKA[1,2], Iwona S. STACHLEWSKA[1], Wojciech KUMALA[1], Tomasz PETELSKI[3], Przemysław MAKUCH[3], Douglas L. WESTPHAL[4], and Bogdan ZAGAJEWSKI[5]

[1]Institute of Geophysics, Faculty of Physics, University of Warsaw, Warsaw, Poland; e-mail: mich@igf.fuw.edu.pl

[2]College of Inter-Faculty Individual Studies in Mathematics and Natural Sciences, Warsaw, Poland

[3]Institute of Oceanology, Polish Academy of Sciences, Sopot, Poland

[4]Marine Meteorology Division, Naval Research Laboratory, Monterey, CA, USA

[5]Department of Geoinformatics and Remote Sensing, Faculty of Geography and Regional Studies, University of Warsaw, Warsaw, Poland

Abstract

This paper is focused on Saharan dust transport to Central Europe/ Poland; we compare properties of atmospheric Saharan dust using data from NAAPS, MACC, AERONET as well as observations obtained during HyMountEcos campaign in June 2012. Ten years of dust climatology shows that long-range transport of Saharan dust to Central Europe is mostly during spring and summer. HYSPLIT back-trajectories indicate airmass transport mainly in November, but it does not agree with modeled maxima of dust optical depth. NAAPS model shows maximum of dust optical depth (~0.04-0.05, 550 nm) in April-May, but the MACC modeled peak is broader (~0.04). During occurrence of mineral dust over Central-Europe for 14% (NAAPS) / 12% (MACC) of days dust optical depths are above 0.05 and during 4% (NAAPS) / 2.5% (MACC) of days

Ownership: Institute of Geophysics, Polish Academy of Sciences;

dust optical depths exceed 0.1. The HyMountEcos campaign took place in June-July 2012 in the mountainous region of Karkonosze. The analysis includes remote sensing data from lidars, sunphotometers, and numerical simulations from NAAPS, MACC, DREAM8b models. Comparison of simulations with observations demonstrates the ability of models to reasonably reproduce aerosol vertical distributions and their temporal variability. However, significant differences between simulated and measured AODs were found. The best agreement was achieved for MACC model.

Key words: aerosol, mineral dust, MACC, NAAPS, DREAM, aerosol transport model, remote sensing.

1. INTRODUCTION

Tropospheric aerosol influence on the global climate system, via direct and indirect radiative forcing, is important for understanding climate changes and still has a lot of uncertainties in geophysical studies (IPCC 2014). Among different types of aerosols, mineral dust may have a high influence on climate radiative forcing due to the possibility of events with large aerosol load and large aerosol optical depth. Natural sources of mineral dust aerosols (mainly silicates) are responsible for approximately 30% (Jimenez *et al.* 2009) of the total aerosol optical depth (AOD) in the atmosphere. Global soil-derived mineral dust emissions were estimated to be from 60 to 3000 Tg/yr by Duce (1995) and 1840 Tg/yr by Schutgens *et al.* (2012), which makes it the aerosol with the highest emissions globally. Relevant modification of radiation flux by mineral dust comes from scattering and absorption in both short and long-wavelength spectrum (Chen *et al.* 2011). The influence on climate is complex in the case of mineral aerosols and it could lead to either warming or cooling (Chand *et al.* 2009). Over surfaces with a relatively high albedo, over 0.3 (*e.g.*, over desert, snow), the top of the atmosphere (TOA) radiative forcing of mineral dust is usually positive and it will warm the climate system. On the other hand, over dark surfaces of albedo, lower than 0.15 (*e.g.*, over oceans, coniferous forests), the TOA radiative forcing of mineral aerosols usually is negative and make the climate system cooler. In the range between light and dark surfaces, where the albedo is higher than 0.15 and lower than 0.30 (Balkanski *et al.* 2007), the influence on climate can be in both directions and it depends on additional factors (*e.g.*, particles shape and size distribution, or particles refractive index). Those variables could be very different on a regional scale, which makes an estimation of mineral dust participation in global radiative forcing significantly uncertain (Balkanski *et al.* 2007). From an observer's point of view, selecting an appropriate technique for determining mineral dust radiative forcing is a complex issue due to three main factors: assumptions in particle shapes and related parameterizations of non-spherical particles used in many retrieval algorithms (Wang *et al.* 2013); insufficient number of measurements in the infrared spectral range where, in general, dust strongly interferes with radiation with the same strength as solar radiation (Vogelmann *et al.* 2003, Markowicz *et al.* 2003); and difficulties in a strict distinction of aerosol type during measurements (Sinha *et al.* 2012).

The most important sources of mineral dust particles over southern, western and Central Europe are the arid and semi-arid regions in Northern Africa, dominated by the presence of the world's largest hot desert, Sahara (Prospero *et al.* 2002). The amount of Saharan dust production plays an important role for the climate of the whole Earth (Guerrero-Rascado *et al.* 2009). The Saharan dust transport over Europe strongly depends on complex meteorological conditions (Di Sarra *et al.* 2001) which makes it irregular, with greater intrusion's frequency and dust load amount during spring and summer months (Papayannis *et al.* 2008, Pisani *et al.* 2011). Studies of the North African dust emission and transport reveal the highest production of dust during the May-August period (Engelstaedter *et al.* 2006), however the main transport trajectories are different between March-May and July-August (Isrealevich *et al.* 2003). Spring trajectories show transport of dust emitted during sand storms in the western direction, to the area of the Atlantic Ocean, where dust on higher altitudes can reach the central and northern areas of Europe. Summer transport of the dust is mainly in the northern direction which results in high occurrence of dust events in southern Europe and in many cases it eventually directs the transport through the Alps to reach Central Europe (Varga *et al.* 2013). Most of these dust events occur in the Mediterranean area and only a few of them reach the borders of Poland (Papayannis *et al.* 2005, 2008; Mona *et al.* 2012). Dust climatological studies based on simulations of the NAAPS model in the period of 9 years (1998-2006) showed that dust events in Poland occurred mainly in the spring (with the highest annual peak in May) and autumn, during October and November (Maciszewska *et al.* 2010). Regarding the irregularity of dust events appearance, it is uncertain what is the real impact of Saharan dust on aerosol radiative forcing over Central Europe, which motivates further research of dust impact on radiative forcing over Poland.

The best source of information on the dust physical and chemical properties comes from field campaigns (Formenti *et al.* 2008, Osborne *et al.* 2008, Heintzenberg 2009, Kandler *et al.* 2009, Tesche *et al.* 2009, Gross *et al.* 2011, Marsham *et al.* 2013) and by the model simulation. Lidar techniques are useful in atmospheric aerosol studies as they can provide data about aerosol properties with high temporal and spatial resolutions. Retrieval of information on the aerosol properties from lidar measurements is complex and requires the use of multi-wavelength lidars and/or additional data from different devices along with several assumptions (Weitkamp 2005). The other important remote sensing method involves passive observations with sun-photometers, which integrate optical properties of a whole air column providing accurate AOD and Angstrom exponent measurements. Measurements from sun-photometers can be used as an input for lidar retrieval algorithms (Landulfo *et al.* 2003, Lopes *et al.* 2013). Apart from remote sensing methods, *in situ* measurements of absorption coefficient with aethalometers and scattering coefficient with nephelometers are of great importance. Data obtained by the two latter instruments contribute to improve the information on the lowermost air layer, which usually is invisible for lidars due to the incomplete overlap between the emitted laser beam and the receiver's field of view (Guerrero-Rascado *et al.* 2010, Wandinger and Ansmann 2002). Thus, the athelometer-nephelometer (with polar nephelometer capable of measuring backward scattering) combined observations can deliver data for assumptions of a lidar ratio necessary for a simple elastic lidar or ceilometer data retrieval (Markowicz *et al.* 2008). Due to difficulties in conducting systematic field measurements and a sparse grid of measurement stations, it is very important to collect field data during dust events, which could be used for verifying of model simulations accuracy. Although the presented dust optical depths (DOD) values over Poland are small in comparison to the basin of Mediterranean Sea, it is still around 25% of total AOD in our region and mineral dust is one of the two most important types of aerosols above the boundary layer (together with products of burning). This fact, together with the lack of models' validations in Central Europe/Poland (area far from sources of dust), especially DREAM and NAAPS, was the main motivation for our study.

The aim of research described in this paper is an attempt to utilize different modelling and observation techniques to estimate the seasonal variation of the dust optical properties over Poland. This paper presents the findings of a field campaign in Karpacz, South-Western Poland, during the Hyperspectral Remote Sensing for Mountain Ecosystems (HyMountEcos) project conducted in June-July 2012. Location of the field campaign site is presented on the overview map (Fig. 1). During this campaign, measurements were performed with lidar, ceilometer, sun-photometers. The whole event was simulated by the DREAM8b, NAAPS and MACC models. These models are briefly described in the Section 2. Section 3 is dedicated to the instruments used during the campaign and different retrieval techniques used to evaluate the lidar and ceilometer data. Section 4 describes the long-term variability of the dust optical properties based on the NAAPS and MACC models and AERONET station in Belsk. In Section 5, the results of the field campaign are described (Holben *et al.* 1998), beginning with temporal evolution of lidar and model results and ending with a comparison of vertical profiles of aerosol extinction obtained from lidar, ceilometer and model simulations.

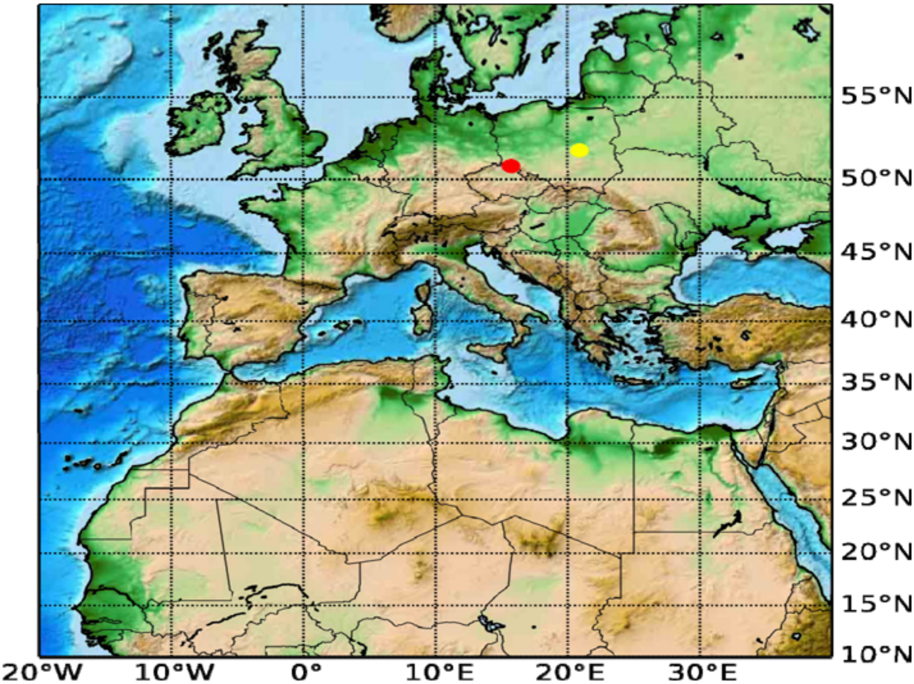


Fig. 1. Overview map with hypsometry and countries boundaries of Western and Central Europe and North Africa. HyMountEcos (Poland) – red circle, Belsk (Poland) AERONET station – yellow circle.

2. AEROSOL TRANSPORT MODELS

To simulate mineral dust optical properties, the Dust Regional Atmospheric Modeling (DREAM8b) (Perez *et al.* 2006a, b) was used. For better time resolution the model was especially run for the HyMountEcos event analysis only, while for the climatological study the data from the public repository of DREAM8b simulation hosted by Barcelona Supercomputing Center was used. To initialize the model for case study description, we used two months of meteorological data prior the period of HyMountEcos campaign. The original DREAM model (Nickovic *et al.* 2001) is a model developed to simulate and predict the atmospheric cycle of a mineral dust aerosol on a regional scale. The model is based on a partial differential nonlinear Eulerian-type equation for a dust mass continuity. Fundamental for all models of atmospheric mineral dust cycle is the parameterization and the conditioning of the dust production phase. In the DREAM the parameterization of aeolian erosion of soil is driven by the soil moisture, the type of soil, type of vegetation, and the atmospheric surface turbulence. As an input for the production components, a global data set on land cover is used with additional data from the Food and Agriculture Organization of the United Nations (FAO) 4 km soil texture data set is required to determine particle size parameters. Grid points from arid and semiarid categories of the global U.S. Geological Survey (USGS) 1 km vegetation data set are treated as potential sources of dust. Particle size distribution is divided into the 8 size bins with the following effective radii: 0.15, 0.25, 0.45, 0.78, 1.3, 2.2, 3.8, 7.1 micrometers. The initial atmospheric and boundary conditions are the 12 UTC 0.5×0.5 degree global National Centers for Environmental Prediction (NCEP) forecast data sets obtained via the Global Forecast System (GFS) model. The 24 h forecast from the day before defines the initial conditions of a dust cycle for the next forecast.

Dust optical depth (DOD) from DREAM8b model is computed from the following equation

$$\tau(\lambda) = \sum_{i=1}^{8} \tau_i(\lambda) = \sum_{i=1}^{8} \frac{3}{4 r_i \rho_i} M_i Q_{\text{ext}}^i(\lambda) \tag{1}$$

where r_i is an effective radius, ρ_i is a particles mass density, M_i is a column mass loading, Q^i_{ext} is an effective extinction cross section for each particle bin. Within each aerosol size bin, dust particles are assumed to have a time-invariant sub-bin lognormal distribution with number median radius of the distribution 0.2986 and geometric standard deviation of 2.0 (Perez *et al.* 2006a, b). The effective extinction cross-section for each particle bin is calculated for spherical particles based on Lorentz–Mie theory. The dust refractive index at 550 nm is assumed to be $1.53 + 0.0055i$ (Hess *et al.* 1998); however, recent studies propose lower values of refractive index, between –0.0005 and –0.0014 (McConnell *et al.* 2008). The Angstrom exponent is computed from Eq. 1 applied for wavelengths of 550 and 1000 nm.

The NAAPS re-analysis model (Witek *et al.* 2007, Zhang *et al.* 2008) is used to predict the spatial distribution of the aerosol concentration and optical properties from 1998 to 2006 and between 2011 and 2012. NAAPS is based on a modification to the model developed by Christensen (1997) with its transition to the Fleet Numerical Meteorology and Oceanography Center (FNMOC). The NAAPS model output is available as 1×1 degree, at 6-hour intervals and 25 sigma-coordinate levels. Model solves the advection-diffusion equation at each grid point for each species. The advection and turbulent mixing is controlled by Navy Operational Global Atmospheric Prediction System (NOGAPS) (Hogan and Rosmond 1991, Hogan and Brody 1993), a dynamic model providing global meteorological fields. Satellite-derived aerosol observations from MODIS assimilated into NAAPS provide estimates of AOD above oceans (Zhang *et al.* 2008). The current version of NAAPS includes gaseous SO_2 and four aerosol components: mineral dust, sea salt, particulate sulphates (SO_4) and smoke. Mineral dust emission areas are characterized by the U.S. Geological Survey (USGS) Land Cover Characteristic Database (Anderson *et al.* 1976). Dust is lifted from the surface whenever the friction velocity exceeds a threshold value (0.6 m/s) and the surface moisture is less than 30%. The employed emission parameterization is proportional to friction wind (Westphal *et al.* 1988). The NAAPS model includes only one size bin for each aerosol type. Aerosol optical properties, such as AOD, single scattering albedo, asymmetry parameter and Angstrom exponent for each aerosol type and as well as for external mixture of particles are computed every 6 hours based on optical interface (Maciszewska *et al.* 2010). NAAPS utilizes a database of global sources individual for each of the simulated aerosol species. Source estimates incorporate weather, remote sensing and anthropogenic activity. For each type of emissions, emission factors are defined, which, for smoke, depend on land use, fuel loading, fuel type and frequency of burns in a particular area; for mineral dust the main factors are: type of soil, area of soil patch and humidity.

The MACC global aerosol transport model consists of ECMWF's Integrated Forecasting System (forward model) and a data-assimilation module (Bellouin *et al.* 2013). The forward modules include 12 prognostic variables (11 aerosol mass mixing ratios and one precursor, SO2). All aerosol species are treated as tracers in the forward model vertical diffusion and convection schemes and are advected by the semi-Lagrangian scheme, consistently with all other dynamical fields and tracers (Morcrette *et al.* 2009). Five types of tropospheric aerosols are included: sea salt, desert dust, organic matter, black carbon and sulphate aerosols. Mineral dust and sea salt are represented by 3 different size classes. Desert dust bins are defined with radii between 0.03-0.55, 0.55-0.9, and 0.9-20 μm, which correspond about 10, 20, and 70% of the total dust mass for each aerosol bins. Emissions of dust particles depend on modelled near-surface wind speeds and dust emission potential which is a function of soil morphology (Ginoux *et al.* 2001). AOD of each aerosol species are computed based on the assumption of external mixture and from standard Lorentz-Mie algorithm (Morcrette *et al.* 2009). Data assimilation module includes the ECMWF four-dimension variation which accounts for background and observational errors. The assimilated observation is the MODIS AOD at 550 nm retrieved over ocean and dark land surface. Aerosols of each type are corrected in proportion of their original contribution to the total aerosol mass (Benedetti *et al.* 2009). In this study we used the MACC re-analysis available for the period between 2003 and 2012.

3. INSTRUMENTATION AND DATA EVALUATION DURING FIELD CAMPAIGN

The measurements for the case study were collected during the first part of the HyMountEcos campaign. This international Polish–Czech project was focused on the assessment of the benefit of hyperspectral techniques for monitoring the highly valuable mountain ecosystems of the Giant Mountains (Karkonosze) National Park. The first part of the field campaign started on 26 June 2012 and finished on 10 July 2012. During the experiment, the mobile laboratory of the Institute of Oceanology, Polish Academy of Sciences (IOPAS) and the Institute of Geophysics, Faculty of Physics, University of Warsaw (IGFUW), equipped with remote sensing, *in situ* and meteorological devices were deployed on the outskirts (about 1 km) of a small town of Karpacz in the Karkonosze in south-western part of Poland. The measurements were made at a field station located 690 m a.s.l. (50.765 °N, 15.757 °E), on the northern side of Sniezka, the highest peak of the Karkonosze Mountains (1602 m a.s.l.). The station was situated over 100 m above the bottom of the valley where Karpacz town is located. The measurement area was situated in a forest clearing, approximately 40 m from the wood areas. The surrounding spruce forest protected the clearing from strong winds. The nearest human settlements, which could cause air pollution, were located about 250 m from the field station; however, both were separated by the ravine of a mountain stream, whose ridges were thickly wooded.

The mobile laboratory was equipped with a LB-10 elastic backscattering lidar operating at 532 nm (Raymetrics, Greece), a CHM-15k ceilometer operated at 1064 nm (JenOptik, Germany), a whole-sky camera, two Microtops II sun-photometers (Solarlight, USA), and a weather station WXT510 (Vaisala, Finland).

3.1 Sun-photometers

In this study, the measurements from the two Microtops II sun-photometers were used. The handheld spectral Microtops II sun-photometers (Morys *et al.* 2001) with visible and near-infrared wavelengths allowed to retrieve aerosol optical depth AOD at 380, 500, 675, 870, and 1020 nm. An important issue in data quality assurance involved the proper calibration of the sun-photometers (Smirnov *et al.* 2000). The calibration factors were derived during different dedicated calibration campaigns in 2012 on Tenerife, Spain, and in Sopot, Poland, as well as in 2011 at Zugspitze, Germany. The spectral dependence of the AOD – the Angstrom exponent is sensitive to the calibration coefficients (Shifrin 1995). In this study, we use the Angstrom exponent defined by AOD at two wavelengths (500 and 1020 nm). Uncertainty of Angstrom parameter decreases significantly with the rise of the AOD value and for the AOD of 0.05, 0.1, and 0.2 (at 500 nm) is about 32, 15, and 8%, respectively (Wagner and Silva 2008, Zawadzka *et al.* 2013).

In addition, data from a CIMEL CE 318 sun photometer (www.cimel.fr) mounted at AERONET station in Belsk (51.836°N, 20.789°E, 180 m a.s.l.) are used. The sun photometers CIMEL are a multi-channel, automatic sun-and-sky scanning radiometers that measure direct solar irradiance and sky radiance at the Earth's surface at seven wavelengths (380, 440, 500, 675, 870, 936, and 1020 nm). The AOD is retrieved at 6 channels and 936 nm channel is used to estimate the total water vapor column. In this study we used the lev. 2.0 data collected between 2002 and 2012.

3.2 Ceilometer and lidar

In principle, JenOptik's CHM15k, similarly to other ceilometers (*e.g.*, Vaisala CT25K), is designed to detect cloud base height (Martucci *et al.* 2010) with the use of lidar technology (O'Connor *et al.* 2004), providing reliable information on clouds up to 15 km. However, significantly higher signal-to-noise ratio than for other ceilometers allows to apply CHM15K to determine the mixing height (Eresmaa *et al.* 2006, Münkel *et al.* 2004, Stachlewska *et al.* 2012) and to examine aerosol profiles up to middle troposphere (Sundström *et al.* 2009, Markowicz *et al.* 2012, McKendry *et al.* 2009, Flentje *et al.* 2010, Frey *et al.* 2010, Heese *et al.* 2010).

The CHM15k uses a diode-pumped Nd-YAG laser at 1064 nm, yielding about 8 μJ per pulse at 5-7 KHz repetition rate (Wiegner and Geiß 2012). The CHM15k receiver consists of 12.7 cm lens telescope directing the backscattered laser light to a silicon avalanche photodiodes (APD) with a photon counter. The divergence of the laser beam is 0.1 mrad. The vertical resolution of the instrument is 15 m. During the field campaign discussed in this paper the temporal averaging was set to 30 s.

The Raymetrics elastic lidar LB-10 is designed to perform continuous measurements of aerosol particles. It is based on the second harmonic frequency of a compact Nd:YAG laser, which emits pulses of 20 mJ output energy at 532 nm with a 20 Hz repetition rate. The laser beam diameter is 10 mm with divergence of less than 0.1 mrad. The optical receiver is a Cassegrainian reflecting telescope with a primary mirror of 20 cm diameter, directly coupled to the lidar signal detection box. Analog detection of the photomultiplier current and single photon counting are combined in one acquisition system. The combination of a powerful A/D converter (12 Bit at 40 MHz) with a 250 MHz fast photon counting system increases substantially the dynamic range of the acquired signal, compared to conventional systems and provides a spatial resolution of 7.5 m. The lidar overlap height was estimated to be between 300 and 400 m based on the visual inspection of the vertical variability or the range corrected signal.

Thanks to data from two laser systems operating on two different wavelengths (532 and 1064 nm) it is possible to detect cases where extinction coefficient is higher at infrared than at visible range, which could indicate coarse particles in the atmosphere, which are characteristic for mineral dust occurrence.

3.3 Other data

The data analyzes were supported with observations from a Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) mounted onboard a Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) satellite (Wong *et al.* 2013). In addition, a Hybrid Single Particle Lagrangian Integrated Trajectory Model (HYSPLIT) (Draxler and Rolph 2010) is used to describe the origin of the air masses.

3.4 Dust properties retrieval techniques

To obtain the vertical profiles of the extinction and backscatter coefficient from lidar and ceilometer signals, the standard Klett–Fernald–Sasano approach was used (Klett 1985, Fernald 1984, Sasano *et al.* 1985). This method requires knowledge or an assumption of the aerosol backscatter coefficient at reference altitude and additional information on aerosol optical properties, such as AOD and/or lidar ratio. In case of lidar and ceilometer data, the standard backward and forward methods (Markowicz *et al.* 2008) were applied, respectively. In the first case we assumed that at the reference altitude is 8 km a.g.l. In the case of forward approaches, the initial aerosol extinction coefficient (at 0.25 km) was assumed to be 0.05 km^{-1} at 1064 nm and initial lidar ratio is set of 35 srad at 1064 nm. The last value is typical for mineral dust particles (Wang *et al.* 2008, Pappalardo *et al.* 2013). However, during the consecutive iterations the initial value varied, so that the final aerosol extinction coefficient at 0.25 km usually differs from the starting value. The same concerns the lidar ratio which is adjusted due to the lidar ratio typical for dust events and additionally verified by AOD constraint calculated from Angstrom exponent for coarse mode, with extinction for 1064 taken from ceilometer retrieval. The assumed AOD was validated by the NAAPS and MACC results and fits inside the simulated range of AOD. In both the backward and forward approach, it was assumed that in the upper troposphere it is only the molecular scattering that contributes to the total backscatter coefficient. Assuming an error of 2% of the molecular backscatter coefficient calculated from the radio sounding data, which is accounted for a daily variation of temperature and pressure, the error of the retrieved aerosol extinction coefficient is about 10% (Stachlewska and Ritter 2010). The errors resulting from the assumption of the overlap correction function (remaining unchanged with time between 250 and 600 m) are less than 3% (Stachlewska *et al.* 2010).

4. ANNUAL VARIABILITY OF THE DUST OPTICAL PROPERTIES

In this section we present results of long-term (1998-2006, 2011, 2012) NAAPS, MACC (2003-2012) re-analysis and DREAM8b v 2.0 re-analysis (2006-2012) as well as the 11-year (2002-2012) observation of the AOD at the AERONET station in Belsk (51.836°N, 20.789°E, 180 m a.s.l.). NAAPS is a global aerosol analysis and prediction model, which operates on $1^{\circ} \times 1^{\circ}$ grid with meteorology from NOGAPS with the same resolution. DREAM8b as a regional model has a domain in the range of 5S to 70N and 60W to 90E with $0.3^{\circ} \times 0.3^{\circ}$ horizontal resolution; meteorology is driven by GFS $0.5^{\circ} \times 0.5^{\circ}$ products. Both models use 6-hour update cycle and data is reported with this time step. MACC is the acronym for the atmospheric composition monitoring project and data published by the project repositories comes from the ensemble of models with different parameters and resolutions (Marécal *et al.* 2015). Resolution and input data for described models show that they are suitable for mesoscale simulations and with too sparse resolution for analyzing of microscale events. Figure 2 shows (top) dust AOD at 550 nm, (middle) dust to total AOD ratio at 550 nm, and (bottom) relative number of days with dust AOD higher than 0.05 for both NAAPS (gray bars), MACC (black bars), and DREAM (white bars). The 0.05 threshold for dust events was selected arbitrarily, to select days with dust contribution to the AOD of about 25%. In case of both models, the dust AOD shows annual cycle with minimum (0.01-0.02) during winter and maximum (about 0.05) during spring and summer. Differences between MACC and NAAPS models are small with the exception of the summer months. During this period the dust AODs in case of MACC are larger (about 0.01) than NAAPS data. In case of DREAM model the dust AOD is significantly smaller, with maximum during May and June. The long-term dust AOD at 550 nm is for NAAPS 0.031, MACC 0.028, and DREAM only 0.005, which correspond to mean dust contribution to the total AOD of 17 and 15% in case of NAAPS and MACC, respectively. In case of NAAPS data the ratio of dust to total AOD has a maximum peak in May (0.28) and a minimum in December (0.1). However, the MACC annual amplitude is smaller and this ratio varies between 0.1 and 0.2. Similarly to dust AOD case during summer months, the MACC shows larger dust contribution to the total AOD than the NAAPS model. In addition, there is a positive difference between NAAPS and MACC ratio of dust to total AOD during spring period. Similar temporal

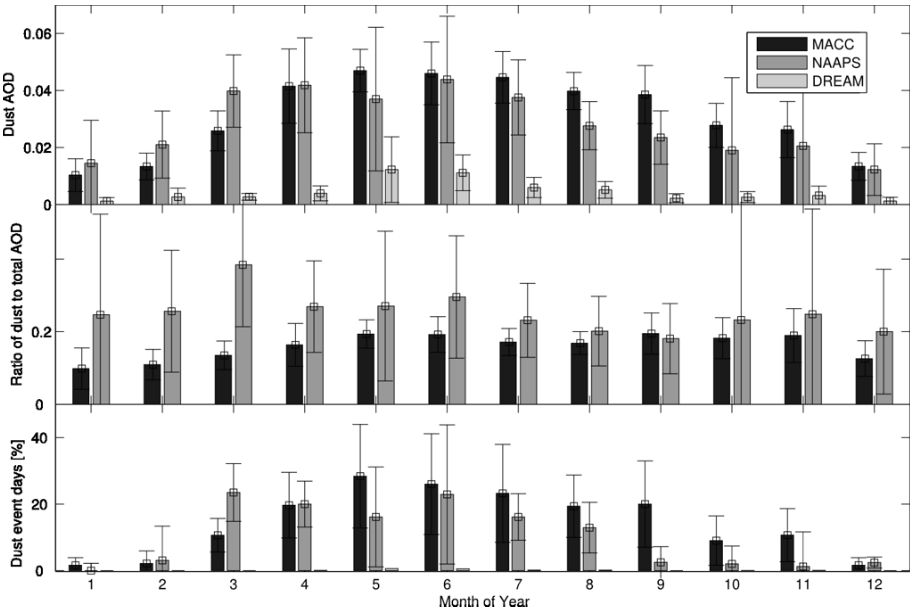


Fig. 2. The long-term monthly mean of dust AOD at 550 nm (top), dust to total AOD (middle), and relative number of days in % with dust AOD larger than 0.05 (bottom) obtained from MACC (black bars), NAAPS (gray bars), and DREAM (white bars) for Poland. Whiskers on bars represent standard deviation.

variation shows the frequency of the dust events. Bottom panel of the Figure 2 shows percentage of days with dust AOD higher than 0.05 (the threshold value was chosen arbitrarily). During May, both models show about 28% of days with dust transport while during winter these values are below 5%. The annual means of these parameters are 12 and 14%, respectively, for MACC and NAAPS model. For threshold of 0.1 the NAAPS models show dust event in May during 9.5% days and only during 4.5% days in case of MACC. In this case, during the whole year, the NAAPS shows about 4% and MACC 2.5% days with a dust event. In simulations from DREAM model the dust events are almost absent, even during the spring maximum.

The largest differences between models appear during summer months and to figure out this discrepancy we did some additional studies. Figure 3 shows the monthly mean dust AOD averaged over Northern Africa (between 30° and 33°N and 10°W and 45°E) based on NAAPS (gray bars) and MACC (black bars). The line with squares corresponds to relative number of days with transport of air masses from Africa to Central Poland. These days are selected using the HYSPLIT back-trajectories ending in Central Poland at 1.5 and 4 km that passed the Northern Africa. Similarly to dust AOD over

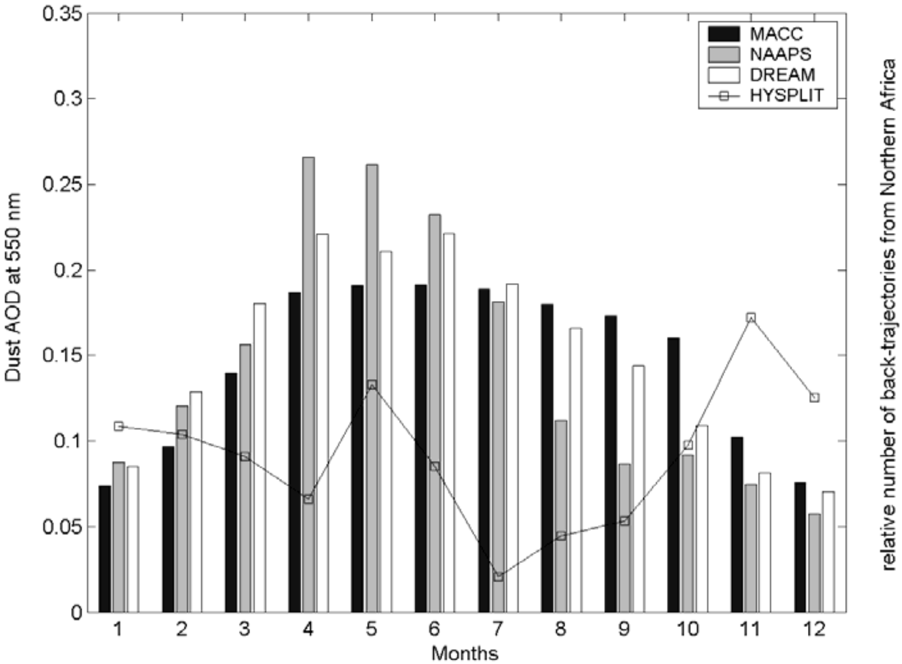


Fig. 3. The monthly mean dust AOD at 550 nm averaged between 30° and 33°N and –10°W and 45°E (North Africa) for MACC (black bars), NAAPS (gray bars), and DREAM (white bars) as well as the relative number of days with back-trajectories crossing North Africa and ending over Central Poland at 12 UTC.

Poland, values for Northern Africa show larger amplitudes in case of NAAPS model. The spring peak exceeds 0.25 at 550 nm while MACC maximum period appears between spring and summer and does not exceed 0.2. For Sahara region the DREAM model shows similar annual cycle as the NAAPS with a significant transition between spring and autumn. NAAPS and DREAM annual cycles are consistent with previous studies which show the maximum of dust emission over Sahel during spring months (Brooks and Legrand 2000, Goudie and Middleton 2001) and late-spring to monsoon onset (Engelstaedter and Washington 2007, Marsham *et al.* 2008). Dust emission is related to the passage of cyclonic fronts in the Northern Sahara in the late winter and spring months and moist convective systems in the Sahel and Southern Sahara in summertime (Alpert and Ziv 1989, Tegen *et al.* 2013). However, more recent results obtained from Multiangle Imaging Spectroradiometer (MISR) by Choobari *et al.* (2014) indicate that AOD over Northern Africa is similar during spring and summer months, which is more consistent with MACC results. However, the back-trajectories show the possible most intensive transport of Saharan air masses to Central Europe in May and November. Both models indicate peak of the dust AOD over Poland in May. During November the intensive transport of Saharan air mass is high but the dust emission is quite low and therefore the dust AOD in Poland is not as large as during May. The minimum of Saharan air mass transport frequency is during summer months (July, August, September) which agrees with reduction of dust AOD simulated by NAAPS. This temporal variation is not predicted by the MACC model which can be explained by data assimilation problems. In case of MACC, the AOD from MODIS is assimilated into the model but each aerosol type is corrected in proportion of its original contribution to the total aerosol mass (Benedetti *et al.* 2009). Therefore, relatively high MACC dust AOD during summer months can be explained by an increase of total AOD during this period. Long-term observation of AOD in Belsk AERONET station shows maximum in July and August. It is important to mention that we used 96-hour back-trajectories simulated by HYSPLIT, which results in a possible loss of data from longer and more complicated dust transport cases. Additionally, due to meteorological situation and processes of deposition, it is possible that not every trajectory from above Africa transports dust to Poland (AERONET (2002-2012)).

More information about model discrepancy can be found from model validation with AERONET observations. Figure 4a and b shows comparison of the total AOD at 550 nm simulated by both models and measured at AERONET in Belsk. In case of Belsk data, the AODs at 55 nm are computed from 500 nm AOD and from Angstrom exponent. In case of MACC model, the mean bias is only 0.02, RMSE is 0.1, and squared correlation co-

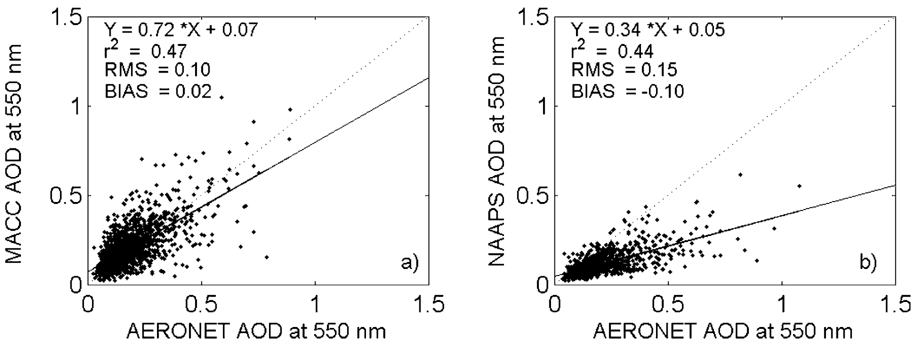


Fig. 4. Comparison of AOD at 550 nm measured in frame of AERONET (lev. 2.0) in Belsk with MACC (a) and NAAPS (b) model. Dotted line corresponds to perfect agreement and solid line is a liner fit.

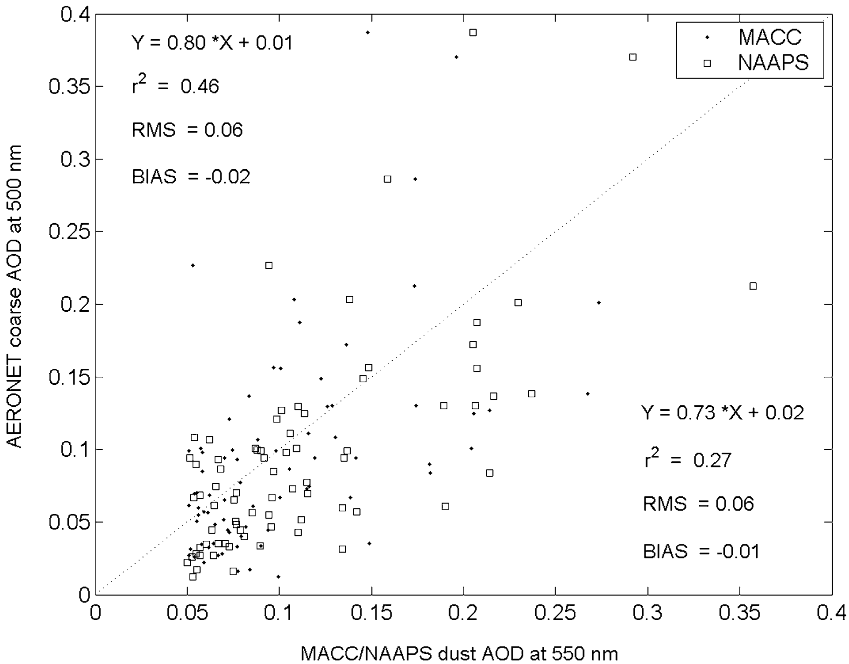


Fig. 5. Comparison of the dust AOD 550 nm obtained from MACC (dots) and NAAPS (squares) models with AERONET AOD of coarse mode at 500 nm retrieved from direct measurements (lev. 2.0). Upper left statistics are for MACC and lower right for NAAPS model.

efficient r^2 is 0.47. The NAAPS model significantly underestimates the AOD (mean bias –0.1); RMSE is 0.15 and r^2 is 0.44. The good agreement for MACC is possibly an effect of AOD assimilation over land. In case of NAAPS model, deficiencies influence simulation accuracy, especially insufficient representation of anthropogenic aerosols and simplified parameterization of aerosol optical properties (only one particle bin size) (Maciszewska *et al.* 2010) as well as the lack of assimilation of the AOD. In case of dust event, defined when the dust AOD for both model is higher than 0.05, we found a better correlation (r^2 is 0.58 and 0.50 for MACC and NAAPS) but a larger bias (0.04 and –0.15 for MACC and NAAPS). Although dust aerosol has better parameterization in the NAAPS than anthropogenic aerosol, we found worst agreements with AERONET data due to the fact that during long-range transport the dust particles are usually mixed with aerosol of other type (Bègue *et al.* 2012; Pavese *et al.* 2012; Papayannis *et al.* 2008). However, the comparison of the AOD obtained from NAAPS and from AERONET coarse mode looks more consistent (Fig. 5). The mean bias is only 0.01 and RMSE is 0.06. Similar results have been found for MACC

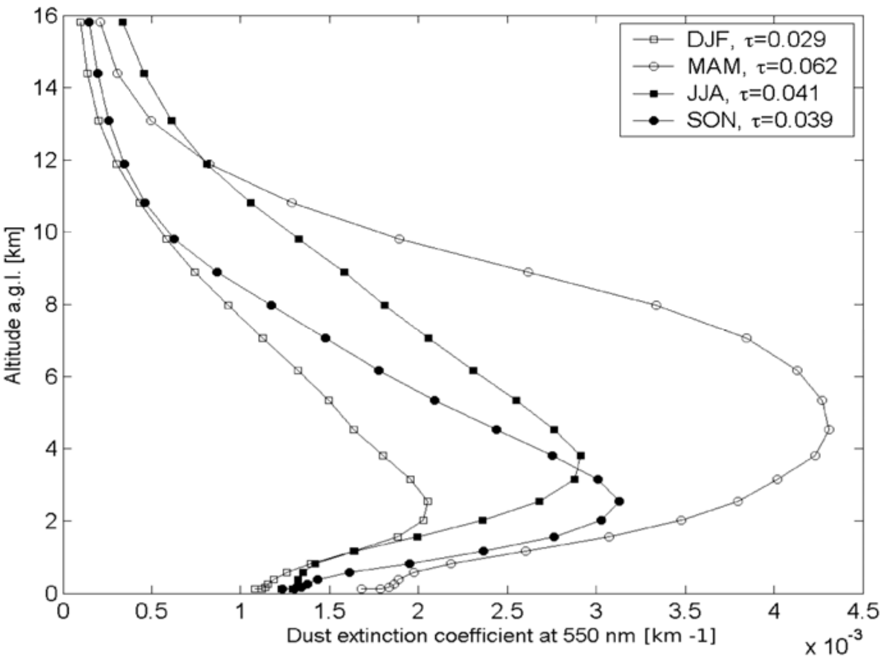


Fig. 6. The mean profiles of dust extinction coefficient obtained from NAAPS re-analysis for December-January-February (open-square line), March-April-May (open-circle line), June-July-August (solid-square line), and September-October-November (solid-circle line).

model. Note that AERONET coarse mode of AOD was estimated for 500 and model AOD at 550 nm. However, in case of coarse mode we cannot expect significant spectral dependence of the AOD. The comparison of the dust AOD by coarse mode AOD is only a simplification because other aerosol species can contribute to it (for example, sea salt or local dust particles, but we do not expect their significant impact on the area of study).

The vertical variability of dust optical properties was obtained only from NAAPS model because the web interface to MACC re-analysis includes only the AOD product. Figure 6 shows the long-term mean of the dust extinction profiles averaged for winter (open squares), spring (open circles), summer (solid squares), and autumn (solid circles). The averaged profile of extinction indicated that dust layers may appear in the upper troposphere. Previous lidar measurements have shown dust mostly in the middle troposphere (Begue *et al.* 2012) but also in the upper troposphere (Ansmann *et al.* 2003). The altitude of maximum dust extinction varies between 2.5 and 3 km in winter and autumn, to 4 km in summer and to 5 km in spring. The maximum of dust extinction coefficient changes between 2 and 4.3×10^{-3} km^{-1}.

Dust is more elevated in summer than in winter, which is in consistence with results of 5-year CALIPSO climatology (Huang *et al.* 2013). After analysis of the whole dataset from three different models we can conclude their performance with the following observations: the NAAPS, although without special assimilation, predicts local emissions well; MACC as an ensemble assimilate data from satellite products, thus predict AODs with high correspondence to observations; DREAM8b is the only model in this set dedicated only to mineral dust transport cycle, which results in the lack of possibility to compare dust AOD to total AOD; simulations of DREAM8b emissions in North Africa fit the results modeled by other models, but after long-range transport simulations to the area of Central Europe/Poland, dust AODs significantly drop, below predictions of other models.

5. RESULTS OF FIELD CAMPAIGN

5.1 Overview of dust event in June 2012

In this section the results of experimental measurements and numerical simulations of a dust event which occurred between 28 and 31 June 2012 are presented. During this period a weak high-pressure system developed over Central and Eastern Europe and an intensive low-pressure system over Great Britain, whereby extensive atmospheric fronts were present. Generally, the synoptic chart predicted south-west circulation and advection of air masses from western Sahara, via the western basin of the Mediterranean Sea to Central Europe.

The series of 5-day backward trajectories generated with the NOAA HYSPLIT model (Fig. 7) indicate that at 18:00 UTC on 28 June 2012 air masses from all the analyzed levels arrived over our site from the Atlantic Ocean through Great Britain. 6 hours later, at 00:00 UTC on 29 June, the situation changed and the calculated source of air masses at the level of 3 and 5 km was in the western Sahara, which suggests a possible Saharan dust transport to Europe. From Africa to the HyMountEcos observational site in Karpacz, the air masses were advected via the Atlantic Ocean and Germany. Twelve hours later, simulation indicated the possible source of mineral dust, mainly on 1 km level, where air masses came from 2.5 km height over the western Sahara. The results from 00:00 UTC on 30 June show the end of potential dust sourcing in Sahara, as all levels had their particle sources across the Atlantic Ocean, from the Iberian Peninsula's coast at 1 km level to the Labrador coast at 5 km level height.

Figure 8 presents simulation of DOD distribution over Central Europe calculated by the MACC, DREAM8b, and NAAPS models. First panel (Fig. 8a) illustrates 00:00 UTC on 29 June 2012. Simulated DOD over Karpacz was MACC 0.19, DREAM8b 0.03, NAAPS 0.11. All models pre-

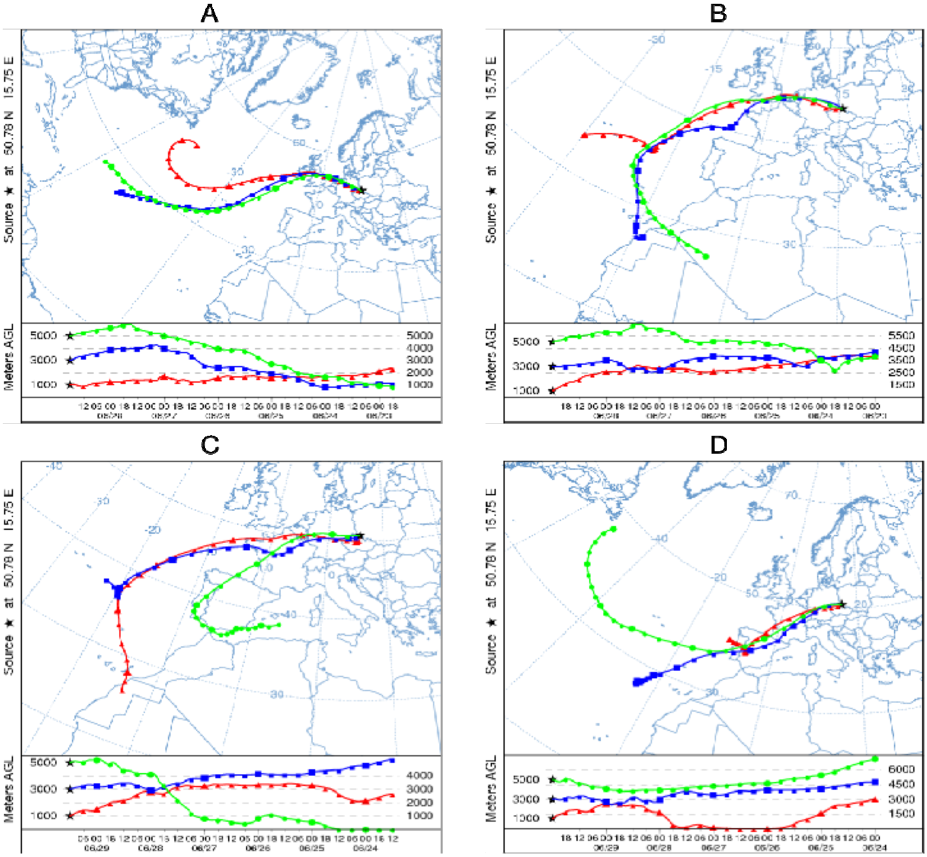


Fig. 7. Hybrid Single Particle Lagrangian Integrated Trajectory Model (HYSPLIT) backward trajectories obtained for Karpacz at: (A) 18 UTC on 28 June 2012, (B) 00 UTC, (C) 12 UTC on 29 June 2012, and (D) 00 UTC on 30 June 2012. HYSPLIT model was running for 5 days with meteorological data from the Global Data Assimilation System (GDAS).

dicted the highest DOD in the western part of Europe. Predicted spatial distribution shows an area of higher DOD coming from the north-west, but the exact location of dust is different among all models. The next panel (Fig. 8b) represents situation at 12:00 UTC on the same day. Models predicted further transport of dust to the east. As on the first panel, the DOD predicted by MACC has highest values (over 0.3 in Germany) and 0.21 over Karpacz. DREAM8b placed Karpacz on an edge of higher DOD with a value of 0.10. NAAPS presents different solution, where Karpacz (0.08) is in the middle between two areas of higher DOD. All models predicted sharp gradient of dust load between Great Britain (low values, 0.0-0.04) and Western Europe

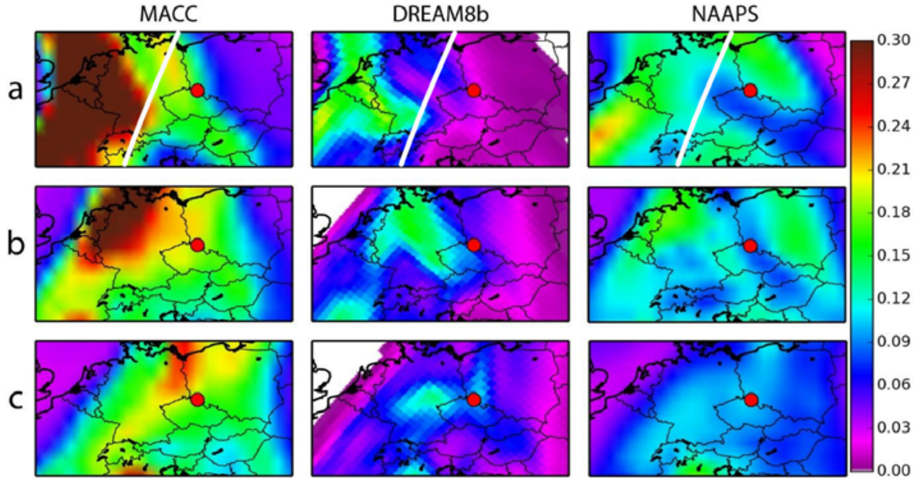


Fig. 8. Mineral dust optical depth at 550 nm between 29 and 30 June 2012 obtained from the MACC, DREAM8b, and NAAPS simulations: (a) 00:00 UTC 29 June 2012, (b) 12:00 UTC 29 June 2012, (c) 00:00 UTC 30 June 2012. For clarity the circle represents the location of the HyMountEcos station in Karpacz. Trajectory of CALYPSO on 29 June 2012 01:30 UTC – white line in panel (a).

(0.15-0.3). The last panel (Fig. 8c) depicts simulations for midnight between 29 and 30 June 2012. All models predicted a decrease of total DOD over Europe and end of dust advection from north-west. MACC simulated highest values of DOD over western borders of Poland, but values in Karpacz are lower than forecasts for previous times. In the DREAM8b and NAAPS simulations, there is a visible decrease of DOD values in most of Europe. However, in the Karpacz area, values are higher than previous time (DREAM8b 0.12, NAAPS 0.11). In the last time step, all three models simulated an area of increased DOD values over North Italy. The source of this increase was the direct transport of dust from Africa to the north, which was independent of the earlier event described above. Complete analysis of predictions by the three models shows that main features of the advection (direction of transport, end of the event), where similar in all models. Deeper comparison reveals discrepancies in spatial and temporal properties of the event. Values of DOD are highest in MACC prediction, reaching 200% of values predicted by NAAPS and DREAM8b. Highest values of DOD predicted by MACC in Central Europe were above 0.30, when NAAPS and DREAM8b proposed ~0.15.

The occurrence of mineral dust over Central Europe was independently confirmed by the CALIOP lidar onboard the CALIPSO satellite. The CALIOP has the ability to collect information on particle linear depolariza-

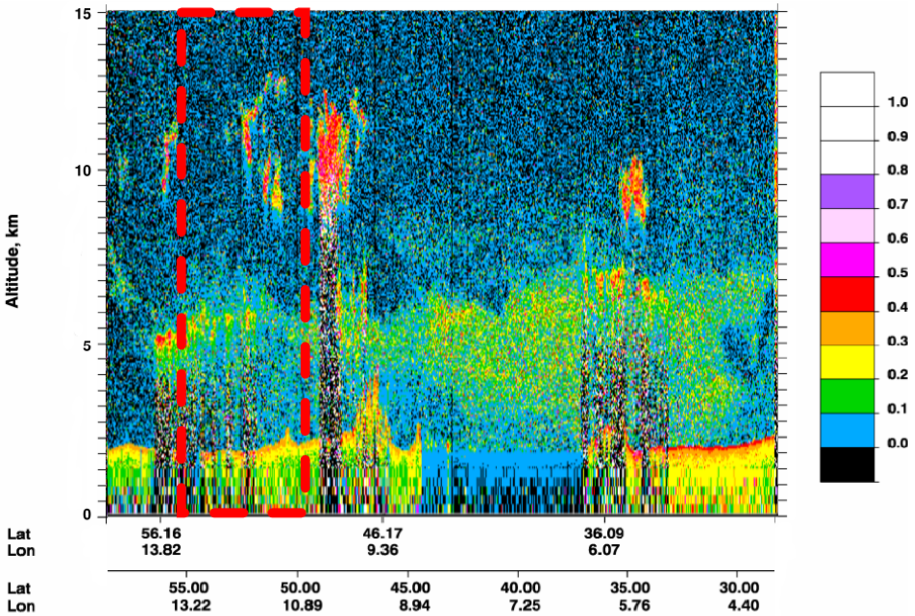


Fig. 9. Transect of particle linear depolarization ratio obtained from CALIOP lidar measurements at 532 nm on 29 June 2012 at 1:30 UTC. Flight path over Central Europe (East Germany). Data from level 2 version 3.02 at 2.5 km horizontal resolution are used here. Central Europe area – red dashed box.

tion ratio at 532 nm, which is helpful in determining the non-sphericity of aerosol particles. On the night of 28/29 June the satellite path crossed Central Europe at 01:30-01:45 UTC. Its overpass was over Central and East Germany, 2.5 degree to the west from Karpacz. The satellite's flight path was not exactly passing over the field station, but according to the HYSPLIT backward trajectories and models simulations, mineral dust was transported over Poland from north-west through Germany. In Fig. 9, illustrating the particle linear depolarization ratio profiles obtained from the CALIOP level 2 data, layers of aerosols with depolarization ratios between 20-30% are visible in the middle troposphere, which is characteristic for Saharan dust (Freudenthaler *et al.* 2009). These layers were observed between 3 and 6 km of altitude, which corresponds with the simulation results (Fig. 6) and seems to be also in accordance with the ground-based observations at the field station near Karpacz (Fig. 12).

5.2 Temporal variability of dust event over Karpacz

Description of temporal changes of the lidar range corrected signal is based on the time composition depicted in Fig. 10a. During the described event, the

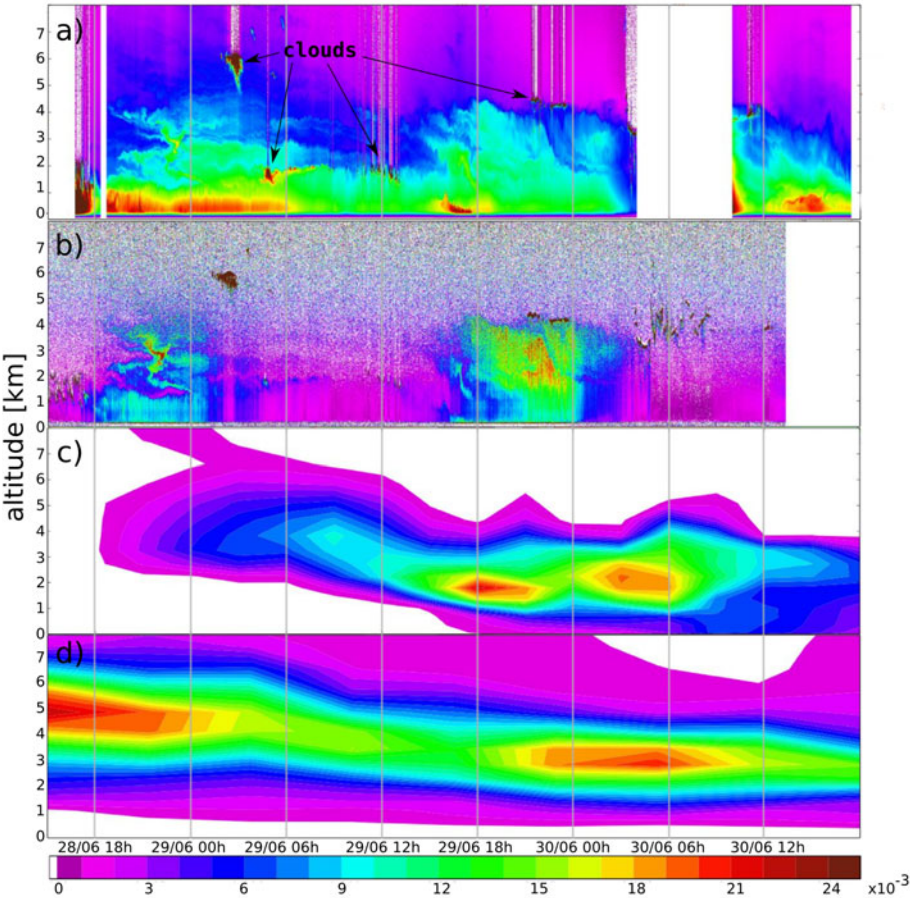


Fig. 10. Temporal variability of vertical structure of the range corrected signal of lidar (a), ceilometer (b), and dust aerosol extinction coefficient [km^{-1}] obtained from the DREAM8b (c) and NAAPS (d) model simulations from 15:00 UTC 28 June 2012 to 18:00 UTC 30 June 2012.

signal above boundary layer (PBL) was mainly dependent on the advected dust concentrations. Lidar measurements with spatial resolution of 7.5 m and time resolution of 1 min provided the most accurate information on this event variability. The first measurements which could be attributed to a dust layer were made before 19:00 UTC on 28 June 2012 at the level of 4 km, and then just before 20:00 UTC an additional layer at 2 km appeared. The intensity of the event has been increasing until midnight. During that time, the observations allowed to distinguish a complicated multilayer structure of dust cloud between 1.5 and 4.5 km. The highest values of the range corrected signal were observed exactly at midnight just below 3 km. After the midnight, the multilayer mixed itself, creating a new layer in the span between 1.5 and 3 km. The high values in the PBL are probably marginally dependent on the dust event (compare with Fig. 10c-d), taking their source most probably from local, likely anthropogenic (coal, oil heating) emissions. This was predicted by NAAPS, which simulated extinction from non-dust particles on the lowest levels. At 06:00 UTC on 29 June 2012, low clouds appeared at 1.5 km level, what made lidar measurements of aerosol particles more difficult to evaluate. The following evening, clouds disappeared at around 18:00 and thus the observations of dust load could be continued. At midnight of 29/30 June 2012 the dust layers were thicker, spanning from 1.2 to 4.5 km, with the main load between 1.5 and 3.5 km. The event was intense during the night of 29/30 June, with a significant decrease of values after 01:30 UTC on 30 June. Note that the gap in lidar measurements between 08:00 and 14:00 on 30 June 2012, is due to a problem with data acquisition. After restart of lidar observations, the dust was still present up to a level of 4 km. Similar vertical structure of aerosol layers has been shown by ceilometer data (Fig. 10b). In this case, a strong range corrected signal appears during nights of 28/29 and 29/30 June. The dust layers between 2 and 4 km are better visible (Fig. 10b) in comparison to lidar data due to larger sensitivity to coarse mode particles and due to neglected Rayleigh scattering. Note that during the daytime the ceilometer has limited sensitivity due to a strong influence of background light on the ceilometer signals.

Figure 10c represents the simulation of the dust event calculated with the use of the DREAM8b model. Extinction coefficient on 24 levels from 0 to 15 km, in intervals of 3 h, was calculated. The model predicted the beginning of the event at 20:00 UTC on 28 June 2012 with an aerosol extinction coefficient of 0.003 km^{-1} at 2.5-5.5 km level. Then, according to simulation, the dust lowered its altitude, reaching the ground at 16:00 on 29 June 2012. The highest peak of aerosol extinction coefficient (more than 0.025 km^{-1}) was simulated from 16:00 to 23:00 UTC on 29 June 2012. The lidar showed high range corrected signal around 19:00 UTC, which generally agreed with the simulation, but was 3 h later than predicted. Afterwards, the simulated event started to weaken. The simulation, in comparison to the lidar measurements, showed an appearance of the dust event later with the clearly visible main load at an altitude of around 2 km, while the lidar detected dust between 0.5 and 4 km, without significant maximum at one altitude.

Figure 10d depicts predictions of extinction coefficient by NAAPS model. The event simulated on this panel started earlier and has local maxima at 18:00 UTC on 28 June with values of extinction coefficient above 0.023 km^{-1} at 5 km altitude. In comparison with observations (Fig. 10a-b) the

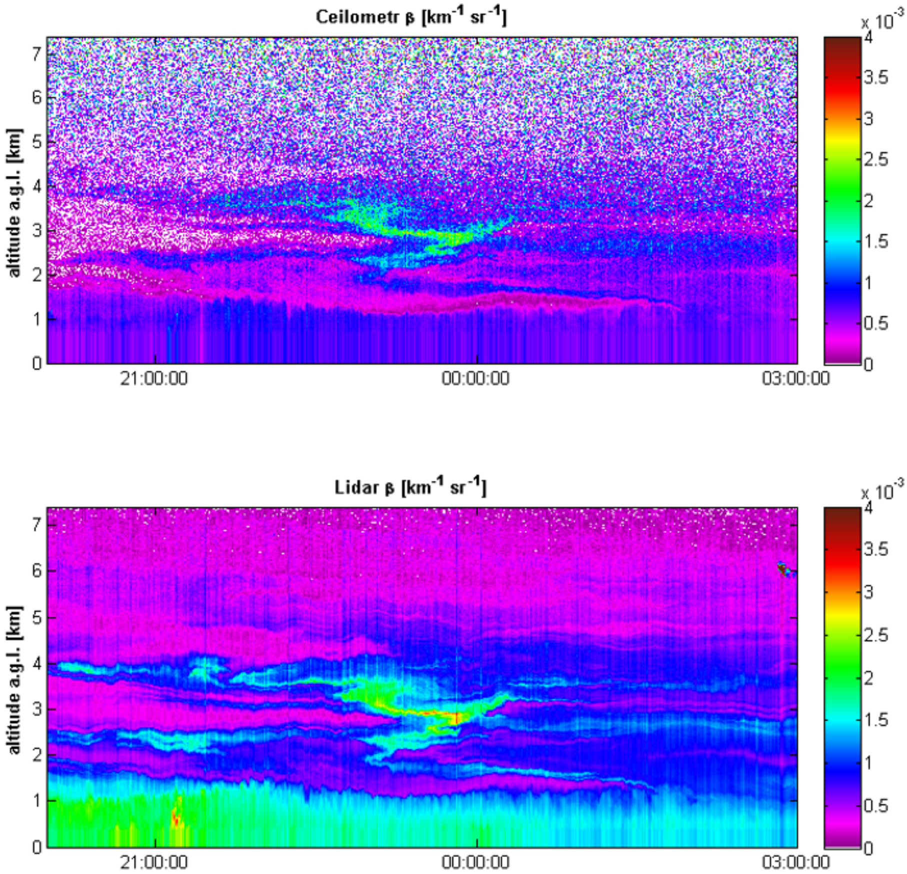


Fig. 11. Aerosol backscatter coefficient retrieved from: (top) ceilometer (1064 nm) and (bottom) lidar (532 nm) observation between 20 UTC 28 June and 2:30 UTC 29 June 2012.

model predicted dust on higher altitude and earlier. Higher values around midnight on 28/29 June better correspond to observations than DREAM8b simulation, which underestimates the intensity of early stages of the dust event. Between 21:00 UTC 29 June and 04:00 UTC 30 June, NAAPS simulated second maxima with values slightly higher than 0.020 km^{-1} between 2 km and 3 km altitude. These maxima correspond with the one simulated by DREAM8b model earlier in a quite similar range. In the whole simulated period, a slow decrease of dust altitude is visible, but it never reached the lowest level.

It is important to remember that the two models operate on a different grid, DREAM8b at 0.33°×0.33° and NAAPS at much sparser grid of 1°×1°;

keeping this in mind, the results from NAAPS show quite accurate prediction of the dust event dynamic with two maxima, but with overestimated altitude of the main dust layer; on the other hand, layers simulated by DREAM8b were on altitudes more similar to measurements, but the maximum at the beginning of the event has not been predicted. Extinction coefficient simulated by DREAM8b and NAAPS reach almost the same values at peaks ~0.022, but NAAPS simulated higher total load with background in the range of 1-7 km of altitude. Temporal variability of results obtained from the lidar measurements clearly shows efficiency of the lidar measurements in determining local diversity of multilayered structure of dust events (Papayannis *et al.* 2007, Guerrero-Rascado *et al.* 2009, Preißler *et al.* 2011).

The temporal evolution of the vertical structure of the retrieved aerosol backscatter coefficients from lidar and ceilometer data is shown in Fig. 11. The upper panel corresponds to the ceilometer backscatter and the bottom one to the lidar backscatter coefficient from 20:00 UTC 28 June 2012 to 02:30 UTC 29 June 2012. We selected this section of measurements for further analysis, because it reveals the most interesting feature of the event. The multilayer of dust particles spans between 2 and 4 km a.g.l. The altitude of almost each aerosol layer decreased with time, which was probably due to air subsidence, advection of the dust to north-west Poland or/and particle sedimentation. Both plots show similar structure of aerosol layers below 4.5 km, especially between 23:00 and 00:30 UTC, when the backscatter coefficient reached the highest values (3.5×10^{-3} – 3.75×10^{-3} km^{-1} sr^{-1} at 532 nm and 2.5×10^{-3} – 2.75×10^{-3} km^{-1} sr^{-1} at 1064 nm).

5.3 Comparison of selected profiles of extinction obtained by model and remote sensors

The comparison of the aerosol extinction coefficient obtained from the lidar and ceilometer measurements and predicted by the DREAM8b and NAAPS models was based on six vertical profiles retrieved during different stages of the dust event. The first three profiles were obtained during the first phase of the dust influx on the night of 28/29 June 2012, and the following three during the event culmination on the night of 29/30 June 2012. The night measurements were selected to avoid problems with high background light affecting ceilometer signals. The aerosol extinction profiles for lidar and ceilometer were calculated with the application of the Klett–Fernald–Sasano algorithm with the AOD calculated as a mean AOD from the NAAPS and MACC simulations for the corresponding corrected wavelengths. Area of uncertainty was defined by standard deviation of the calculated mean AOD. Between simulations time-steps, values were linear interpolated. Profiles were filtered with the Savitzky–Golay algorithm (polynomial of grade 3, data window of 9 measurements). In both cases, a 15 min approximation (±7.5 min) was made, from the hour for which the profile was calculated. Further analyses of obtained profiles are focused on range above PBL (1-1.5 km) due to different overlap function for lidar and ceilometer.

In Fig. 12a, profiles for 28 June 2012 at 22:00 UTC show an extremely low dust load predicted by DREAM8b. On the contrary, the NAAPS model shows a total extinction of 0.052 km^{-1} (0.049 km^{-1} dust) with a maximum at 4.3 km altitude. It is significant that NAAPS predicted increased total extinction coefficient (with a lack of dust) below the PBL which agrees with ob-

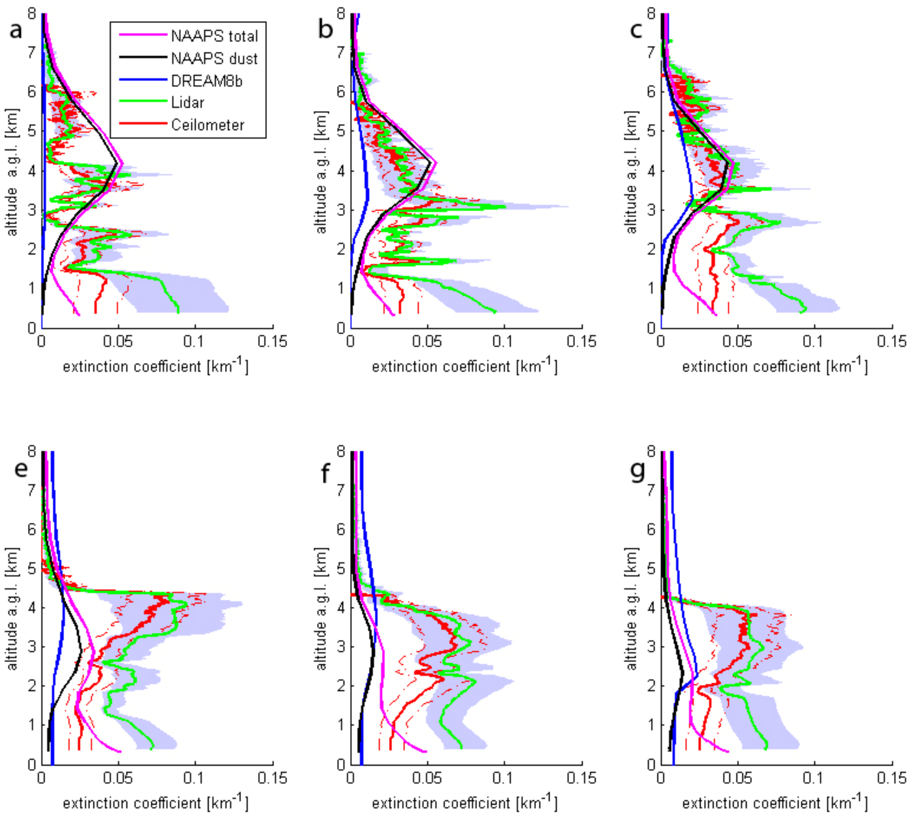


Fig. 12. Aerosol extinction coefficient profiles obtained with the means of lidar at 532 nm (green line), ceilometer at 1064 nm (red line), the DREAM8b model at 550 nm (blue line), the NAAPS model at 550 nm dust extinction coefficient (black line) and NAAPS model total extinction coefficient (magenta line) at 22:00 UTC 28 June (a), 00:30 UTC (b), 02:00 UTC (c), and 21:00 UTC 29 June (d), 00:00 UTC (e), and 04:00 UTC 30 June (f). Gray color and red dotted line represent the area of uncertainty for lidar/ceilometer.

servations based on range corrected signal (Fig. 10). The aerosol extinction profiles, based on both remote sensing measurements, depict some aerosol layers at the range of 1.5-6 km (max 0.058±0.025 km^{-1} at 532 nm and 0.048 ± 0.025 km^{-1} at 1064 nm). The structure presented three layers: 1.5-2.6 km, 3-4.2 km, and 4.8-6 km. Starting from 3.25 km the results from lidar and ceilometer present high similarity of a structure; however, in the upper layers the increase of extinction coefficient obtained from lidar measurement, along with the simultaneous decrease of the extinction coefficient for the ceilometer measurements, may be noticed. This might be caused by the low signal-to-noise ratio of the ceilometer beyond 5 km altitude and by the aerosol advection, which seem to be confirmed by the increase of the aerosol extinction coefficient of lidar at 5-6 km range, however, does not exceed 0.025±0.07 km^{-1}, which is lower by almost a half than at the lower layers at this profile. Differences in profiles below 2 km are caused by different overlap function of both systems and higher number of smaller particles in PBL layer, which are better detected by the shorter wavelength system.

The plot in Fig. 12b depicts results at 00:30 UTC (2.5 h later). Here the first clearly marked increase in the aerosol extinction coefficient value was simulated by the DREAM8b model. The main load of dust was predicted between 2.2 and 5.5 km, with a flat maximum at about 3.4 km and the aerosol extinction coefficient value of about 0.011 km^{-1}. NAAPS predictions were the same in shape as before, with only small decrease of values, 0.055 km^{-1} total extinction and 0.051 km^{-1} dust component. The measurements confirmed the existence of the maximum a little lower, at heights between 3 and 3.3 km, but with substantially higher values of about 0.06 km^{-1} ± 0.023 for ceilometer and 0.1 km^{-1} ± 0.035 for lidar. The course of vertical variability of the extinction coefficient is similar for remote measurements and highlights a few clear local maxima, *e.g.*, at: 1.75, 2.8, and 3.25 km. The high extinction values, confirming the expected transport of the Saharan dust, extend from 1.5 to 5.5 km. The retrieved profile has the highest values of all profiles retrieved during the night of 28/29 June 2012.

In the profile retrieved for the same day at 02:00 UTC (Fig. 12c), 1.5 h later, the maximum values decreased and the vertical structure of aerosol simplified. Values simulated by the DREAM8b model reached the level of 0.021 km^{-1}, remaining at practically the same heights with the maximum at about 3.4 km. NAAPS simulations preserve the earlier shape of the profile, but extinction coefficient decreased to 0.047 km^{-1}. The remote sensors profiles show higher aerosol extinction coefficient than predicted by DREAM8b model, for the entire profile, up to the height of 6.5 km, with two local maxima, at 2.75 and 3.5 km. Results retrieved from remote sensing above 3.5 km fit the values simulated by NAAPS model. The observed remote profiles allow to discern two thick, homogeneous aerosol layers, divided at a height of 3.1 km by the decrease of aerosol extinction coefficient to a one third (0.025 km^{-1} ± 0.02) of the for the maximum level (0.078 km^{-1} ± 0.023). Nearly similar values of aerosol extinction coefficient for devices operating on different wavelength (lidar – 532 nm, ceilometer – 1064 nm) retrieved during the night of 28/29 June 2012 are characteristic for large particles like the mineral dust.

The following night, on 29/30 June 2012, depicted in Fig. 12d-f, was characterized by substantially higher observed aerosol extinction coefficient values. During this night, the DREAM8b model at 21:00 UTC (Fig. 12d) predicted dust from 2 to 7 km, with a maximum of about 0.015 km^{-1} at about 4.2 km. NAAPS model maximum of 0.035 km^{-1} lowered the altitude and was predicted just below 3 km. The measurements correspond to the DREAM8b simulation, with clearly marked maximum values at about 4.3 km with value of about 0.093 km^{-1} ± 0.03 for lidar and 0.083 km^{-1} ± 0.025 for ceilometer. Both profiles show the upper limit of the dust layer at a height of 4.5 km, above which the signal strength decreased to the noise level. Both profiles are corresponding well to each other, with values almost steadily increasing with the height up to the top of the layer.

The profiles at midnight of 29/30 June 2012 (Fig. 12e) show a larger discrepancy between the modeled and the measured values. In comparison with previous profile (Fig. 12d) the DREAM8b model simulated a steady increase in the aerosol extinction coefficient, with the maximum at the level of 3.8 km and span over 1.8-6.5 km, with maximum value of about 0.019 km^{-1}. Values simulated by NAAPS decreased to 0.021 km^{-1} at the same height as before. For the first time in the analyzed profiles both models show a region of great coherence of simulated dust extinction up to 3 km altitude. The measurements show a decrease of the aerosol extinction coefficient values which have reach the values of about 0.084 km^{-1} ± 0.020 (532 nm) at about 2.1 km and between 2.9-3.7 km and 0.079 km^{-1} ± 0.032 (532 nm) at about 2.5 km. Thus, the aerosol layer is not that uniform as previously and no longer has a clear upper limit. Above 4.7 km it vanishes. As before, the lidar and ceilometer measurements are corresponding above the lowermost 1.5 km well to each other, indicating the main dust load at the range of 2.75-3.75 km, with additional thin layers at 2.25 and 2.5 km. This layers had probably separated from the existing increased aerosol load beyond 2.5 km.

Finally, Fig. 12f shows the data from 04:00 UTC on 30 June 2012. Here the DREAM8b predicted lowering of the dust maximum in terms of its height down to about 2.3 km as well as in terms of the aerosol extinction value as this height range reaches 0.024 km^{-1}. From this point, the values steadily decreased until they reached zero at approximately 6.5 km. NAAPS model simulated lower total extinction with flat maximum of 0.020 km^{-1} at the same altitude as DREAM8b. Shape of both simulated profiles shows great discrepancy. The profile from lidar reveals rather flat maxima at about 3 and 3.8 with the values of 0.065 km^{-1} ± 0.028 and a clear decrease of layer at 4 km. Above, only marginal aerosol load remains up to 6.5 km. The top border of the main dust layer is well-marked. The shape of the ceilometer profile is similar to the one in Fig. 12d, with a steady extinction increase with height. The lidar profile has a more similar structure to profiles in Fig. 12e, with an additionally separated thin layer at about 2.1 km. This layer is probably the same as the one at 2.25 km in Fig. 12e.

The comparison of vertical variability of the extinction coefficient among the lidar, the ceilometer, the DREAM8b model and the NAAPS model measurements allows for the following conclusions: the models present very simplified profiles. Extinction values predicted by DREAM8b model are substantially lower than the observed ones. NAAPS model simulated values closer to the observed ones. In most cases, the DREAM8b model correctly predicted the level of the dust transport, together with the location of its maximum value; however, at times the maximum values were substantially lower, as compared to the observed ones, which has already been observed in other studies (Mona *et al.* 2014). Profiles calculated by NAAPS model show limited coherence with profiles retrieved from remote sensing methods. The profiles based on the remote sensing measurements show great coherence with the vertical variability profiles made with the aid of lidar and ceilometer. Possible errors of retrieved extinction coefficient come from the inevitability of subjectively made assumptions about the initial values, resultant from the lack of remote sensing measurements of the complete optical thickness, which was caused by high clouds covering the sun. Thus, it was impossible to make measurements with a sun-photometer.

5.4 Comparison of the observed and simulated AOD

Based on the retrieved profiles of aerosol extinction coefficient from lidar we estimated the total AOD and the AOD above the PBL. To compute the AOD above the PBL, we integrated each profile of the aerosol extinction coefficient in the range above boundary layer height, which was in general between about 1.5 and 7.5 km. The PBL height was estimated based on the maximum gradient method applied to the range corrected signal (Stachlewska *et al.* 2012). Below the overlap height the aerosol extinction coefficient was assumed constant. Due to very small value of particle linear depolarization ratio from CALIPSO and low concentration of dust simulated by DREAM8b model below 1.5 km, the AOD above the PBL can be interpreted, in the first approximation, as the dust AOD. Figure 13a shows the

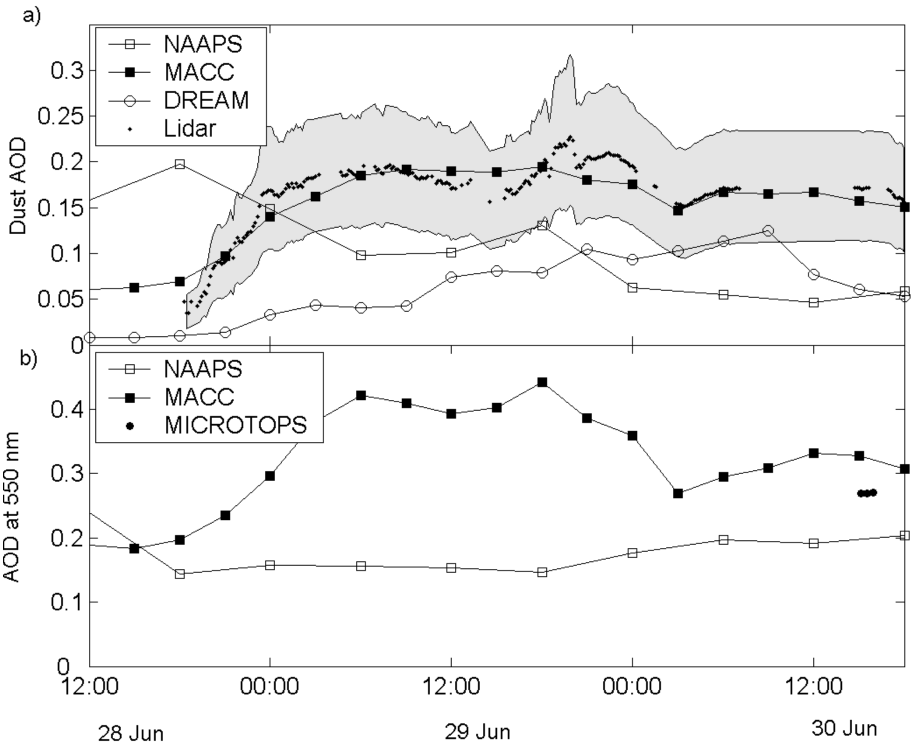


Fig. 13. Temporal variability of: (a) dust AOD and (b) total AOD obtained from the NAAPS (open square line), MACC (solid squared line), DREAM8b model (open circle lines), lidar observation (black dots); and Microtops (dots). Gray color on case of lidar data represents the area of uncertainty.

AOD above the PBL obtained from the NAAPS (line with open squares), MACC (line with solid squares), DREAM8b model (line with circles), and from the lidar (shading). Differences between models and lidar are significant. For example, NAAPS model shows the peak of dust event on 28 June late afternoon while DREAM before noontime of 30 June. The MACC model simulated the beginning of the dust event at midnight 28/29 June and maximum of the dust AOD about 0.2. The lidar AOD above the PBL starts an increase just before the midnight 28/29 similar to MACC model. MACC simulation very well corresponds with lidar measurements reproducing almost the same temporal evolution and values of dust AOD. However, the uncertainty of the dust AOD obtained from lidar measurements is quite large due to high uncertainty of the total AOD applied to Klett–Fernald–Sasano method. Comparison of the total AOD retrieved from NAAPS (line with open squares), MACC (line with solid squares), and the Microtops sun-photometer (dots) (Fig. 13b) reveals similar differences. The AOD from MACC is systematically higher than from NAAPS and a maximum of AOD difference exceeds 0.2 on 29 June. Unfortunately, due to regular occurrence of cirrus clouds, only 3 sun-photometer daytime observations during this dust event were possible. Comparison of the sun-photometer AOD can be done only between 15 and 16 UTC on 30 June. During this period, the NAAPS underestimated the observation AOD by 0.06 and MACC overestimated by 0.04. In case of the Angstrom exponent the NAAPS, MACC, and Microtops values are the following: 0.39, 0.59, and 1.23, respectively. Although the dust AOD contribution to the total AOD is about 50% for MACC and only 25% for NAAPS, the Angstrom exponent estimated from MACC model is significantly higher than from NAAPS. It means that the size distribution of dust particles is dominated by fine mode. In case of DREAM8b we estimated the dust Angstrom exponent of 0.28 which also indicates reduction of large dust particles, in consistence with dust transport. HYSPLIT analysis (Fig. 7) has shown several-day long-range transport trough Western Europe, which enabled a significant part of the coarse-mode dust to be removed.

6. SUMMARY AND FINAL CONCLUSION

In this paper we attempted to provide seasonal variability of dust AOD based on NAAPS and MACC re-analysis. The models are able to reproduce with reasonable skill the observed long term seasonal mean AOD. The comparison with CIMEL sun-photometer measurements shows that the MACC model tends to slighty overestimate the AOD (mean bias is 0.02) and NAAPS to significantly underestimate the AOD (mean bias is 0.1). Similar comparisons for long-term monthly mean value indicate that NAAPS underestimates the CIMEL value by 0.05, especially during spring season (about 30%). For the MACC model we found an overestimation of the AOD in May and June (25-30%) and an underestimation during winter (30-45%). Although the mean AOD bias for MACC is smaller than for the NAAPS model, MACC shows significant inconsistencies with observations during winter. It is probably because of the data assimilation procedures which include the AOD from MODIS observations. The MODIS AOD retrieval is limited only to days with clear sky, which during winter season are usually related to low temperature and high local emissions, which result in smog conditions and high values of AOD; thus in the case of small number of those days, the MODIS assimilation to the MACC can produce some bias. Results from both models are consistent with estimated annual cycles of dust AOD except for the summer months. Simulated dust AOD shows annual cycle with minimum (0.01-0.02) during winter and maximum (about 0.05) during spring and summer. During summer, the dust AODs in case of MACC are about 0.04, while the NAAPS value is about 0.02. Similarly to dust AOD case during summer months, the MACC shows larger dust contribution to the total AOD than NAAPS model. Generally, the MACC dust contribution during spring and summer months is almost flat, while the NAAPS model shows peak value in May.

The annual mean of dust AOD estimated from AERONET and model data is 0.038 ± 0.016. The monthly mean dust AOD has a maximum in April (0.057 ± 0.03) and minimum in December and January 0.027 ± 0.01. The relative number of days with dust event (dust AOD larger than 0.05) is about 28% in May and below 5% in winter. Significant dust events (dust AOD above 0.1) appear during about 9.5% (NAAPS) and 4.5% (MACC) days in May.

The second phase of model validation was based on the HyMountEcos campaign which took place in June 2012. Aerosol optical properties measured by remote sensing instruments and simulated by NAAPS, MACC, and DREAM8b model during the Saharan dust transport were compared. Major correspondences were found in aerosol extinction and backscatter coefficient profiles (shape and values) retrieved from the lidar and ceilometer data. The differences observed in some parts of profiles are characteristic for mineral dust events where larger particles are expected, which results in a relatively high signal at 1064 nm in comparison with 532 nm. Discrepancies in profiles below PBL have their source in different overlap function of both systems and possible appearance of smaller particles from local emissions. The ceilometer detected dust layer above the boundary layer up to 6 km a.g.l. during the night. The comparison of vertical variability of the aerosol extinction coefficient among lidar/ceilometer with the NAAPS and DREAM8b model shows moderate agreement. Differences in values and altitudes of maxima of aerosol backscatter coefficient simulated by the models in comparison to lidar/ceilometer data were found, but the main character of the event was preserved. It is likely due to the model simplifications in dust emissions, deposition and advection parameterization, which are related to the spatial and temporal resolution of the models and often limit their simulation capabilities for processes of larger scale. Although the DREAM8B model includes 8 size bins and NAAPS only one, the profiles of aerosol extinction from NAAPS are more consistent with lidar data. Generally, both the extinction coefficient and dust AOD for DREAM8b data values are smaller than NAAPS data. However, the uncertainty of the dust AOD obtained from Klett–Ferrnald–Sasano method applied to lidar data is large, the results agree quite well with the MACC model. Measurements done with the lidar delivered detailed information on temporal evolution of the dust event with especially interesting data representing multilayered structures of mineral dust transported by the airmass. Thanks to transport of the dust above PBL it was possible to calculate estimated dust AOD from lidar measurements and compare it with values simulated by models.

Acknowledgements. This research was supported with funding of the National Grants No. 1283/B/P01/2010/38 and No. 1276/B/P01/2010/38 of the Ministry of Science and Higher Education of Poland, both coordinated by the IGF UW. We kindly acknowledge the NASA Langley Research Center Atmospheric Science Data Center for the provision of the CALIPSO products used in this study. This work has been supported in part by the European Community under the contract FP7/2007-2013 (MACC project). We acknowledge Brent Holben for the use of the data from the Belsk AERONET station and P. Flatau from Naval Research Laboratory in Monterey for providing NAAPS data.

References

AERONET (2002-2012), AERONET climatology, level 2.0 – quality assured data: Belsk, Aerosol Robotic Network, available from: http://aeronet.gsfc.nasa.gov/new_web/V2/climo_new/Belsk_500.html.

Alpert, P., and B. Ziv (1989), The Sharav cyclone: observations and some theoretical considerations, *J. Geophys. Res.* **94**, D15, 18495-18514, DOI: 10.1029/JD094iD15p18495.

Anderson, J.R., E.E. Hardy, J.T. Roach, and R.E. Witmer (1976), *A Land Use and Land Cover Classification System for Use with Remote Sensor Data*, U.S. Geol. Surv. Prof. Paper, Vol. 964, 28 pp.

Ansmann, A., J. Bösenberg, A. Chaikovsky, A. Comerón, S. Eckhardt, R. Eixmann, V. Freudenthaler, P. Ginoux, L. Komguem, H. Linné, Á.L. Márquez, V. Matthias, I. Mattis, V. Mitev, D. Müller, S. Music, S. Nickovic, J. Pelon, L. Sauvage, P.Sobolewsky, M.K. Srivastava, A. Stohl, O. Torres, G. Vaughan, U. Wandinger, and M. Wiegner (2003), Long-range transport of Saharan dust to northern Europe: The 11-16 October 2001 outbreak observed with EARLINET, *J. Geophys. Res.* **108**, D24, 4783, DOI: 10.1029/2003JD003757.

Balkanski, Y., M. Schulz, T. Claquin, and S. Guibert (2007), Reevaluation of mineral aerosol radiative forcings suggests a better agreement with satellite and AERONET data, *Atmos. Chem. Phys.* **7**, 81-95, DOI: 10.5194/acp-7-81-2007.

Bègue, N., P. Tulet, J.-P. Chaboureau, G. Roberts, L. Gomes, and M. Mallet (2012), Long-range transport of Saharan dust over northwestern Europe during

EUCAARI 2008 campaign: Evolution of dust optical properties by scavenging, *J. Geophys. Res*. **117**, D17, D17201, DOI: 10.1029/2012JD017611.

Bellouin, N., J. Quaas, J.-J. Morcrette, and O. Boucher (2013), Estimates of aerosol radiative forcing from the MACC re-analysis, *Atmos. Chem. Phys*. **13**, 4, 2045-2062, DOI: 10.5194/acp-13-2045-2013.

Benedetti, A., J.-J. Morcrette, O. Boucher, A. Dethof, R.J. Engelen, M. Fisher, H. Flentje, N. Huneeus, L. Jones, J.W. Kaiser, S. Kinne, A. Mangold, M. Razinger, A.J. Simmons, and M. Suttie (2009), Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: 2. Data assimilation, *J Geophys. Res.* **114**, D13, D13205, DOI: 10.1029/2008JD011115.

Brooks, N., and M. Legrand (2000), Dust variability over northern Africa and rainfall in the Sahel. **In:** S.J. Mc Larsen and D. Kiiverton (eds.), *Linking Land Surface Change to Climate Change*, Kluwer.

Chand, D., R. Wood, T.L. Anderson, S.K. Satheesh, and R.J. Charlson (2009), Satellite-derived direct radiative effect of aerosols dependent on cloud cover, *Nat. Geosci*. **2**, 3, 181-184, DOI: 10.1038/NGEO437.

Chen, L., G. Shi, S. Qin, S. Yang, and P. Zhang (2011), Direct radiative forcing of anthropogenic aerosols over oceans from satellite observations, *Adv. Atmos. Sci*. **28**, 4, 973-984, DOI: 10.1007/s00376-010-9210-4.

Choobari, O.A., P. Zawar-Reza, and A. Sturman (2014), The global distribution of mineral dust and its impacts on the climate system: A review, *Atmos. Res.* **138**, 152-165, DOI: 10.1016/j.atmosres.2013.11.007.

Christensen, J.H. (1997), The Danish Eulerian Hemispheric Model – a three-dimensional air pollution model used for the Arctic, *Atmos. Environ*. **31**, 24, 4169-4191, DOI: 10.1016/S1352-2310(97)00264-1

Di Sarra, A., T. Di Iorio, and M. Cacciani (2001), Saharan dust profiles measured by lidar at Lampedusa, *J. Geophys. Res. Atmos.* **106**, D10, 10335-10347.

Draxler, R.R., and G.D. Rolph (2010), HYSPLIT (HYbrid Single-Particle Lagrangian Integrated Trajectory) Model access via NOAA ARL READY, NOAA Air Resources Laboratory, Silver Spring, MD, available from: http://ready.arl.noaa.gov/HYSPLIT.php.

Duce, R.A. (1995), Sources, distributions, and fluxes of mineral aerosols and their relationship to climate. **In:** R. Charlson and J. Heintzenberg (eds.), *Aerosol Forcing of Climate*, Wiley, New York, 43-72.

Engelstaedter, S., and R. Washington (2007), Atmospheric controls on the annual cycle of North African dust, *J. Geophys. Res. Atmos.* **112**, D3, D03103, DOI: 10.1029/2006JD007195.

Engelstaedter, S., I. Tegen, and R. Washington (2006), North African dust emissions and transport, *Earth Sci. Rev.* **79**, 1-2, 73-100, DOI: 10.1016/j.earscirev.2006.06.004.

Eresmaa, N., A. Karppinen, S.M. Joffre, J. Rasanen, and H. Talvitie (2006), Mixing height determination by ceilometers, *Atmos. Chem. Phys.* **6**, 1485-1493.

Fernald, F.G. (1984), Analysis of atmospheric lidar observations: some comments, *Appl. Opt*. **23**, 5, 652-653.

Flentje, H., B. Heese, J. Reichardt, and W. Thomas (2010), Aerosol profiling using the ceilometer network of the German Meteorological Service, *Atmos. Meas. Tech*. **3**, 3643-3673, DOI: 10.5194/amtd-3-3643-2010.

Formenti, P., J.L. Rajot, K. Desboeufs, S. Caquineau, S. Chevaillier, S. Nava, A. Gaudichet, E. Journet, S. Triquet, S. Alfaro, M. Chiari, J. Haywood, H. Coe, and E. Highwood (2008), Regional variability of the composition of mineral dust from western Africa: Results from the AMMA SOP0/ DABEX and DODO field campaigns, *J. Geophys. Res. Atmos*. **113**, D20, D00C13, DOI: 10.1029/2008JD009903.

Freudenthaler, V., M. Esselborn, M. Wiegner, B. Heese, M. Tesche, A. Ansmann, D. Müller, D. Althausen, M. Wirth, A. Fix, G. Ehret, P. Knippertz, C. Toledano, M. Garhammer, and M. Seefeldner (2009), Depolarization-ratio profiling at several wavelengths in pure Saharan dust during SAMUM 2006, *Tellus B* **61**, 1, 165-179, DOI: 10.1111/j.1600-0889.2008.00396.x.

Frey, S., K. Poenitz, G. Teschke, and H. Wille (2010), Detection of aerosol layers with ceilometer and the recognition of the mixed layer depth. **In:** *Proc. Int. Symp. for Advancement of Boundary Layer Remote (ISARS)*, 3646-3647.

Ginoux, P., M. Chin, I. Tegen, J.M. Prospero, B. Holben, O. Dubovik, and S.-J. Lin (2001), Sources and distributions of dust aerosols simulated with the GOCART model, *J. Geophys. Res.* **106**, D17, 20255-20273, DOI: 10.1029/ 2000JD000053.

Goudie, A.S., and N.J. Middleton (2001), Saharan dust storms: nature and consequences, *Earth Sci. Rev*. **56**, 1-4, 179-204, DOI: 10.1016/S0012-8252 (01)00067-8.

Gross, S., M. Tesche, V. Freudenthaler, C. Toledano, M. Wiegner, A. Ansmann, D. Althausen, and M. Seefeldner (2011), Characterization of Saharan dust, marine aerosols and mixtures of biomass-burning aerosols and dust by means of multi-wavelength depolarization and Raman lidar measurements during SAMUM 2, *Tellus B* **63**, 4, 706-724, DOI: 10.1111/j.1600-0889. 2011.00556.x.

Guerrero-Rascado, L., F.J. Olmo, I. Avilés-Rodríguez, F. Navas-Guzmán, D. Pérez-Ramírez, H. Lyamani, and L. Alados Arboledas (2009), Extreme Saharan dust event over the southern Iberian Peninsula in September 2007: active and passive remote sensing from surface and satellite, *Atmos. Chem. Phys*. **9**, 21, 8453-8469.

Guerrero-Rascado, J.L., M.J. Costa, D. Bortoli, A.M. Silva, H. Lyamani, and L. Alados-Arboledas (2010), Infrared lidar overlap function: an experimental determination, *Opt. Express* **18**, 19, 20350-20359, DOI: 10.1364/ OE.18.020350.

Heese, B., H. Flentje, D. Althausen, A. Ansmann, and S. Frey (2010), Ceilometer-lidar inter-comparision: backscatter coefficient retrieval and signal-to-noise ratio determination, *Atmos. Meas. Tech*. **3**, 3907-3924, DOI: 10.5194/amtd-3-3907-2010.

Heintzenberg, J. (2009), The SAMUM-1 experiment over Southern Morocco: overview and introduction, *Tellus B* **61**, 1, 2-11, DOI: 10.1111/j.1600-0889.2008.00403.x.

Hess, M., P. Koepke, and I. Schult (1998), Optical properties of aerosols and clouds: The software package OPAC, *Bull. Am. Meteorol. Soc*. **79**, 5, 831-844, DOI: 10.1175/1520-0477(1998)079<0831:OPOAAC>2.0.CO;2.

Hogan, T.F., and L.R. Brody (1993), Sensitivity studies of the Navy's global forecast model parameterizations and evaluation of improvements to NOGAPS, *Mon. Weather Rev*. **121**, 8, 2373-2395, DOI: 10.1175/1520-0493(1993)121<2373:SSOTNG>2.0.CO;2.

Hogan, T.F., and T.E. Rosmond (1991), The description of the Navy operational global atmospheric prediction system, *Mon. Weather Rev.* **119**, 8, 1786-1815, DOI: 10.1175/1520-0493(1991)119<1786:TDOTNO>2.0.CO;2.

Holben, B.N., T.F. Eck, I. Slutsker, D. Tanré, J.P. Buis, A. Setzer, E. Vermote, J.A. Reagan, Y.J. Kaufman, T. Nakajima , F. Lavenu, I. Jankowiak, and A. Smirnov (1998), AERONET – A federated instrument network and data archive for aerosol characterization, *Remote Sens. Environ*. **66**, 1, 1-16, DOI: 10.1016/S0034-4257(98)00031-5.

Huang, L., J.H. Jiang, J.L. Tackett, H. Su, and R. Fu (2013), Seasonal and diurnal variations of aerosol extinction profile and type distribution from CALIPSO 5-year observations, *J. Geophys. Res. Atmos*. **118**, 10, 4572-4596, DOI: 10.1002/jgrd.50407.

IPCC (2014), Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (T.F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)), Cambridge University Press, Cambridge, 1535 pp.

Israelevich, P.L., E. Ganor, Z. Levin, J.H. Joseph (2003), Annual variations of physical properties of desert dust over Israel, *J. Geophys. Res*. **108**, D13, 4381, DOI: 10.1029/2002JD003163.

Jimenez, J.L., M.R. Canagaratna, N.M. Donahue, A.S.H. Prevot, Q. Zhang, J.H. Kroll, P.F. DeCarlo, J.D. Allan, H. Coe, N.L. Ng, A.C. Aiken, K.S. Docherty, I.M. Ulbrich, A.P. Grieshop, A.L. Robinson, J. Duplissy, J.D. Smith, K.R. Wilson, V.A. Lanz, C. Hueglin, Y.L. Sun, J. Tian, A. Laaksonen, T. Raatikainen, J. Rautiainen, P. Vaattovaara, M. Ehn, M. Kulmala, J.M. Tomlinson, D.R. Collins, M.J. Cubison, E.J. Dunlea, J.A. Huffman, T.B. Onasch, M.R. Alfarra, P.I. Williams, K. Bower, Y. Kondo, J. Schneider, F. Drewnick, S. Borrmann, S. Weimer, K. Demerjian, D. Salcedo, L. Cottrell, R. Griffin, A. Takami, T. Miyoshi,

S. Hatakeyama, A. Shimono, J.Y. Sun, Y.M. Zhang, K. Dzepina, J.R. Kimmel, D. Sueper, J.T. Jayne, S.C. Herndon, A.M. Trimborn, L.R. Williams, E.C. Wood, A.M. Middlebrook, C.E. Kolb, U. Baltensperger, and D.R. Worsnop (2009), Evolution of organic aerosols in the atmosphere, *Science* **326**, 5959, 1525-1529, DOI: 10.1126/science.1180353.

Kandler, K., L. Schuetz, C. Deutscher, M. Ebert, H. Hofmann, S. Jackel, R.P. Knippertz, K. Lieke, A. Masling, A. Petzold, A. Schladitz, B. Weinzier, A. Wiedensohler, S. Zorn, and S. Weinbruch (2009), Size distribution, mass concentration, chemical and mineralogical composition and derived optical parameters of the boundary layer aerosol at Tinfou, Morocco, during SAMUM 2006, *Tellus B* **61**, 1, 32-50, DOI: 10.1111/j.1600-0889.2008.00385.x.

Klett, J.D. (1985), Lidar inversions with variable backscatter/extinction values, *Appl. Opt.* **24**, 11, 211-220, DOI: 10.1364/AO.24.001638.

Landulfo, E., A. Papayannis, P. Artaxo, A.D.A. Castanho, A.Z. de Freitas, R.F. Souza, N.D. Vieira Junior, M.P.M.P. Jorge, O.R. Sánchez-Ccoyllo, and D.S. Moreira (2003), Synergetic measurements of aerosols over São Paulo, Brazil using LIDAR, sunphotometer and satellite data during the dry season, *Atmos. Chem. Phys.* **3**, 5, 1523-1539, DOI: 10.5194/acp-3-1523-2003.

Lopes, F.J.S., E. Landulfo, and M.A. Vaughan (2013), Evaluating CALIPSO's 532 nm lidar ratio selection algorithm using AERONET sun photometers in Brazil, *Atmos. Meas. Tech.* **6**, 11, 3281-3299, DOI: 10.5194/amt-6-3281-2013.

Maciszewska, A., K. Markowicz, and M. Witek (2010), Multi year analysis of the aerosol optical thickness over Europe, *Acta Geophys.* **58**, 6, 1147-1163, DOI: 10.2478/s11600-010-0034-5.

Marécal, V., V.H. Peuch, C. Andersson, S. Andersson, J. Arteta *et al.* (2015), A regional air quality forecasting system over Europe: the MACC-II daily ensemble production. *Geosci. Model Dev.* **8**, 9, 2777-2813, DOI: 10.5194/gmd-8-2777-2015.

Markowicz, K.M., P.J. Flatau, A.M. Vogelmann, P.K. Quinn, and D. Bates (2003), Modeling and observations of infrared radiative forcing during ACE-Asia, *Quart. J. Roy. Meteorol. Soc.* **129**, 594, 2927-2947.

Markowicz, K.M., P.J. Flatau, A.E. Kardas, K. Stelmaszczyk, and L. Woeste (2008), Ceilometer retrieval of the boundary layer vertical aerosol extinction structure, *J. Atmos. Ocean. Techn.* **25**, 6, 928-944, DOI: 10.1175/ 2007JTECHA 1016.1.

Markowicz, K.M., T. Zielinski, S. Blindheim, M. Gausa, A.K. Jagodnicka, A.E. Kardas, W. Kumala, Sz.P. Malinowski, M. Posyniak, T. Petelski, and T. Stacewicz (2012), Study of vertical structure of aerosol optical properties by sun photometers and ceilometer during macron campaign in 2007, *Acta Geophys.* **60**, 5, 1308-1337, DOI: 10.2478/s11600-011-0056-7.

Marsham, J.H., D.J. Parker, C.M. Grams, C.M. Taylor, and J.M. Haywood (2008), Uplift of Saharan dust south of the intertropical discontinuity, *J. Geophys. Res. Atmos*. **113**, D21, D21102.

Marsham, J.H., M. Hobby, C.J.T. Allen, J.R., Banks, M. Bart *et al*. (2013), Meteorology and dust in the central Sahara: Observations from Fennec supersite-1 during the June 2011 Intensive Observation Period, *J. Geophys. Res*. **118**, 10, 4069-4089, DOI: 10.1002/jgrd.50211.

Martucci, G., C. Milroy, and C.D. O'Dowd (2010), Detection of cloud-base height using Jenoptik CHM15K and Vaisala CL31 ceilometers, *J. Atmos. Ocean. Technol*. **2**, 305-318.

McConnell, C.L., E.J. Highwood, H. Coe, P. Formenti, B. Anderson *et al*. (2008), Seasonal variations of the physical and optical characteristics of Saharan dust: Results from the Dust Outflow and Deposition to the Ocean (DODO) experiment, *J. Geophys. Res*. **113**, D14, DOI: 10.1029/2007JD009606.

McKendry, I.G., D. van der Kamp, K.B. Strawbridge, A. Christen, and B. Crawford (2009), Simultaneous observations of boundary-layer aerosol layers with CL31 ceilometer and 1064/532 nm lidar, *Atmos. Environ*. **43**, 36, 5847-5852, DOI: 10.1016/j.atmosenv.2009.07.063.

Mona, L., Z. Liu, D. Müller, A. Omar, A. Papayannis, G. Pappalardo, N. Sugimoto, and M. Vaughan (2012), Lidar measurements for desert dust characterization: an overview, *Adv. Meteorol.* **2012**, 356265, DOI: 10.1155/2012/356265.

Mona, L., N. Papagiannopoulos, S. Basart, J. Baldasano, and I. Binietoglou *et al*. (2014), EARLINET dust observations vs. BSC-DREAM8b modeled profiles: 12-year-long systematic comparison at Potenza, Italy, *Atmos. Chem. Phys*. **14**, 16, 8781-8793, DOI: 10.5194/acp-14-8781-2014.

Morcrette, J.-J. *et al*. (2009), Aerosol analysis and forecast in the European centre for medium-range weather forecasts integrated forecast system: forward modeling, *J. Geophys. Res*. **114**, D06206, DOI: 10.1029/2008JD011235.

Morys, M., Mims III, F.M. Hagerup, S. Anderson, S.E. Baker, A. Kia, and J. Walkup (2001), Design calibration, and performance of MICROTOPS II handheld ozone monitor and Sun photometer, *J. Geophys. Res*. **106**, D13, 14573-14582, DOI: 10.1029/2001JD900103.

Münkel, C., S. Emeis, W.J. Mueller, and K.P. Schaefer (2004), Aerosol concentration measurements with a lidar ceilometer: Results of a one year measuring campaign. **In:** K. Schaefer *et al.* (eds.), Remote Sensing of Clouds and the Atmosphere VIII, International Society for Optical Engineering (*SPIE Proc*. **5235**), 486-496.

Nickovic, S., G. Kallos, A. Papadopoulos, and O. Kakaliagou (2001), A model for prediction of desert dust cycle in the atmosphere, *J. Geophys. Res. Atmos*. **106**, D16, 18113-118129, DOI: 10.1029/2000JD900794.

O'Connor, E.J., A.J. Illingworth, and R.J. Hogan (2004), A technique for autocalibration of cloud lidar, *J. Atmos. Ocean. Technol*. **21**, 5, 777-786, DOI: 10.1175/1520-0426(2004)021<0777:ATFAOC>2.0.CO;2.

Osborne, S.R., B.T. Johnson, J.M. Haywood, A.J. Baran, M.A.J. Harrison *et al*. (2008), Physical and optical properties of mineral dust aerosol during the Dust and Biomass-burning Experiment, *J. Geophys. Res.* **113**, D00C03, DOI: 10.1029/2007JD009551.

Papayannis, A., D. Balis, V. Amiridis, G. Chourdakis, G. Tsaknakis, C. Zerefos, A.D.A. Castanho, S. Nickovic, S. Kazadzis, and J. Grabowski (2005), Measurements of Saharan dust aerosols over the Eastern Mediterranean using elastic backscatter-Raman lidar, spectrophotometric and satellite observations in the frame of the EARLINET project, *Atmos. Chem. Phys.* **5**, 8, 2065-2079.

Papayannis, A., H.Q. Zhang, V. Amiridis, H.B. Ju, G. Chourdakis, G. Georgoussis, C. Pérez, H.B. Chen, P. Goloub, R.E. Mamouri, S. Kazadzis, D. Paronis, G. Tsaknakis, and J.M. Baldasano (2007), Extraordinary dust event over Beijing, China, during April 2006: Lidar, Sun photometric, satellite observations and model validation, *J. Geophys. Res.* **34**, 7, L07806, DOI: 10.1029/2006GL029125.

Papayannis, A., V. Amiridis, L. Mona, G. Tsaknakis, D. Balis, J. Bösenberg, A. Chaikovski, F. De Tomasi, I. Grigorov, I. Mattis, V. Mitev, D. Müller, S. Nickovic, C. Pérez, A. Pietruczuk, G. Pisani, F. Ravetta, V. Rizi, M. Sicard, T. Trickl, M.Wiegner, and M. Gerding (2008), Systematic lidar observations of aerosol optical properties during Saharan dust intrusions over Europe, in the frame of EARLINET (2000-2002): Statistical analysis and results, *J. Geophys. Res.* **113**, D10, D10204, DOI: 10.1029/2007JD009028.

Pappalardo, G., L. Mona, G. D'Amico, U. Wandinger, M. Adam,A. Amodeo *et al*. (2013), Four-dimensional distribution of the 2010 Eyjafjallajökull volcanic cloud over Europe observed by EARLINET, *Atmos. Chem. Phys.* **13**, 8, 4429-4450, DOI: 10.5194/acp-13-4429-2013.

Pavese, G., M. Calvello, F. Esposito, L. Leone, and R. Restieri (2012), Effects of Saharan dust advection on atmospheric aerosol properties in the West-Mediterranean area, *Adv. Meteorol.* **2012**, 730579, DOI: 10.1155/2012/730579.

Perez, C., S. Nickovic, G. Pejanovic, J. M. Baldasano, and E. Ozsoy (2006a), Interactive dust-radiation modeling: A step to improve weather forecasts, *J. Geophys. Res*. **111**, D16, D16206, DOI: 10.1029/2005JD006717.

Perez, C., S. Nickovic, J.M. Baldasano, M. Sicard, F. Rocadenbosch, and V.E. Cachorro (2006b), A long Saharan dust event over the western Mediterranean: Lidar, Sun photometer observations, and regional dust modeling, *J. Geophys. Res.* **111**, D15, D15214, DOI: 10.1029/2005JD006579.

Pisani, G., A. Boselli, N. Spinelli, and X. Wang (2011), Characterization of Saharan dust layers over Naples (Italy) during 2000-2003 EARLINET project, *Atmos. Res.* **102**, 3, 286-299, DOI: 10.1016/j.atmosres.2011.07.012.

Preißler, J., F. Wagner, S.N. Pereira, and J.L. Guerrero-Rascado (2011), Multi-instrumental observation of an exceptionally strong Saharan dust outbreak over Portugal, *J. Geophys. Res.* **116**, D24, D24204, DOI: 10.1029/2011JD016527.

Prospero, J.M., P. Ginoux, O. Torres, S.E. Nicholson, and T.E. Gill (2002), Environmental characterization of global sources of atmospheric soil dust identified with the Nimbus 7 total ozone mapping spectrometer (toms) absorbing aerosol product, *Rev. Geophys.* **40**, 1, 2-1-2-31, DOI: 10.1029/2000RG000095.

Sasano, Y., E.V. Browell, and S. Ismail (1985), Error caused by Rusing a constant extinction/backscattering ratio in the lidar solution, *Appl. Opt.* **24**, 22, 3929-3932, DOI: 10.1364/AO.24.003929.

Schutgens, N., M. Nakata, and T. Nakajima (2012), Estimating aerosol emissions by assimilating remote sensing observations into a global transport model, *Remote Sens.* **4**, 11, 3528-3543, DOI: 10.3390/rs4113528.

Shifrin, K.S. (1995), Simple relationships for the Angstrom parameter of disperse systems, *Appl. Opt.* **34**, 21, 4480-4485, DOI: 10.1364/AO.34.004480.

Sinha, P., D. Kaskaoutis, R. Manchanda, and S. Sreenivasan (2012), Characteristics of aerosols over Hyderabad in southern Peninsular India: synergy in the classification techniques, *Ann. Geophys*. **30**, 9, 1393-1410, DOI: 10.5194/angeo-30-1393-2012.

Smirnov, A., B.N. Holben, T.F. Eck, O. Dubovik, and I. Slutsker (2000), Cloud screening and quality control algorithms for the AERONET database, *Rem. Sens. Env.* **73**, 3, 337-349, DOI: 10.1016/S0034-4257(00)00109-7.

Stachlewska, I.S., and C. Ritter (2010), On retrieval of lidar extinction profiles using Two-Stream and Raman techniques, *Atmos. Chem. Phys*. **10**, 6, 2813-2824, DOI: 10.5194/acp-10-2813-2010.

Stachlewska, I.S., K.M. Markowicz, and M. Piądłowski (2010), On forward Klett's inversion of ceilometer signals. **In:** *25th ILRC International Laser Radar Conference, 5-9 July 2010, St. Petersburg, Russia.*

Stachlewska, I.S., M. Piądłowski, S. Migacz, A. Szkop, A.J. Zielińska, and P.L. Swaczyna (2012), Ceilometer observations of the boundary layer over Warsaw, Poland, *Acta Geophys.* **60**, 5, 1386-1412, DOI: 10.2478/s11600-012-0054-4.

Sundström, A.-M., T. Nousiainen, and T. Petäjä (2009), On the quantitative lowlevel aerosol measurements using ceilometer-type lidar, *J. Atmos. Ocean. Technol*. **26**, 11, 2340-2352, DOI: 10.1175/2009JTECHA1252.1.

Tegen, I., K. Schepanski, and B Heinold (2013), Comparing two years of Saharan dust source activation obtained by regional modelling and satellite observations, *Atmos. Chem. Phys.* **13**, 5, 2381-2390, DOI: 10.5194/acp-13-2381-2013.

Tesche, M., A. Ansmann, D. Mueller, D. Althausen, I. Mattis *et al.* (2009), Vertical profiling of Saharan dust with Raman lidars and airborne HSRL in southern Morocco during SAMUM, *Tellus B* **61**, 1, 144-164, DOI: 10.1111/j.1600-0889.2008.00390.x.

Varga, G., J. Kovács, and G. Újvári (2013), Analysis of Saharan dust intrusions into the Carpathian Basin (Central Europe) over the period of 1979-2011, *Global Planet Change* **100**, 333-342, DOI: 10.1016/j.gloplacha.2012. 11.007.

Vogelmann, A., P. Flatau, M. Szczodrak, K. Markowicz, and P. Minnett (2003), Observations of large greenhouse effects for anthropogenic aerosols, *Geophys. Res. Lett.* **30**, 12, 1654-1657.

Wagner, F., and A.M. Silva (2008), Some considerations about Angström exponent distributions, *Atmos. Chem. Phys*. **8**, 3, 481-489.

Wandinger, U., and A. Ansmann (2002), Experimental determination of the lidar overlap profile with Raman lidar, *Appl. Opt*. **41**, 3, 511-514, DOI: 10.1364/AO.41.000511.

Wang, X., A. Boselli, L. D'Avino, G. Pisani, N. Spinelli, A. Amodeo, A. Chaikovsky, M. Wiegner, S. Nickovic, A. Papayannis, M.R. Perrone, V. Rizi, L. Sauvage, and A. Stohl (2008), Volcanic dust characterization by EARLINET during Etna's eruptions in 2001-2002, *Atmos. Environ*. **42**, 5, 893-905, DOI: 10.1016/j.atmosenv.2007.10.020.

Wang, Z., H.H. Zhang, X. Jing, and X. Wei (2013), Effect of non-spherical dust aerosol on its direct radiative forcing, *Atmos. Res.* **120**, 112-126, DOI: 10.1016/j.atmosres.2012.08.006.

Weitkamp, C. (ed.) (2005), Lidar: Range-resolved Optical Remote Sensing of the Atmosphere, Springer, New York.

Westphal, D.L., O.B. Toon, and T.N. Carlson (1988), A case study of mobilization and transport of Saharan dust, *J. Atmos. Sci.* **45**, 15, 2145-2175, DOI: 10.1175/1520-0469(1988)045<2145:ACSOMA>2.0.CO;2.

Wiegner, M., and A. Geiß (2012), Aerosol profiling with the Jenoptik ceilometer CHM15kx, *Atmos. Meas. Tech.* **5**, 8, 1953-1964, DOI: 10.5194/amt-5-1953-2012.

Witek, M.L., P.J. Flatau, P.K. Quinn, and D.L. Westphal (2007), Global sea-salt modeling: Results and validation against multicampaign shipboard measurements, *J. Geophys. Res*. **112**, D8, D08215, DOI: 10.1029/2006JD 007779.

Wong, M.S., M.I. Shahzad, J.E. Nichol, K.H. Lee, and P.W. Chan (2013), Validation of MODIS, MISR, OMI, and CALIPSO aerosol optical thickness using ground-based sunphotometers in Hong Kong, *Int. J. Remote Sens*. **34**, 3, 897-918, DOI: 10.1080/01431161.2012.720739.

Zawadzka, O., K. Markowicz, A. Pietruczuk, T. Zielinski, and J. Jaroslawski (2013), Impact of urban pollution emitted in Warsaw on aerosol properties, *Atmos. Environ*. **69**, 15-28, DOI: 10.1016/j.atmosenv.2012.11.065.

Zhang, J., J.S. Reid, D.L. Westphal, N.L. Baker, and E.J. Hyer (2008), A system for operational aerosol optical depth data assimilation over global oceans, *J. Geophys. Res*. **113**, D10, D10208, DOI: 10.1029/2007JD009065.

Received 7 February 2015
Received in revised form 12 November 2015
Accepted 8 December 2015

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2591-2608
DOI: 10.1515/acgeo-2016-0101

Imprints of Natural Phenomena and Human Activity Observed During 10 Years of ELF Magnetic Measurements at the Hylaty Geophysical Station in Poland

Zenon NIECKARZ

Institute of Physics, Jagiellonian University, Kraków, Poland;
e-mail: zenon.nieckarz@uj.edu.pl

Abstract

Current human activity produces strong electromagnetic pollution. The power spectrum in the extremely low frequency (ELF, 3-3000 Hz) range is mainly polluted by anthropogenic narrow spectral lines at 16.66, 50, and 60 Hz and their harmonics. Meanwhile, signatures connected with natural phenomena appearing in the Earth's atmosphere, ionosphere and magnetosphere are also observed in the same frequency range.

This paper presents the amplitude behaviour of the anthropogenic lines in the years 2005-2014 based on the 10 years of activity of the Hylaty station situated in southeast Poland. The analysis includes, *i.a.*, an assessment of the correctness of the choice of the Bieszczady mountains as a location for the installation of an ELF station for long-term geophysical and climatological studies.

Key words: electromagnetic pollution, human activity, ELF waves.

1. INTRODUCTION

The measurement of magnetic field components in the frequency domain from nearly DC to a few hundred Hertz at stations located on the surface of the Earth is a widely used method in geophysical, meteorological, climatological, and earth science studies. Different signatures of the following phenomena are observed in the records obtained: structured and unstructured pearl pulsation (Pc1) elicited by geomagnetic storms, spectral resonance structures (SRS) associated with activity in the Alfvén resonator (IAR), and Schuman resonances (SchR) connected with lightning activity inside the Earth-ionosphere (E-i) cavity.

Measurements of ELF waves are performed in several regions of the world: in the USA at Hollister (36.8°N, 121.5°W) and Rhode Island (41.6°N, 71.6°W); in Israel at Mitzpe-Ramon (30.5°N, 34.4°E); in Japan at Moshiri (142.2°E, 44.3°N); in Russia at Lekhta (33.9°E, 64.4°N), and Vernadsky (65.3°S, 64.2°W); and the Ukrainian ELF station at Antarctica. General and detailed descriptions are available in many papers (Fraser-Smith and Helliwell 1994, Price *et al.* 1999, Hobara *et al.* 2000, Bezrodny *et al.* 2007). In Europe a few stations also work continuously, such as Nagycenk (47.6°N, 16.7°E) in Hungary, and Eskdalemuir in United Kingdom (55.314°N, 3.206°W), which are described in detail in papers (Sátori *et al.* 1996, Beggan *et al.* 2012). Most of these stations correctly measure the signals at frequencies below 3 Hz, which fall within the ULF (ultra low frequency) range. It should be noted that the ULF wave range is widely used in geophysics studies, *i.e.*, by INTERMAGNET Magnetic Observatories, which use sampling intervals of one or a few seconds to observe Earth's magnetic field (Jankowski and Sucksdorff 1996, Love and Chulliat 2013, Turbitt *et al.* 2013).

Two centres in Poland perform measurements of ELF waves. One of them is the Institute of Geophysics, Polish Academy of Sciences, which founded two ELF stations (Neska and Sátori 2006), one in the Polish Polar Station at Spitsbergen (77.0°N, 15.5°E) and second in the Central Geophysical Observatory in Belsk (51.8°N, 20.8°E). The other centre is the Astronomical Observatory of the Jagiellonian University in Kraków, which has provided continuous ELF measurements at the Hylaty ELF station since 2006. The station is situated in southeast Poland (Bieszczady Mountains, 49.2°N, 22.5°E). The data are gathered and shared by Cracow ELF Group (Kulak *et al.* 2014).

The aim of this paper is to look at ELF signals recorded in Hylaty station for identification and characterization of anthropogenic traces as well as for analysing their influence on the quality of study of the natural terrestrial and extra-terrestrial sources of signals.

2. DATA SET

In this paper, data collected by the Hylaty ELF station have been used. Figure 1 shows the location of the Hylaty ELF station. The geographic and geomagnetic coordinates are (49.2°N, 22.5°E) and (44.6°N, 96.8°E), respectively and the *L*-shell parameter is 1.98. Figure 2 shows the block diagram of the station. It contains a receiver and antennae. The receiver is placed in an underground container. Two orthogonal magnetic antennae are placed about 100 m away and oriented in the North-South and East-West directions. The EM waves are transversal in nature and their ELF range propagates within the Earth-ionosphere cavity. For this reason, the magnetic antenna positioned along the EW axis has the highest reception sensitivity to waves propagating along the NS axis and, by analogy, the antenna positioned along the NS axis receives best the waves propagating along the EW axis, as predicted by the Hertz dipole's directional reception characteristic.

The antennae are 1 m long and are designed as active magnetic antennae powered by the receiver unit. The station produces 2.5 MB of data per hour with a sampling frequency of 175 Hz and a dynamic range of 16 bits, and has the 3 dB frequency bandwidth of 0.03 to 55 Hz. More details of this station are described by Kulak *et al.* (2014).

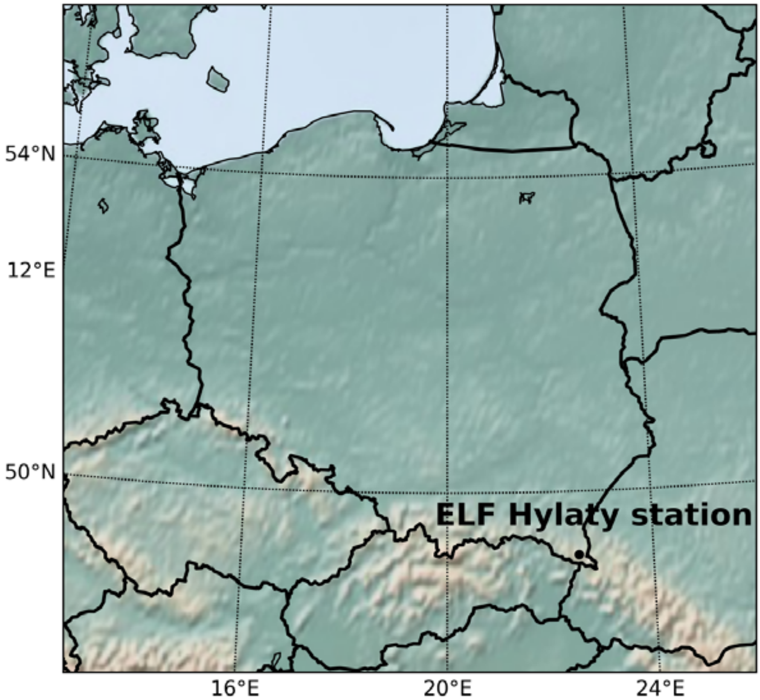


Fig. 1. Map showing the location of the ELF station in Poland.

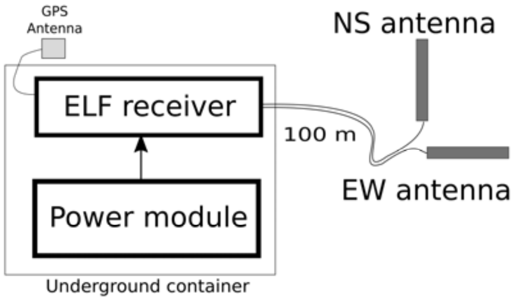


Fig. 2. Block diagram of the Hylaty ELF station.

Data sets from two magnetic antennae and a thermometer inside an underground container have been collected during the station's ten years of operation. Figure 3 (solid line) presents the variability of the daily average temperature during the years 2005-2014. The lowest instantaneous value of temperature was not colder than +1.4°C. The daily average minimum (+1.4°C) and maximum (+16.6°C) temperatures were recorded during the days 19-23 February 2012 and 1-4 August 2005, respectively. The average temperature for all ten years was 7.9 ± 4.2°C.

The nearest station to Hylaty in Poland is operated by the Institute of Meteorology and Water Management and is located at Lesko at a distance of 36 km. The data are available via FTP protocol: ftp://ftp.ncdc.noaa.gov/pub /data/gsod. The change of daily air temperature measured by this hydrological and meteorological station is shown in Fig. 3 (as dots). The close correlation between both sets of temperatures is clearly visible.

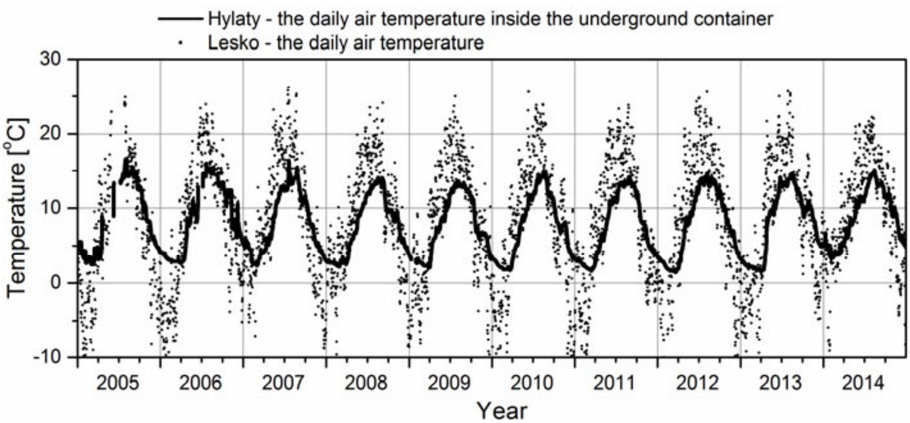


Fig. 3. The variation of the daily air temperature at Lesko (dots) and the daily air temperature inside the underground container (line) during the 10-year period (2005-2014).

The annual maximum and minimum air temperatures in Lesko were earlier than in the temperatures recorded by the ELF station and this response is expected in accordance with general scientific principles (Salazar 2006). The location of the ELF receiver in an underground chamber prevents the temperature from lowering below zero degrees Celsius and the device freezing. This effect achieves what is confirmed by Fig. 3. Daily temperature variations are not observed inside the container.

3. ULF/ELF SIGNATURES OF NATURAL PHENOMENA

3.1 Pc1 and Pi1 pulsations

Extensive study of the dynamic power spectrum calculated from magnetic field components allows one to distinguish the signatures of the different natural phenomena mentioned in the introduction. In the lower part of the frequency domain of the power spectrum, a signature connected with processes in the Earth's magnetosphere and solar wind is observed. Solar flares and related plasma injections into the Earth's magnetosphere causing the intensification of current circuits are commonly thought to be responsible for the waves registered by magnetometers (Engebretson *et al.* 2008). Oscillations with quasi-sinusoidal waveform were called pulsations continuous (Pc1) and those with waveforms that are more irregular are called pulsations

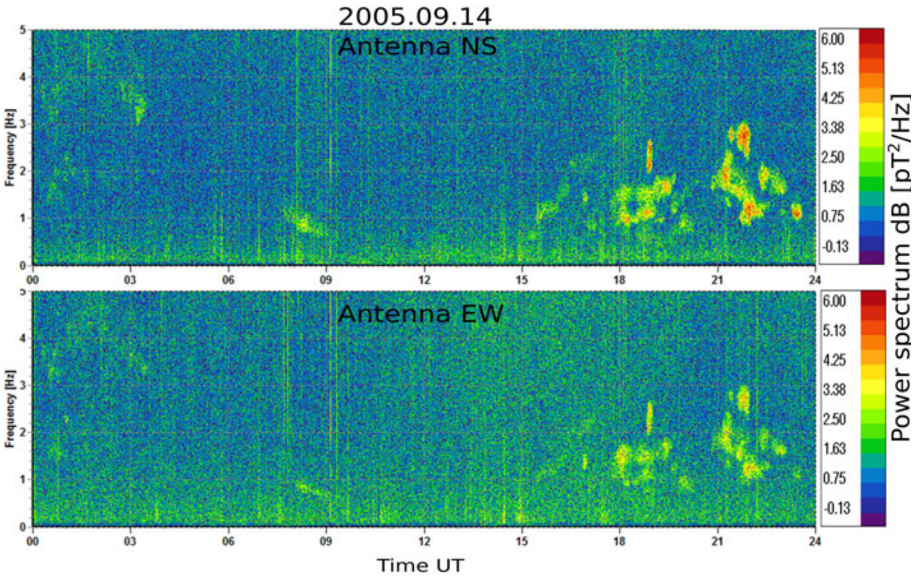


Fig. 4. The example of geomagnetic continuous pulsations (Pc1). The dynamic power spectrum is calculated on the basis of signals from NS (top panel) and EW (bottom panel) antennae collected on 14 September 2005 by the Hylaty ELF station.

irregular (Pi1). As a result, regular and irregular signals in the frequency band below 5 Hz can be observed on the Earth's surface (Kangas *et al.* 1998).

Figure 4 presents the dynamic power spectrum in the frequency range (0.03-5 Hz) for 14 September 2005 collected by the ELF station through NS and EW antennae. Geomagnetic continuous pulsations (Pc1) below 5 Hz are visible at different times during this day. The most intense Pc1 structures are observed at 18:00-24:00 UT. The Pc1 are generally more intense in the NS antenna than the EW, which is also demonstrated through differences between the NS spectrum (top panel) and the EW spectrum (bottom panel) seen in Fig. 4.

3.2 The Spectral Resonance Structures (SRS)

This is a type of signal that can also be observed in the ULF/ELF range below 5 Hz, due to ionospheric Alfvén resonator (IAR) excitation by negative-polarity cloud-to-ground (C-G) discharges whose polarity is that of the charge in the region in which the lightning leaders originated. In a similar manner, intense continuing currents associated with +CG discharges could also trigger IAR (Shalimov and Bösinger 2008). As a result of this excitation, the magnetic coils installed on the Earth's surface register signals containing spectral resonance structures (SRS). Occasionally these are also observed during winter thunderstorms in sprite-producing power spectra (Surkov *et al.* 2010). The SRS are mainly evident during night-time, independent of season, at different latitudes (Bösinger *et al.* 2002, Hebden *et al.* 2005, Odzimek *et al.* 2006, Semenova and Yahnin 2008).

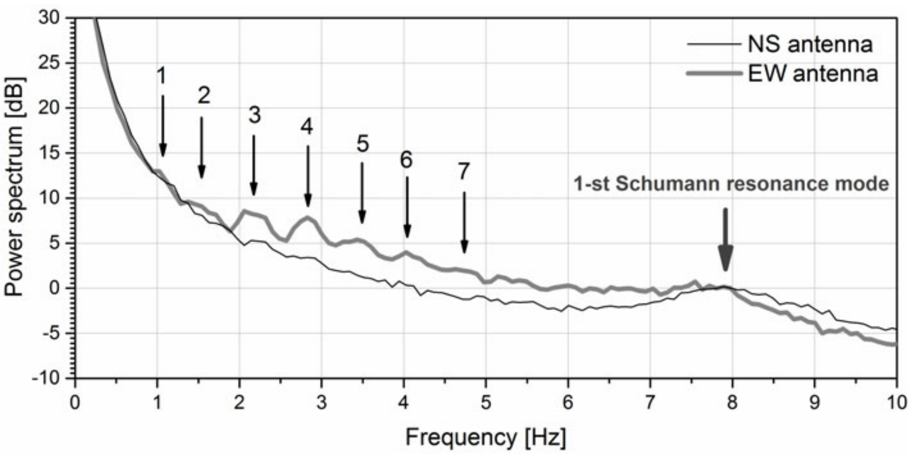


Fig. 5. The power spectrum calculated from ELF signals collected from 18:00 to 19:00 UT by the Hylaty ELF station on 24 September 2005. The seven peaks of the SRS are shown.

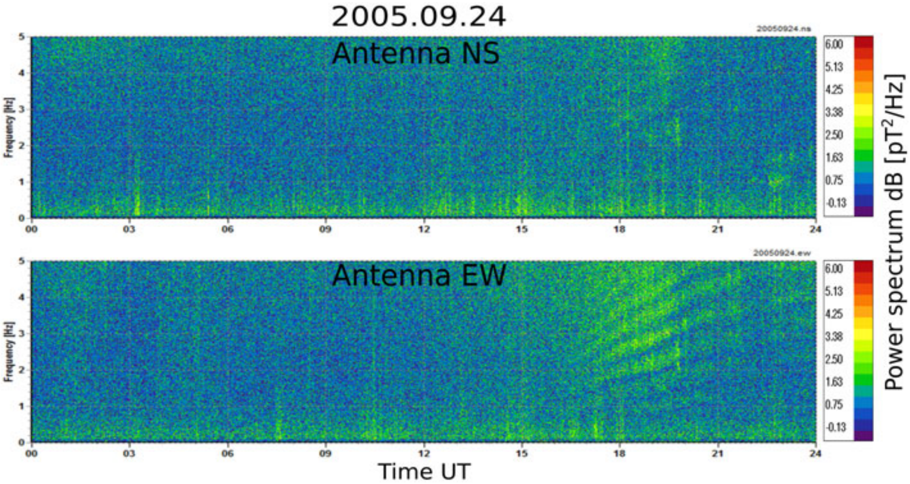


Fig. 6. The example of spectral resonance structures (SRS). The dynamic power spectrum of ELF signals is calculated on the basis of signals from NS (top panel) and EW (bottom panel) antennae collected by the Hylaty ELF station on 24 September 2005.

The SRS related to the Earth's ionosphere conductivity profile is shown in Fig. 5 as seven peaks in the power spectrum, and in Fig. 6 in which the dynamic power spectrum calculated for 24 September 2005 is presented. In the case of this particular day, we can see the most intense phenomena between 17:00 and 21:00 UT in the EW antenna (bottom panel).

3.3 Schumann resonance (SchR)

In many papers published over the last three decades, Schumann resonance phenomena have been used as a special diagnostic method for climate studies (Williams 2005). The growth in the number of studies was initiated by a paper about the connection between tropical air temperature and the intensity of Schumann resonance phenomena (Williams 1992). The Earth's surface and its ionosphere reflect electromagnetic (EM) waves in the extremely low frequency range of 3-3000 Hz (known as ELF) well. A resonant spherical cavity is thus formed in which electromagnetic waves propagate. This effect was theoretically predicted by Schumann (1952) but the first observation was made by Balser and Wagner (1960).

The strongest sources of ELF signals are lightning discharges in this cavity and these are sources of EM waves with a flat (for ELF range) and wide frequency range. The Schumann resonance is a natural phenomenon generally always visible on the power spectrum. For data from the Hylaty station, seven peaks of Schumann resonance can be found in both panels of Fig. 7 at:

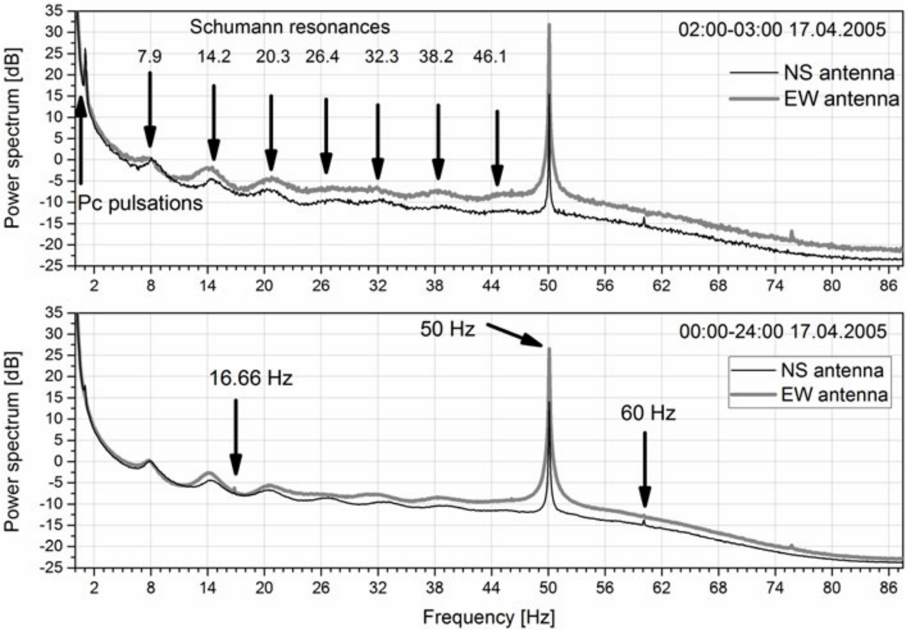


Fig. 7. The power spectrum is calculated based on ELF signals collected by the Hylaty ELF station on 17 April 2005. The top and bottom panels present the hourly (02:00-03:00 UT) and daily average power spectrum, respectively.

7.9, 14.2, 20.3, 26.4, 32.3, 38.2, and 46.1 Hz, respectively. Peak number eight of this phenomenon is masked by a strong peak at 50 Hz in the power spectrum. The next peaks are cut off by the anti-aliasing filter used at this station.

Figure 8 illustrates the dynamic power spectrum with colour-coding of amplitudes of signals with specific frequencies (vertical axis) and time (horizontal axis). The strong horizontal smudges/stripes that span the whole 24-hour periods correspond to subsequent Schumann resonances. In this particular figure, they are especially pronounced for the first three SchR modes, *i.e.*, 7.9, 14.2, 20.3 Hz. The changing intensity/colour of the smudges corresponds to SchR peak amplitude.

Daily changes in the intensity of Schumann resonance peaks are visible in Fig. 8, *e.g.*, in EW antenna it is maximum during UT afternoon hours. Information contained in the amplitudes of the SchR peaks allows one to study global and continental thunderstorm activity (Nieckarz *et al.* 2009, Dyrda *et al.* 2014). In the meantime, the analysis of frequency changes permits the study of the condition of the Earth-ionosphere cavity (Sátori *et al.* 2007) and its changes due to exposure to the Sun (Kulak *et al.* 2003, Sátori *et al.* 2005).

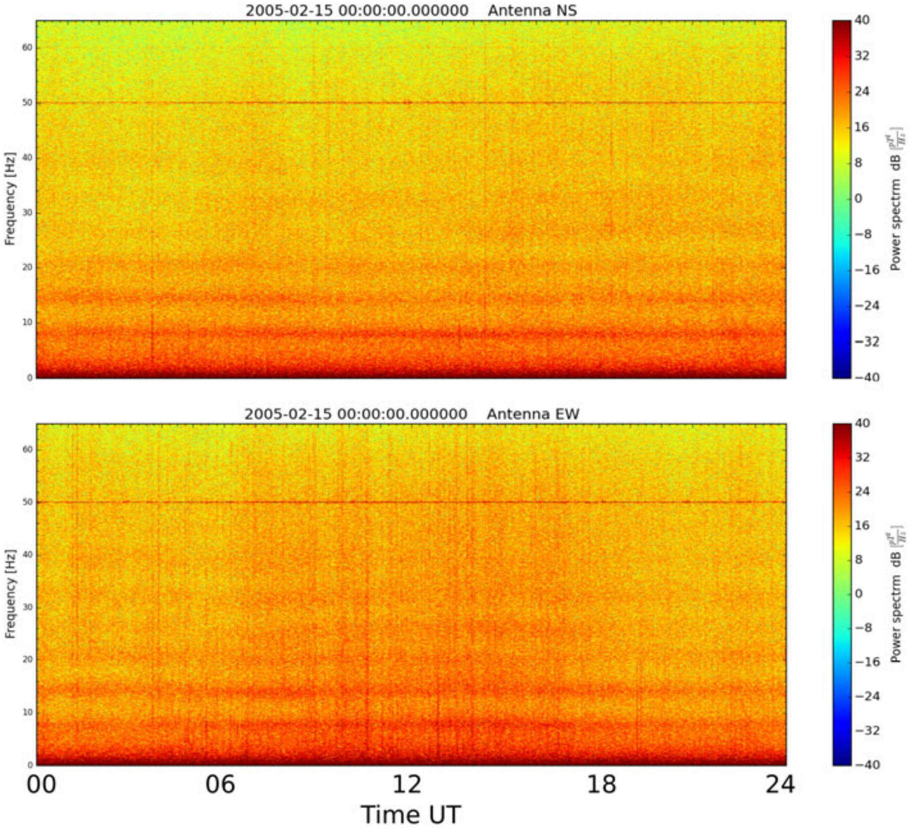


Fig. 8. The dynamic power spectrum of ELF signals is calculated on the basis of signals from NS (top panel) and EW (bottom panel) antennae collected on 15 Feb 2005 at the Hylaty ELF station. See text for details.

4. RESULTS

The most common method for assessing the occurrence of phenomena in the recorded signals, especially for studying the resonance phenomena, is the power spectrum (PS) calculated using the Fast Fourier Transform algorithm (FFT). An example of a PS recorded at the Hylaty ELF station is shown in Fig. 7. The narrow spectral lines, which are results of human activity, are clearly visible in this spectrum and are indicated by the arrows on the bottom panel.

The daily average PS was analysed. The assessment of the anthropogenic line activity was performed using three kinds of indices, AN, BG, and RLB, where AN is a net amplitude of the analysed narrow spectral line calculated as this line's amplitude read directly from the spectrum minus the spectral amplitude identified in the nearest background of this line and marked BG. RLB is an index defined as the ratio of line's net amplitude (AN) to the nearest background amplitude (BG), which can be recorded with the formula RLB = AN/BG. It was possible to calculate the indices AN, BG and RLB, each for the three spectral lines, 16.66, 50, and 60 Hz, respectively. These indices were independently computed for each of the NS and EW antennae.

Indices AN50, BG50, and RLB50 are presented in Fig. 9 for both antennae, NS and EW, during the years 2005-2014. No significant linear trend in time domain is observed for any of the indices. There is, however, a seasonal variation.

A narrow spectral peak at 50 Hz, which predominantly originates from the Polish and European power grids, is always clearly visible. The values of the AN50 index calculated for both antennae are illustrated on the top panel in Fig. 9. Both $AN50_{NS}$ and $AN50_{EW}$ indices have two seasonal maximums, in winter and summer, even if they are neither very distinct in each year nor simultaneous on all antennae. The index $AN50_{EW}$ is always larger than $AN50_{NS}$ during the entire period of study.

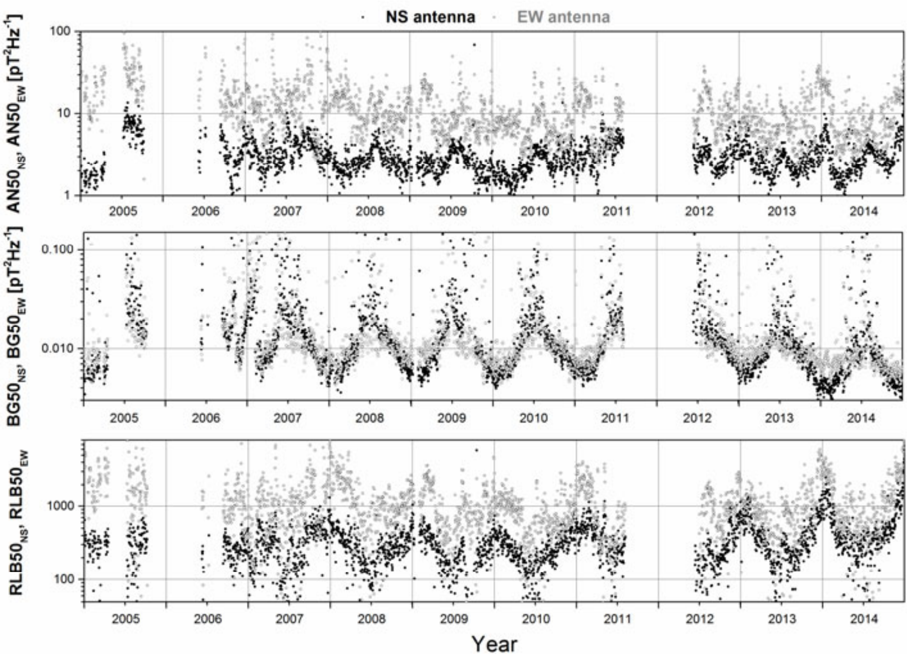


Fig. 9. The variability of indices: AN50 (top panel), BG50 (middle panel), and RLB50 (bottom panel) at both antennae during the years 2005-2014. The results are not available in certain periods due to technical problems.

On both antennae, the timeline of the BG50 index displays strong and repeatable seasonal variability with a single peak in summer and a single trough in winter. Values of the $BG50_{NS}$ and $BG50_{EW}$ indices strongly correspond with the amplitude of Schumann resonances in this range of the power spectrum (see Fig. 7). It is generally known from observations that resonance amplitudes depend on the intensity of the global thunderstorm activity (*i.e.*, the frequency of discharges, amplitudes of discharge currents and on the length of discharge channels) and on the distance between thunderstorm areas and the ELF station (Kulak *et al.* 2006), even if these dependencies are not linear. In a simplified approach, the values of the BG50 indices from both antennae provide an approximated picture of the global thunderstorm activity and in terms of its nature is compatible with the current state of research in this area. Different types of global thunderstorm activity observation indicate that the northern hemisphere is much more active than the southern counterpart (Christian *et al.* 2003). This is explained by the asymmetrical distribution of landmass and the resulting tracts of humidity and air circulation around the globe. Notably, tropical landmass provides the greatest contribution to global thunderstorm activity (Williams 2005).

The RLB50 index, defined as the AN to BG ratio, was devised to provide quantitative assessment of the contribution from the anthropogenic 50 Hz spectral line to the amplitude of the natural signal in this part of the spectrum. As a result, it was observed that while both RLB50 indices are higher in winter than in summer, the $RLB50_{EW}$ variety is always greater than $RLB50_{NS}$ across the study period. It is worth noting that the location of the 50 Hz peak in the frequency domain corresponds with the location of the eighth mode of Schumann resonance (around 51.5 Hz). It should also be noted that more than ten modes could be observed (Fullekrug 2005) when using a very wide band receiver.

The spectral line 60 Hz, which is mainly characteristic of the North American power grids, is weakly visible in datasets recorded at the Hylaty station. Usually it is only the average power spectrum produced for a longer period of time (more than 30 min) that enables spectral lines to be observed. It should be noted that the frequency of 60 Hz is in the top part of the declining transmission characteristics of station (see Fig. 7), but it is strong enough to study its behaviour in the period analysed. Figure 10 presents an analysis of variability of the indices AN60, BG60, and RLB60 during the years 2005-2014 for each of the NS and EW antennae. Both AN60 indices display seasonality with peaks in winter and summer with the former peak normally better isolated than the latter. The $AN60_{NS}$ index is greater than $AN60_{EW}$ across the study period. The pattern of the BG60 index in time has the same origin and a nearly identical shape to BG50. The only significant difference is the low values of BG60 due to the partial suppression of this spectrum range

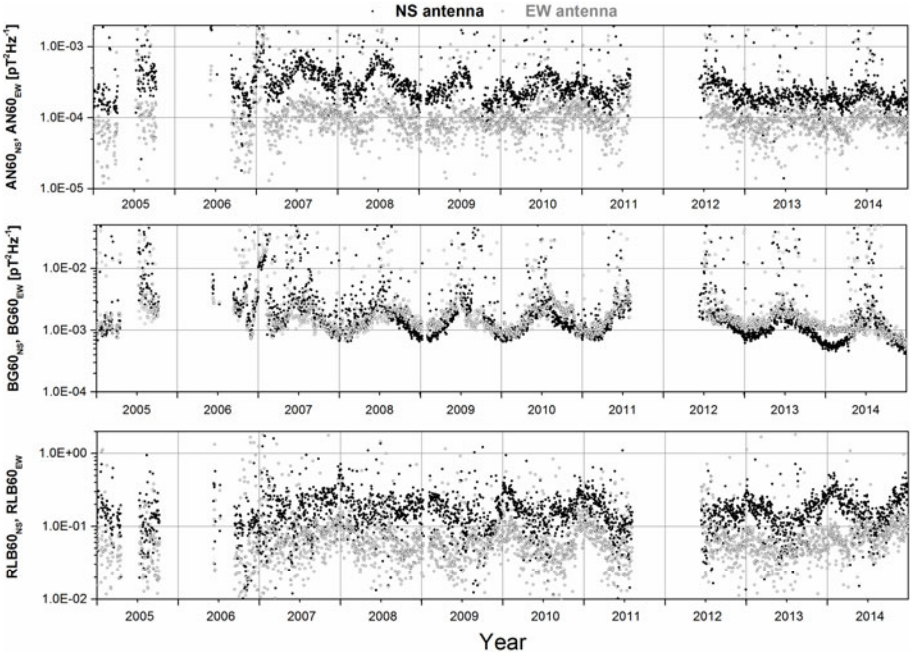


Fig. 10. The variability of indices: AN60 (top panel), BG60 (middle panel), and RLB60 (bottom panel) at both antennae during the years 2005-2014. The results are not available in certain periods due to technical problems.

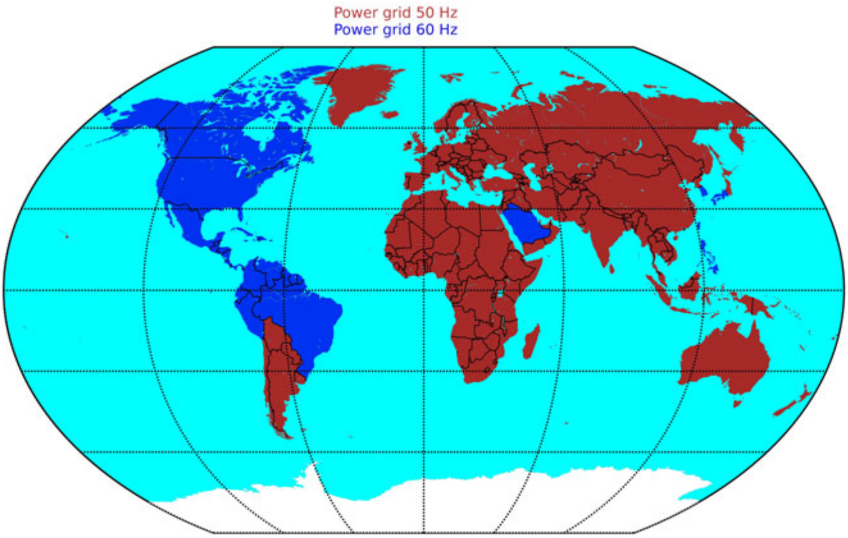


Fig. 11. The world map of electrification systems based on Hughes (1983) and Neidhofer (2011). The brown colour indicates the area where power lines use a frequency equal to 50 Hz and the navy colour indicates the areas powered by 60 Hz.

by the station's filter. The lower panel in Fig. 10 depicts the timeline of the $RLB60_{NS}$ and $RLB60_{EW}$ indices during the years 2005-2014. Any significant linear trend in time domain is observed for either index, but seasonal variations are present. In winter both indices are higher than in the summer period but always index $RLB60_{NS}$ is bigger than $RLB60_{EW}$. Index $RLB60_{NS}$ is, on average, 6.9 times larger than $RLB60_{EW}$.

It is known that different countries use different voltages and one of two values for the frequency of the power grid. The use of frequencies in the power systems of the World is presented in Fig. 11. The national power grids around the globe operate at frequencies equal to 50 or 60 Hz (Hughes 1983, L'Abbate *et al.* 2007, Neidhofer 2011).

The minor spectral lines at 16.66 Hz are poorly visible (see Fig. 7). The variability of the AN16, BG16, and RLB16 indices for both antennae during the study period is presented in Fig. 12. The timeline of the AN16 index from both antennae reveals no clear seasonality. During the period 2007-2009, $AN16_{EW}$ was systematically greater than $AN16_{NS}$, while for the rest of the period the two are approximately equal. Also in this case, the timelines of both BG16 varieties have the same origin and a nearly identical nature to their BG50 counterparts.

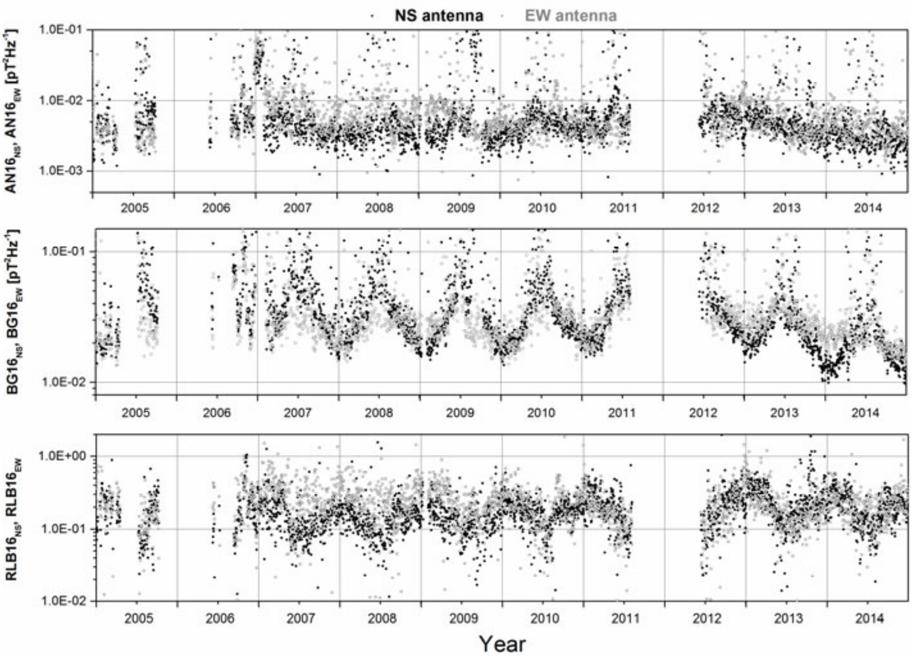


Fig. 12. The variability of indices: AN16 (top panel), BG16 (middle panel), and RLB16 (bottom panel) at both antennae during the years 2005-2014. The results are not available in certain periods due to a technical problem.

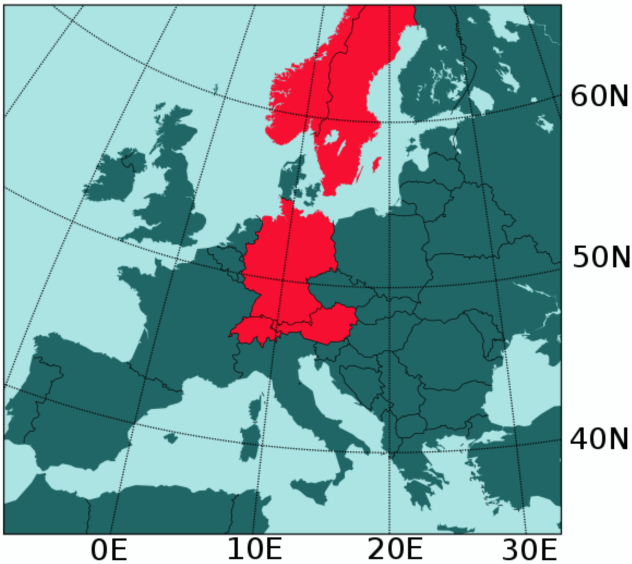


Fig. 13. A map of the electrification systems for European railways (based on Frey 2012). European electric rail tractions operate at DC voltage and AC 50 Hz voltage (blue) as well as AC 16.66 Hz (red).

The seasonal variability of both indices, $RLB16_{NS}$ and $RLB16_{EW}$, is observed. Again, both indices are higher in winter than in summer. In the period 2007-2009, $RLB16_{EW}$ is slightly greater than $RLB16_{NS}$ on a permanent basis. In the rest of the period, both indices are comparable. Generally, the amplitudes of the line spectrum 16.66 Hz in a dataset collected at the Hylaty station are small and vary in time. It is known that this line comes from several European railway systems: Germany, Austria, Switzerland, Sweden, and Norway (Frey 2012). Figure 13 shows a map of the electrification systems of railways in Europe. It is noticeable that the Austrian railway network, with that frequency system, is close to that in the southeast corner of Poland where the Hylaty ELF station is located.

5. CONCLUSIONS

The amplitudes of all anthropogenic spectral lines (16.66, 50, and 60 Hz) are stable over a long time scale. Unexpected disorders that affect the quality of measurements were not observed. This means that the Bieszczady mountains are a good location for this type of measurement. The regular seasonal intermittency is visible in almost all BG, AN50, AN60, and RLB indices.

However, the anthropogenic lines are narrow spectral lines but could disturb the observed natural power spectrum of Schumann resonances. When there is an enormous intensity of these lines, the analysis of Schuman resonances is impossible. Generally, the index $AN50_{EW}$ is greater than $AN60_{NS}$, whereas the index $AN60_{NS}$ is higher than $AN60_{EW}$ in all study periods.

In the Bieszczady mountains the relatively low amplitude of the spectral lines 50 and 60 Hz permits us to study around 6-7 peaks of Schumann resonance effectively. Because the train power line 16.66 Hz is not strong, analysis of the 2nd (~14 Hz) peak of Schuman resonance is also possible at all times. The problem with the huge amplitude of the train power lines is demonstrable in those regions where railways are supplied via 16.66 Hz, which includes Germany, Switzerland, Austria, Sweden, Norway and the immediate borders of these countries (see Fig. 13). For example, this problem exists in ELF recordings in the Nagycenk (Hungary) observatory close to the Austrian border. In consequence, special stop-band filters have been used in the recording system (Sátori *et al.* 1996).

Intensive human activity imposes strong constraints on the spatial opportunities for locating stations. It is clearly visible that the correct choice of location for an ELF station is important and has a strong influence on the quality of magnetic measurements for geophysical purposes.

Acknowledgments. The paper was partially financed by the National Science Centre (NCN, Poland) grants NCN-2012/04/M/ST10/00565 and N N306 039040, and also grant financed by the Jagiellonian University No. WFAIS-FOCUS 139/F/ZN/2016.

References

Balser, M., and C.A. Wagner (1960), Observations of Earth-ionosphere cavity resonances, *Nature* **188**, 4751, 638-641, DOI: 10.1038/188638a0.

Beggan, C.D., T. Gabillard, A. Swan, S. Flower, and A. Thomson (2012), Investigation of global lightning using Schumann resonances measured by high frequency induction coil magnetometers in the UK. **In:** *AGU Fall Meeting, Lightning and Atmospheric Electricity in Thunderstorms V, San Francisco, USA*, AE23B-0333.

Bezrodny, V., O. Budanov, A. Koloskov, M. Hayakawa, V. Sinitsin, Y. Yampolski, and V. Korepanov (2007), The ELF band as a possible spectral window for seismo-ionospheric diagnostics, *Sun Geosphere* **2**, 2, 88-95.

Bösinger, T., C. Haldoupis, P.P. Belyaev, M.N. Yakunin, N.V. Semenova, A.G. Demekhov, and A.V. Angelopoulus (2002), Spectral properties of the ionospheric Alfvén resonator observed at a low-latitude station (L = 1.3), *J. Geophys. Res.* **107**, A10, SIA4-1 – SIA4-9, DOI: 10.1029/2001JA005076.

Christian, H.J., R.J. Blakeslee, D.J. Boccippio, W.L. Boeck, D.E. Buechler, K.T. Driscoll, S.J. Goodman, J.M. Hall, W.J. Koshak, D.M. Mach, and M.F. Stewart (2003), Global frequency and distribution of lightning as observed from space by the optical transient detector, *J. Geophys. Res.* **108**, D1, 4005, DOI: 10.1029/2002JD002347.

Dyrda, M., A. Kulak, J. Mlynarczyk, M. Ostrowski, J. Kubisz, A. Michalec, and Z. Nieckarz (2014), Application of the Schumann resonance spectral decomposition in characterizing the main African thunderstorm center, *J. Geophys. Res. Atmos.* **119**, 23, 13338-13349, DOI: 10.1002/2014JD022613.

Engebretson, M.J., M.R. Lessard, J. Bortnik, J.C. Green, R.B. Horne, D.L. Detrick, A.T. Weatherwax, J. Manninen, N.J. Petit, J.L. Posch, and M.C. Rose (2008), Pc1-Pc2 waves and energetic particle precipitation during and after magnetic storms: Superposed epoch analysis and case studies, *J. Geophys. Res*. **113**, A1, A01211, DOI: 10.1029/2007JA012362.

Fraser-Smith, A.C., and R.A. Helliwell (1994), Overview of the Stanford University/ Office of Naval Research ELF/VLF radio noise survey. **In**: J.M. Goodman (ed.), *Proc. 1993 Ionospheric Effects Symposium, SRI International, Arlington, Virginia*, 502-509.

Frey, S. (2012), *Railway Electrification System and Engineering*, White Word Publications.

Fullekrug, M. (2005), Detection of thirteen resonances of radio waves from particularly intense lightning discharges, *Geophys. Res. Lett.* **32**, 13, L13809, DOI: 10.1029/2005GL023028.

Hebden, S.R., T.R. Robinson, D.M. Wright, T. Yeoman, T. Raita, and T.A. Bösinger (2005), Quantitative analysis of the diurnal evolution of ionospheric Alfvén resonator magnetic resonance features and of changing IAR parameters, *Ann. Geophys*. **23**, 5, 1711-1721.

Hobara, Y., N. Iwasaki, T. Hayashida, T. Tsuchiya, E.R. Williams, M. Sera, Y. Ikegami, and M. Hayakawa (2000), New ELF observation site in Moshiri, Hokkaido, Japan, and the results of preliminary data analysis, *J. Atmos. Electr.* **20**, 2, 99-109.

Hughes, T.P. (1983), *Networks of Power: Electrification in Western Society 1880-1930*, The Johns Hopkins University Press, Baltimore.

Jankowski, J., and C. Sucksdorff (1996), *Guide for Magnetic Measurements and Observatory Practice*, International Association of Geomagnetism and Aeronomy, Warsaw, Poland.

Kangas, J., A. Guglielmi, and O. Pokhotelov (1998), Morphology and physics of short-period magnetic pulsations, *Space Sci. Rev.* **83**, 3, 435-512, DOI: 10.1023/A:1005063911643.

Kulak, A., S. Zieba, S. Micek, and Z. Nieckarz (2003), Solar variations in extremely low frequency propagation parameters: 1. A two-dimensional telegraph equation (TDTE) model of ELF propagation and fundamental parameters of

Schumann resonances, *J. Geophys. Res*. **108**, A7, 1270, DOI: 10.1029/2002JA009304.

Kulak, A., J. Mlynarczyk, S. Zieba, S. Micek, and Z. Nieckarz (2006), Studies of ELF propagation in the spherical shell cavity using a field decomposition method based on asymmetry of Schumann resonance curves, *J. Geophys. Res.* **111**, A10, A10304, DOI: 10.1029/2005JA011429.

Kulak, A., J. Kubisz, S. Klucjasz, A. Michalec, J. Mlynarczyk, Z. Nieckarz, M. Ostrowski, and S. Zieba (2014), Extremely low frequency electromagnetic field measurements at the Hylaty station and methodology of signal analysis, *Radio Sci*. **49**, 6, 361-370, DOI: 10.1002/2014RS005400.

L'Abbate, A., G. Fulli, F. Starr, and S.D. Peteves (2007), Distributed power generation in Europe: Technical issues for further integration, *JCR Sci. Tech. Reports*, European Commission, EUR 23234.

Love, J., and A. Chulliat (2013), An international network of magnetic observation, *EOS* **94**, 42, 373-384.

Neidhofer, G. (2011), 50-Hz frequency [history]: how the standard emerged from a European jungle, *IEEE Power Energy Mag.* **9**, 4, 66-81, DOI: 10.1109/MPE.2011.941165.

Neska, M., and G. Sátori (2006), Schumann resonance observation at Polish Polar Station at Spitsbergen and in Central Geophysical Observatory in Belsk, Poland, *Prz. Geofiz*. **3-4**, 189-198 (in Polish).

Nieckarz, Z., S. Zięba, A. Kułak, and A. Michalec (2009), Study of the periodicities of lightning activity in three main thunderstorm centers based on Schumann resonance measurements, *Month. Weath. Rev.* **137**, 12, 4401-4409, DOI: 10.1175/2009MWR2920.1.

Odzimek, A., A. Kulak, A. Michalec, and J. Kubisz (2006), An automatic method to determine the frequency scale of the ionospheric Alfven resonator using data from Hylaty station, Poland, *Ann. Geophys*. **24**, 8, 2151-2158.

Price, C., M. Finkelstein, B. Starobinets, and E. Williams (1999), A new Schumann resonance station in the Negev desert for monitoring global lightning activity. **In:** *Proc. 11th Int. Conf. on Atmospheric Electricity, 7-11 June 1999, Guntersville, Alabama*, NASA, Marshall Space Flight Center, Alabama, 695-697.

Salazar, A. (2006), Energy propagation of thermal waves, *Eur. J. Phys.* **27**, 6, 1349-1355, DOI: 10.1088/0143-0807/27/6/009.

Sátori, G., J. Szendroi, and J. Vero (1996), Monitoring Schumann resonances – I. Methodology, *J. Atmos. Sol. Terr. Phys*. **58**, 13, 1475-1481, DOI: 10.1016/0021-9169(95)00145-X.

Sátori, G., E.R. Williams, and V. Mushtak (2005), Response of the Earth-ionosphere cavity resonator to the 11-year solar cycle in X-radiation, *J. Atmos. Sol. Terr. Phys*. **67**, 6, 553-562, DOI: 10.1016/j.jastp.2004.12.006.

Sátori, G., M. Neska, E. Williams, and J. Szendroi (2007), Signatures of the day-night asymmetry of the Earth-ionosphere cavity in high time resolution Schumann resonance records, *Radio Sci.* **42**, 2, RS2S10, DOI: 10.1029/2006RS003483.

Schumann, W.O. (1952), On the free oscillation of a conducting sphere, which is surrounded by an air layer and an ionospheric shell, *Z. Naturforsch.* **7a**, 149-154 (in German).

Semenova, N.V., and A.G. Yahnin (2008), Diurnal behaviour of the ionospheric Alfven resonator signatures as observed at high latitude observatory Baentsburg, *Ann. Geophys*. **26**, 8, 2245-2251.

Shalimov, S., and T. Bösinger (2008), On distant excitation of the ionospheric Alfvén resonator by positive cloud-to-ground lightning discharges, *J. Geophys. Res.* **113**, A2, A02303, DOI: 10.1029/2007JA012614.

Surkov, V.V., Y. Matsudo, M. Hayakawa, and S.V. Goncharov (2010), Estimation of lightning and sprite parameters based on observation of sprite-producing lightning power spectra, *J. Atmos. Sol. Terr. Phys*. **72**, 5-6, 448-456, DOI: 10.1016/j.jastp.2010.01.001.

Turbitt, C., J. Matzka, J. Rasson, B. St-Louis, and D. Stewart (2013), An instrument performance and data quality standard for intermagnet one-second data exchange. **In:** *Proc. 15th IAGA Workshop on Geomagnetic Observatory Instruments, Data Acquisition and Processing*, No. 03/13, 186-188.

Williams, E.R. (1992), The Schumann resonance: A global tropical thermometer, *Science* **256**, 5060, 1184-1187, DOI: 10.1126/science.256.5060.1184.

Williams, E.R. (2005), Lightning and climate: A review, *Atmos. Res.* **76**, 1-4, 272-287, DOI: 10.1016/j.atmosres.2004.11.014.

Received 20 January 2016
Received in revised form 13 May 2016
Accepted 15 June 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2609-2629
DOI: 10.1515/acgeo-2016-0083

Extremely Cold Winter Months in Europe (1951-2010)

Robert TWARDOSZ[1], Urszula KOSSOWSKA-CEZAK[2], and Sebastian PEŁECH[1]

[1]Institute of Geography and Spatial Management, Jagiellonian University, Kraków, Poland; e-mail: r.twardosz@uj.edu.pl

[2]Department of Climatology, Faculty of Geography and Regional Studies, Warsaw University, Warszawa, Poland

Abstract

Investigation of extreme thermal conditions is important from the perspective of global warming. Therefore, this study has been undertaken in order to determine the frequency, timing and spatial extent of extremely cold months in winter time at 60 weather stations across Europe over a sixty-year period from 1951 to 2010. Extremely cold months (ECMs) are defined as months in which the average air temperature is lower than the corresponding multi-annual average by at least 2 standard deviations. Half of all the ECMs occurred in the years 1951-1970 (33 out of 67). The lowest number of ECMs was recorded in the decade 1991-2000, but since the beginning of the 21st century, their density and territorial extent has started to increase again. The extremely cold months with ECMs of the greatest spatial extent, covering at least one third of the stations (over 20 stations), included: February 1954 (22), February 1956 (36), January 1963 (25), and January 1987 (23 stations).

Key words: extremes temperature, temperature anomaly, winter seasons, Europe.

1. INTRODUCTION

Winter weather conditions, including air temperature, over the European land mass are chiefly determined by the North Atlantic Oscillation. During

---

Ownership: Institute of Geophysics, Polish Academy of Sciences;

the positive phase of the NAO, relatively warm air from over the Atlantic flows in over Europe, whereas during its negative phase there is advection of cold, often freezing, Arctic air from the north or continental polar air from the east (Hirschi and Sinha 2007, Cattiaux *et al.* 2010, Wang *et al.* 2010, Ouzeau *et al.* 2011, Buchan *et al.* 2014). Longer spells of such advection, which are usually caused by the build-up of high-pressure blocking systems, lead to periods with anomalously low air temperatures in various parts of the continent (*e.g.*, Kossowska-Cezak 1997, Jaagus 2006, Bardin 2007, Isayev and Sherstyukov 2008, Sidorenkov and Orlov 2008, Ugryumov and Khar'kova 2008, Van den Besselaar *et al.* 2010). Such periods of strong declines in temperature have a number of effects on human life and activity and disturb the cycle of natural processes (Błażejczyk and McGregor 2007, Maignan *et al.* 2008). Descriptions of such freezing winters can be found in historical chronicles with accounts of the ice-bound Baltic Sea and rivers frozen down to the bottom or trees cracking in frost (Girguś and Strupczewski 1965). The most noteworthy winters in more recent times include three winters in the early 1940s which had a crucial effect on WWII. During one of them, in January 1942, the temperature on the Eastern Front dropped to –56°C (Brönnimann 2005).

Sometimes, extremely low temperatures last throughout the winter (Twardosz and Kossowska-Cezak 2016), and even continue until March and beyond, as during the extremely cold winter of 1928/29 in Poland (Gumiński 1931). Yet, extremely cold months tend to be much more frequent than ECWs. Therefore, the goal of the study is to determine the frequency and timing of extremely cold winter months (ECMs), as well as the areas where the respective extremely cold winter months were recorded in Europe between the mid-20th century and 2010. The study is a continuation of the authors' research into extremely cold winters in Europe (Twardosz and Kossowska-Cezak 2016). Literature includes other studies on the topic, *e.g.*, Baur (1954), Graham *et al.* (2006), Hirschi and Sinha (2007), Hirschi (2008), Cattiaux *et al.* (2010), Wang *et al.* (2010), Ouzeau *et al.* (2011), and Buchan *et al.* (2014), yet they concern different areas of Europe and different periods.

2. DATA AND METHODS

The study is based on average monthly values (between December and February) of air temperature in the years 1951-2010 recorded at 60 weather stations on the European continent and the British Isles (Table 1). For the most part, they are stations located in lowland areas up to 300 m a.s.l. As with earlier studies by the authors (Twardosz and Kossowska-Cezak 2013a, b; 2015a, b) the data for the research was obtained from the generally available European Climate Assessment & Dataset (ECA&D, www.eca.knmi.nl)

Table 1

Long-term average temperatures in winter months and the numbers of exceptionally cold months (ECMs) in Europe (1951-2010)

Station		T_{av} [°C]			No. of ECMs
No.	Name	Dec	Jan	Feb	
$\varphi < 40°N$					
1	Lisbon	12.0	11.3	12.2	3
2	Almeria	13.4	12.4	13.0	4
3	Crotone	10.5	9.3	9.5	5
4	Athens	12.0	10.3	10.6	5
$\varphi = 40$-$45°N$					
5	La Coruña	11.1	10.4	10.6	3
6	Madrid	6.5	6.0	7.5	6
7	Bordeaux	6.5	5.8	6.7	7
8	Barcelona	9.8	9.0	9.7	3
9	Marseilles	7.7	6.7	7.7	7
10	Rome	8.6	7.5	8.3	5
11	Split	9.2	7.8	8.2	7
12	Belgrade	2.7	0.9	2.9	7
13	Sofia	0.5	–1.2	0.8	6
14	Konstanca	3.2	0.9	1.8	8
15	Istanbul	8.3	6.1	6.3	6
16	Simferopol	2.3	–0.1	0.5	9
17	Sochi	8.2	6.2	6.3	3
18	Makhachkala	2.8	0.6	0.8	7
$\varphi = 45$-$50°N$					
19	Brest	7.2	6.4	6.4	7
20	Paris	5.2	4.4	5.3	11
21	Zurich	1.0	0.0	1.2	8
22	Würzburg	1.2	0.1	1.2	9
23	Vienna	1.3	–0.1	1.7	7
24	Debrecen	0.2	–1.8	0.3	9
25	Chernivtsi	–1.8	–3.9	–2.4	7
26	Zaporozhe	–1.4	–3.8	–3.1	9
27	Rostov on the Don	–1.4	–3.9	–3.4	5
28	Astrakhan	–2.1	–4.9	–4.6	7
$\varphi = 50$-$55°N$					
29	Valentia	7.8	6.9	6.8	6
30	London	5.3	4.5	4.5	7
31	De Bilt	3.4	2.5	2.8	11
32	Berlin	1.3	0.1	0.8	7

to be continued

Table 1 (continuation)

Station		T_{av} [°C]			No. of ECMs
No.	Name	Dec	Jan	Feb	
33	Warsaw	–0.7	–2.6	–1.8	8
34	Minsk	–3.7	–6.0	–5.3	6
35	Kiev	–2.3	–4.6	–3.8	8
36	Kursk	–5.1	–7.6	–7.4	7
37	Saratov	–6.6	–9.3	–9.3	6
38	Orenburg	–9.8	–13.1	–12.9	6
φ = 55-60°N					
39	Edinburgh	4.1	3.6	3.8	7
40	Oslo	–2.5	–3.7	–3.7	5
41	Copenhagen	2.4	0.8	0.6	6
42	Stockholm	–0.5	–2.3	–2.7	7
43	Liepaja	0.1	–2.0	–2.6	9
44	St. Petersburg	–4.1	–6.6	–6.7	7
45	Moscow	–5.7	–8.1	–7.6	8
46	Vologda	–8.6	–11.5	–10.7	3
47	Kazan	–9.0	–12.0	–11.4	7
48	Yekaterinburg	–11.3	–13.7	–12.3	7
φ = 60-65°N					
49	Bergen	2.7	2.0	1.8	8
50	Trondheim	–1.4	–2.6	–2.3	6
51	Vaasa	–4.6	–6.6	–7.3	7
52	Kajaani	–8.5	–11.1	–11.1	7
53	Arkhangelsk	–9.6	–13.0	–12.0	7
54	Syktyvkar	–11.9	–15.0	–13.4	4
55	Ivdel	–16.0	–18.8	–16.6	6
φ = 65-70°N					
56	Bodö	–0.6	–1.6	–1.9	4
57	Sodankyla	–11.9	–13.9	–13.2	4
58	Naryan-Mar	–13.8	–17.6	–17.3	6
59	Pechora	–15.7	–18.9	–17.5	7
φ > 70°N					
60	Vardö	–3.2	–4.5	–5.0	3
$\sum$					387

(Klein Tank *et al.* 2002). The data in the database offers a very high spatial resolution and a large number of complete, good-quality data series, which are verifiable for homogeneity (Wijngaard *et al.* 2003). The database has been also used by many other researchers (*e.g.*, Cony *et al.* 2008, Van den

Besselaar *et al.* 2010, Moberg *et al.* 2006). For the purposes of the present study, extremely cold months (ECMs) are those during which the average temperature is lower than the multi-annual (1951-2010) average temperature at a given station by at least 2 standard deviations ($t \le t_{av} - 2\sigma$). The authors have previously used the method to investigate extremely cold winter seasons and extremely hot summer seasons and months (Twardosz and Kossowska-Cezak 2013a, b; 2015a, b; 2016). This is a relative method, in which reference is made to the average thermal conditions in a given area. This criterion has been used by many other researchers, including Hansen *et al.* (2012).

3. FREQUENCY OF EXTREMELY COLD WINTER MONTHS

In the years 1951-2010, 387 cases of extremely cold months (ECMs) were recorded across the 60 stations in Europe, which occurred during 67 months of the 60-year period. On average, one ECM occurred at nearly 6 stations (5.8). Extremely mild months, the existence of which was determined by an analogous method, were half as frequent (34 months), and the areas they covered were half as big – 3 stations (3.1) (Kossowska-Cezak *et al.* 2016). As has been shown by the authors (Twardosz and Kossowska-Cezak 2016), the years 1951-2010 saw 103 cases of extremely cold winters. Thus, the ratio of the number of ECMs to ECWs is 3.75, while the same ratio for extremely mild months is merely 2.62. Furthermore, among the 103 cases of extremely cold winters in 19 cases there were no ECMs at all, and in 50 winters there was only one ECM. On the other hand, among the 387 ECMs only 123 occurred during extremely cold winters, and the other 264 ECWs were not connected to such winters. The above figures clearly indicate that individual ECMs are a frequent phenomenon in Europe.

Table 2

Number of exceptionally cold months (ECMs) and the number of stations where ECMs were recorded in Europe (1951-2010)

Period	Number of months	Number of stations
1951/52* – 1959/60	16	99
1960/61 – 1969/70	17	99
1970/71 – 1979/80	9	42
1980/81 – 1989/90	9	85
1990/91 – 1999/2000	7	17
2000/01 – 2009/10	9	45
1951/52 – 2009/10	67	387

*⁾No data from winter 1950/1951.

Table 3

Exceptionally cold winter months (no. of stations) in Europe (1951-2010)

Winter	Dec	Jan	Feb	Σ
1950/51	x	–	1	(1)
1951/52	–	1	–	1
1952/53	–	2	1	3
1953/54	3	11	22	36
1954/55	–	–	1	1
1955/56	12	–	36	48
1956/57	1	3	–	4
1957/58	1	–	–	1
1958/59	2	–	1	3
1959/60	1	–	–	1
...				
1962/63	6	25	7	38
1963/64	9	6	–	15
1964/65	–	–	4	4
1965/66	1	2	11	14
1966/67	1	–	–	1
1967/68	1	4	–	5
1968/69	2	8	2	12
1969/70	9	–	2	11
1970/71	3	–	–	3
1971/72	–	10	–	10
...				
1975/76	–	1	3	4
1976/77	1	2	–	3
1977/78	–	–	1	1
1978/79	16	4	–	20
...				
1980/81	1	2	–	3
1981/82	7	–	–	7
...				
1984/85	2	19	19	40
1985/86	1	–	11	12
1986/87	–	23	–	23
...				
1990/91	1	–	–	1
1991/92	6	–	–	6
...				
1995/96	1	–	–	1
...				

to be continued

Table 3 (continuation)

Winter	Dec	Jan	Feb	Σ
1997/98	–	–	3	3
1998/99	4	1	–	5
1999/2000	–	1	–	1
...				
2001/02	10	–	–	10
2002/03	14	–	1	15
...				
2005/06	1	2	–	3
...				
2007/08	–	1	–	1
...				
2009/10	1	2	–	3
2010/11	13	x	x	(13)
No. of stations with ECMs	131	130	126	387
No. of ECMs	29	21	17	67

In the six decades, from 1951 to 2010, the frequency of ECMs fluctuated significantly (Table 2). Half of the ECMs (33 out of 67) and over half of their cases (the number of stations where ECMs were recorded was 198 out of 387) were recorded in the first two decades. Those years saw 3 out of the 4 ECMs that were recorded by more than 20 stations (Table 3). In the decades that followed, the number of ECMs remained quite steady (7-9 in 10 years), but the decade 1980/81-1989/90 was distinguished by a relatively high number of cases (85), with another ECM recorded by 20 stations and 2 ECMs by 19 stations. The last decade of the 20th century saw the least number of ECMs, and clearly the lowest number of cases, *i.e.*, merely 17, which means that the ECMs in that decade covered smaller areas. In the first decade of the 21st century, ECMs registered by 10 or more stations started to slowly re-appear. The first decade of the 21st century also saw the greatest number of extremely mild winter months (10 out of 34) and of their cases (40 out of 105). It was the only decade to see fewer ECMs than mild months (9 and 10, respectively), while in the first 2 decades under study there were 4 times more of them (33 and 8, respectively).

The number of ECMs at individual stations varied from 3 to 11, but typically ranged from 6 to 7 (Table 1). The lowest numbers (3-5 per station) were recorded in the southernmost parts of Europe, mainly in the Iberian Peninsula. ECMs were most numerous in the central part of the continent, in the area stretching from the Scandinavian Peninsula to the Balkans, with a wedge reaching westwards as far as France – more than 8 ECMs, with a peak of 11 in Paris and De Bilt.

4. SPATIAL EXTENT AND THERMAL CHARACTERISTICS OF ECMS

A calendar of ECMs is presented in Table 3, showing the year and month, as well as the number of cases, *i.e.*, stations where an ECM occurred. As can be seen in the calendar, a large proportion of the ECMs covered a very small area, only 1-3 stations, that is, not more than 5% of the stations. There were 40 such montfIGhs out of the 67 in the 6 decades. Another 7 ECMs were recorded by not more than 6 stations (10%). These 47 ECMs comprising not more than 10% of the stations will be largely disregarded in the study presented below. A compilation of the remaining 20 ECMs is given in Table 4. It is worth noting that out of the 34 extremely mild winter months (not shown), only 5 covered more than 10% of the stations (December 1960 – 7 stations, February 1990 – 19, February 2002 – 7, December 2006 – 10, and January 2007 – 13 stations).

The first 10-year period of 1951-1960 recorded nearly ¼ of all the ECMs – 16 out of 67, but most of them covered very small areas or only a single station. Only 4 ECMs had a larger extent.

The ECM of January 1954 covered 11 stations in south-central Europe (Table 4). The temperature anomaly Δt ranged between –3.2°C in Split and –4.4°C in Istanbul, and –6 to –8°C in the remaining area, with the greatest anomaly in Zaporizhia, where $\Delta t = -8.7$°C. In Split and Istanbul, the average monthly temperatures were above zero, and the lowest temperature was recorded in Saratov, where $t = -17.1$°C.

The next month of the same year, the ECM of February 1954, covered 22 stations (Fig. 1, Table 4) which were nearly all the same as in previous month plus east and central Europe from Berlin and Vienna to Yekaterinburg and Makhachkala. The temperature anomaly Δt ranged between –4.3°C in Istanbul (which was the only station where t was above zero), –5 to –6°C in the west and –9 to –11°C in the east with the greatest anomalies in Astrakhan $\Delta t = -14$°C (the largest of all the months) and Rostov $\Delta t = -13.6$°C. The lowest average temperature was recorded in Orenburg $t = -24.4$°C. February 1954 was among the coldest months in the 60-year period. More temperature characteristics from selected weather stations are presented in Table 5.

The ECM of December 1955 covered 12 stations in southeastern Europe. A temperature anomaly of $\Delta t = -5$°C was seen in Bodö and Vardø only, with the other stations recording anomalies of more than –10°C, including Syktyvkar $\Delta t = -13.5$°C and Pechora $\Delta t = -13.1$°C. The lowest monthly average temperatures were in Pechora $t = -28.8$°C and in Ivdel $t = -28.7$°C.

Table 4

Exceptionally cold winter months in Europe recorded at 7 or more stations (1951-2010)

Year	Month	No. of stations	Stations (no. according to Table 1)	ECW
1954	Jan	11	11, 14, 15, 16, 24, 25, 26, 27, 35, 36, 37	ECW 1953/54 at 16 stations; 3 ECMs each in Istanbul (15) and Simferopol (16)
	Feb	22	2, 12, **14**, 15, 16, **18**, 23, 24, 25, **26**, **27**, **28**, 32, 33, 34, 35, 36, 37, 38, 43, 47, 48	
1955	Dec	12	44, 45, **46**, 47, 48, 52, 53, 54, 55, 56, 59, 60	ECW 1955/56 at 10 stations
1956	Feb	36	1, 2, 3, **5**, **6**, **7**, **8**, **9**, **10**, **11**, 12, 13, 14, 16, **19**, **20**, **21**, **22**, **23**, 24, 25, 26, 27, **31**, **32**, 33, 34, 35, 36, 37, 41, 43, 44, 45, 47, 48	
1963	Jan	25	7, 11, 12, 13, 14, 19, 20, 21, 22, 23, 24, 25, 26, **29**, **30**, **31**, 32, 33, 34, 35, 36, 39, 41, 45, 49	ECW 1962/63 at 24 stations. 3 ECMs each in Paris (20), Wurzburg (22), and De Bilt (31)
	Feb	7	19, 20, 22, 30, 31, 39, 49	
1963	Dec	9	12, 19, 20, 21, 22, 23, 24, 25, 31	ECW 1963/64 at 3 stations
1966	Feb	11	42, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59	ECW 1965/66 at 7 stations
1969	Jan	8	28, 37, 38, 45, 47, 48, 54, 55	ECW 1968/69 at 7 stations
1969	Dec	9	9, 11, 20, 21, 22, 23, 32, **33**, 43	ECW 1969/70 at 4 stations
1972	Jan	10	16, **18**, 26, 27, **28**, 37, 38, 47, 48, 55	
1978	Dec	16	34, 36, 40, 42, 43, 44, 45, 46, 47, 50, 51, 52, 53, 54, 58, 59	ECW 1978/79 at 6 stations
1981	Dec	7	31, 39, 40, 41, 49, 50, 56	
1985	Jan	19	6, 7, **8**, 9, 10, 20, 21, 22, 23, 30, 31, 41, 42, 43, 51, 52, 53, 57, 58	ECW 1984/85 at 9 stations
	Feb	19	14, 15, 16, 17, 24, 26, 33, 34, 35, 42, 43, 44, 51, 52, 53, 55, 56, 57, 59	
1986	Feb	11	19, 20, 21, 22, 29, 30, 31, 32, 33, 39, 49	
1987	Jan	23	7, 9, 19, 20, 21, 30, 32, 33, 34, 35, 36, 40, 41, 42, 43, 44, 45, **46**, 49, 50, 51, 52, 57	
2001	Dec	10	4, 8, 11, 12, **13**, 14, 24, 25, 26, 35	
2002	Dec	14	16, **18**, 25, 26, 27, **28**, 33, 35, 36, 37, 38, 43, 45, 47	
2010	Dec	13	19, 20, 22, 29, 30, 31, 32, 39, 40, **41**, 42, 49, 50	

Notice: Station numbers printed in bold mean that the average winter temperature at that station met the formula $t \leq t_{av} - 3\sigma$.

Table 5

Thermal characteristic of the extremely cold winter months (ECMs) in Europe (1951-2010)

Station		Temperature [°C]				No. of days with T_{max}		
No.	Name	$T_{av.}$	Δt	T_{max}	T_{min}	<0°C	<–10°C	<–20°C
1954 February								
15	Istanbul	2.0*	–4.3	5.4	–1.3	3	–	–
28	Astrakhan	**–18.6**	–14.0	–12.8	–23.4	28	25	–
35	Kiev	–12.9	–9.1	–9.2	–16.6	27	15	–
47	Kazan	–21.6	–10.2	–17.7	–26.0	28	26	9
1956 February								
1	Lisbon	**7.8***	–4.4	11.8	3.9	–	–	–
20	Paris	**–3.5***	–8.8	0.3	–7.2	14	–	–
23	Vienna	**–8.5***	–10.2	–5.2	–11.7	26	4	–
32	Berlin	**–8.4***	–9.2	–5.1	–12.7	25	3	–
35	Kiev	–13.1*	–9.3	–8.7	–16.5	24	14	1
44	St. Petersburg	–14.8*	–8.1	–10.1	–19.3	27	10	5
45	Moscow	–18.5*	–10.9	–13.9	–22.8	29	19	5
48	Yekaterinburg	–20.1*	–7.8	–15.5	–24.1	29	22	8
1963 January								
20	Paris	–1.6*	–6.0	0.7	–3.9	13	–	–
29	Valentia	**2.3***	–4.6	4.9	–0.3	–	–	–
32	Berlin	–7.3*	–7.4	–4.4	–10.7	24	1	–
41	Copenhagen	–4.5	–5.3	–2.2	–6.9	22	–	–
45	Moscow	–15.9	–7.8	–12.9	–19.5	31	20	2
1978 December								
40	Oslo	–8.0	–5.5	–5.9	–10.3	31	4	–
59	Pechora	–29.6*	–13.9	–25.2	–33.8	31	29	22
1985 January								
10	Rome	4.1*	–3.4	8.1	0.0	3	–	–
53	Arkhangelsk	–25.3*	–12.3	–20.1	–29.7	31	25	18
58	Narjan Mar	–27.3*	–9.7	–22.7	–31.9	29	27	23
1985 February								
16	Simferopol	–8.0*	–8.5	–3.2	–12.1	18	6	–
57	Sodankyla	–25.1*	–11.9	–19.6	–31.1	28	25	10
1987 January								
44	St. Petersburg	–17.9*	–11.3	–15.0	–21.0	31	20	10
49	Bergen	–2.9	–4.9	–0.5	–5.6	15	1	–
2002 December								
18	Machaczkała	**–5.4***	–8.2	–0.9	–9.9	16	1	–
35	Kiev	–8.4*	–6.1	–5.4	–11.4	28	1	–
2010 December								
29	Valentia	**4.2***	–3.6	6.8	1.5	–	–	–
42	Sztokholm	–6.6*	–6.1	–4.2	–9.3	27	1	–

*) The lowest in 60 years;

Notice: A value in bold means that the temperature meets the criterion $t \le t_{av} - 3\sigma$.

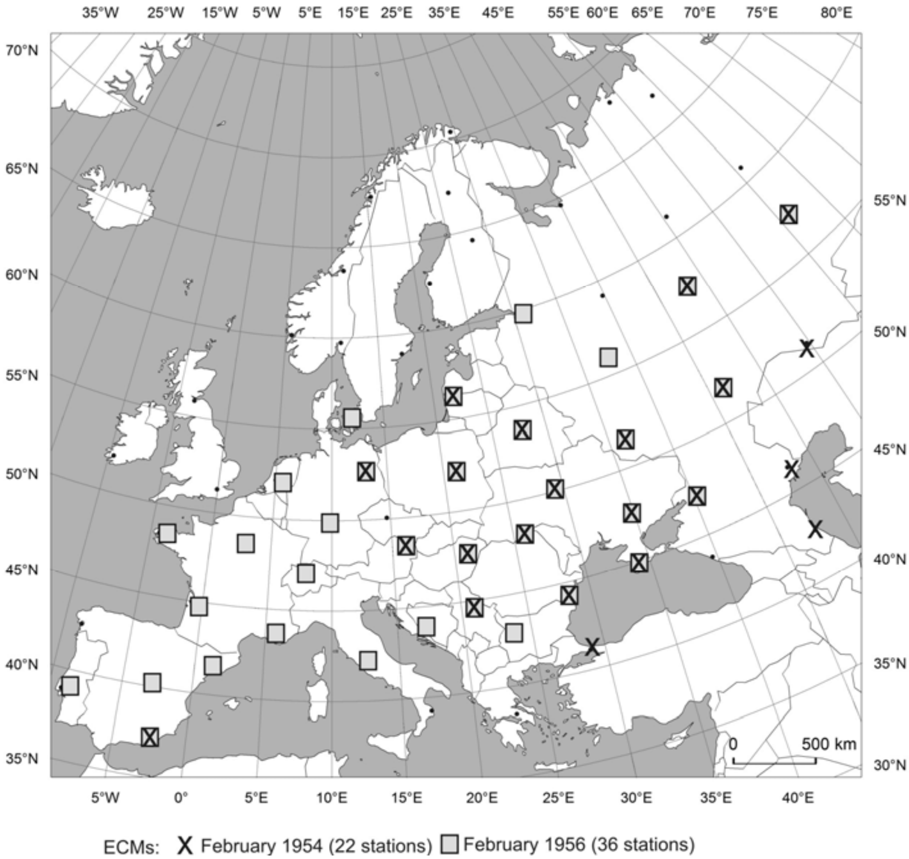


Fig. 1. Stations with extremely cold months: February 1954 (22 stations) and February 1956 (36 stations).

February 1956 was another ECM in the same winter of 1955/56. It was the ECM that extended over the largest number of stations during the 6 decades, *i.e.*, 36 stations in southern and central Europe (Fig. 1, Table 4). In Saint Petersburg, Moscow, Kazan, and Yekaterinburg it was the second ECM during the same winter. It was one of the coldest months in the 6 decades under study. The deviation of average monthly temperature, t, from the multi-annual average, t_{av}, in the ECM of February 1956 exceeded 3 standard deviations at 14 stations. A temperature anomaly Δt of approx. –3 to –6°C was only recorded in the westernmost and southernmost areas (average temperature t above zero), while in the remaining area Δt was from –8 to –11°C, with the greatest value in Saratov, $\Delta t = -11.3$°C, where the lowest temperature, $t = -20.6$°C, was also recorded.

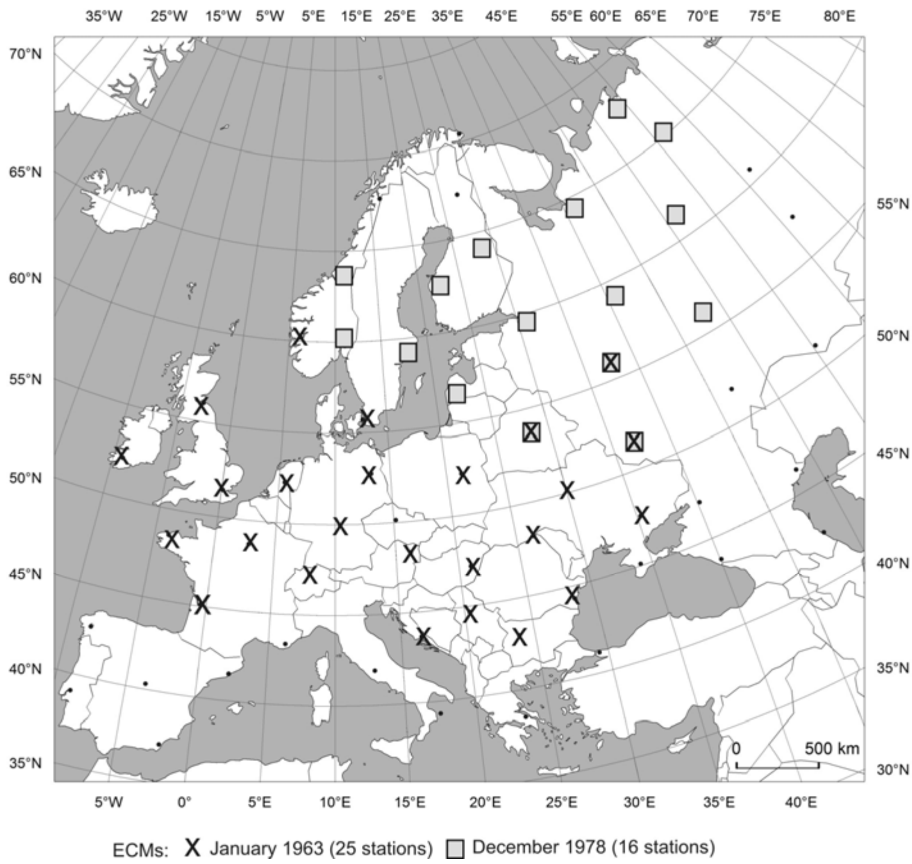


Fig. 2. Stations with extremely cold months: January 1963 (25 stations) and December 1978 (16 stations).

In the next decade, 1960/61 − 1969/70, ECMs were as frequent as in the previous one (17 ECMs) yet with more ECMs covering a greater territory.

The ECM of January 1963 covered 25 stations (in 6 stations December 1962 was also an ECM) in western and central Europe (Fig. 2, Table 4). In most of the stations, it was the coldest January in the 60-year period and in many the coldest winter month. The temperature anomaly Δt ranged between –4 and –6°C in the west and south (overall, t in the westernmost areas was above zero), and in the east between –8 and –9°C, including the greatest anomaly in Warsaw, $\Delta t = -9.8$°C. The lowest mean temperature was recorded in Kursk, $t = -16.8$°C. The lower negative air temperature anomalies in winter in Western Europe are attributable to the mitigating effect of the waters of the Atlantic Ocean (Hirschi and Sinha 2007), as opposed to the cooled land in central, and even more so, Eastern Europe, which is conducive to significant falls in air temperature.

The ECM of February 1963 was recorded by 7 stations of northwestern Europe and was a second ECM in the same winter. The temperature anomaly Δt was approx. from –4 to –6°C. The lowest mean temperature was in Würzburg, $t = -5.4$°C.

The winter of 1962/63 was extremely cold at 24 stations in Western Europe. The winter was among those with the largest extent and one of the two coldest (in addition to the winter of 1955/56) in the 60-year period, and in the southernmost and westernmost areas of the continent – the only extremely cold winter during the 6 decades.

The ECM of December 1963 occurred at 9 stations in a band running between Brest and De Bilt in the northwest and Belgrade and Chernivtsi in the southeast. The temperature anomaly Δt was between *ca.* –3°C in the westernmost part (here temperature t was above 0°C, in Brest $t = 4.2$°C) and –5°C in the east, including Vienna $\Delta t = -5.7$°C. The lowest temperature, in Chernivtsi, was $t = -6.6$°C.

January 1964 was an ECM in 6 stations and February 1965 in 4 stations in the south of Europe. A larger extent was that of the ECM of February 1966, which was registered by 11 stations on the Scandinavian Peninsula and in northwest Russia. The temperature anomaly Δt was between *ca.* –7°C on the coast of the Scandinavian Peninsula and –12 – –13°C in the northwestern areas covered by that ECM. The highest, $\Delta t = -12.9$°C, and the lowest temperature of all the months under study, $t = -30.2$°C, was recorded in Naryan-Mar, and not much higher in Pechora, $t = -29.8$°C.

Again, 4 stations within the same area that recorded the ECM of January 1968 (Δt from ~ –9 to –11°C), and another ECM in January 1969 at 8 stations in southeastern Europe, including, as in February 1966, Syktyvar and Ivdel. The temperature anomaly Δt was from –8 to –11°C, with $\Delta t = -11.7$° in Orenburg. The lowest average monthly temperature, $t = -27.1$°C, was recorded in Ivdel, the lowest $\Delta t = -13.0$°C in Astrakhan. In the easternmost areas, it was the coldest January in the 60-year period.

In the same year, December 1969 was an ECM covering 9 stations in Western Europe. The temperature anomaly in Marseilles, Split and Paris Δt was between –3 and –4°C (with the average temperature above zero, up to 6°C in Split) and –6 and –7°C in the east, with the greatest anomaly in Warsaw, $\Delta t = -7.8$°C; Warsaw also recorded the lowest temperature, $t = -8.5$°C. It was the coldest December in the years 1951–2010 in Zurich, Berlin, and Warsaw.

The decade that followed was characterised by far fewer ECMs. There were 9 of them, including only two ECMs recorded by more than 10% of the stations. This included the ECM of January 1972, which occurred at 10 stations in southeastern Europe. Despite the rather vast area, Δt was quite uniform, ranging between approx. –8 and –10°C, Simferopol being the only exception, with $\Delta t = -7.1$°C (and with the highest $t = -7.0$°C). The greatest anomalies were recorded in Astrakhan, $\Delta t = -10.6$°C, and Yekaterinburg, $\Delta t = -10.5$°C. The lowest average temperature was recorded in Ivdel $t = -28.2$°C. Here, as well as in some other stations in the east, the coldest January of the six decades under study was recorded.

The ECM of December 1978 comprised 16 stations in the Scandinavian Peninsula and in northeastern Europe (Fig. 2, Table 4). The temperature anomaly Δt in the westernmost and southernmost areas of ECMs ranged from ~ –5 to –7°C (the lowest in Stockholm $\Delta t = -4.9$°C, where the highest $t = -5.4$°C was also registered). The anomaly in the north was more than –10°C, the greatest in Pechora, $\Delta t = -13.9$°C, which also saw the lowest temperature, $t = -29.6$°C. For some stations it was the coldest December in the six decades. In the north, the winter of 1978/79 was extremely cold throughout (Twardosz and Kossowska-Cezak 2016).

The 1980s also saw 9 ECMs, but out of these 5 covered more than 10% of the stations. The smallest area, 7 stations, was that of the ECM of December 1981, which occurred in northwestern Europe, including the western areas of the Scandinavian Peninsula. The temperature anomaly Δt was between approx. –4 and –5°C in the southern areas covered by the ECM (De Bilt, Copenhagen) and –6 and –7°C in the north, with the greatest anomaly in Trondheim, $\Delta t = -7.3$°C. It was the coldest period in the six decades for the coast of the Norwegian Sea, with the lowest average temperature $t = -9.0$°C being recorded in Oslo.

January 1985 and February 1985 were ECMs during the winter of 1984/85, which extended over large areas – 19 stations each (Fig. 3, Table 4). The ECM of January 1985 covered southwestern Europe, the Baltic countries and the northernmost and easternmost areas. The temperature anomaly Δt ranged from –2 to –4°C in the south and west, and –5 to –7°C in the central part. The northernmost and easternmost areas had a Δt of ~ –10 to –12°C, with the highest anomaly in Kajaani $\Delta t = -13.2$°C. The average monthly temperature on the coast of the Mediterranean Sea was above zero, and below –20.0°C in the northeast, with the lowest temperature in Naryan Mar, $t = -27.3$°C. Naryan Mar also saw the coolest January in the 60-year period.

The ECM of February 1985 occurred in the same 6 stations as the previous month on the Baltic Sea, in Finland and Arkhangelsk, reaching beyond them to the south of the above area, across Poland and Belarus, as far as the coast of the Black Sea (Fig. 3, Table 4), as well as to the north to Pechora and Ivdel. The temperature anomaly of Δt was only lower than –5°C in Bodö, Istanbul and Sochi, and ranged between approx. –8 and –11°C at the

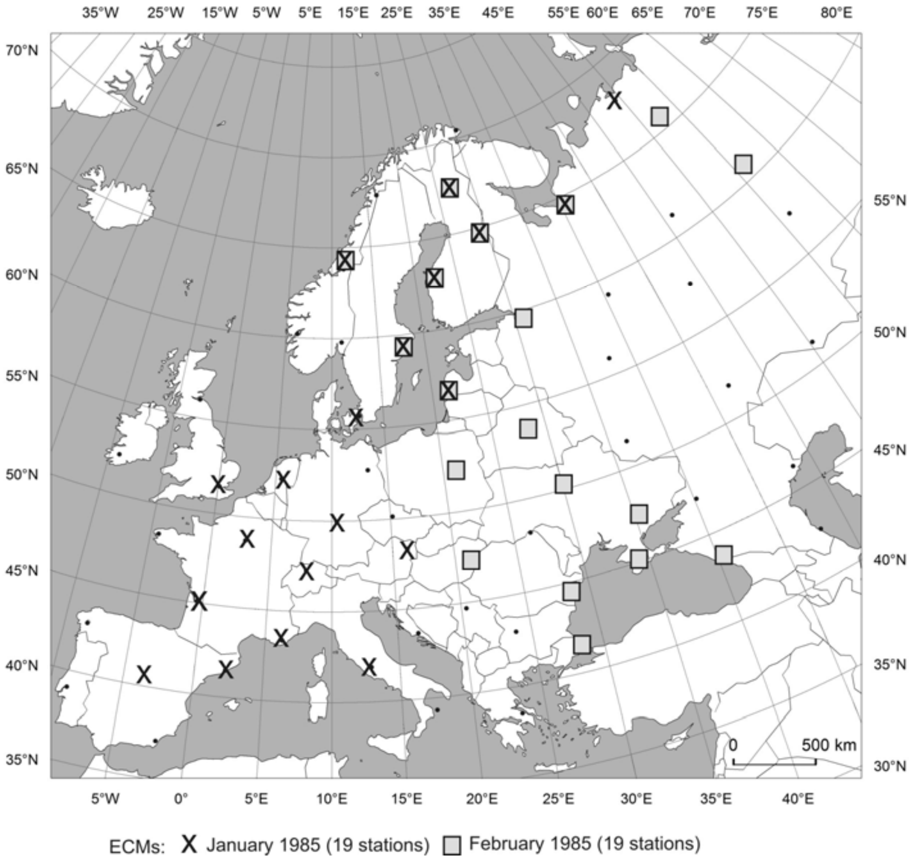


Fig. 3. Stations with extremely cold months: January 1985 (19 stations) and February 1985 (19 stations).

other stations, with the greatest anomaly, $\Delta t = -11.9$°C, in Sodankylä. Temperatures were above zero in Istanbul and Sochi (but in Simferopol $t = -8.0$°C), and below –20ºC in the north, with the lowest temperature in Pechora, $t = -28.5$°C. At some individual stations, including Istanbul and Saint Petersburg, it was the coldest February in the 60-year period.

In the following year, the ECM of February 1986 also occurred. It spanned 11 stations in Western Europe. The temperature anomaly in the westernmost parts (Valentia, Edinburgh, Bergen), showed a Δt from ~ –4°C with up to –6 – –7°C in the rest of the area and with the greatest anomaly, $\Delta t = -7.4$°C, in Warsaw. Warsaw also recorded the lowest temperature $t = -9.6$°C. The temperature t in Valentia and Brest was above 0°C.

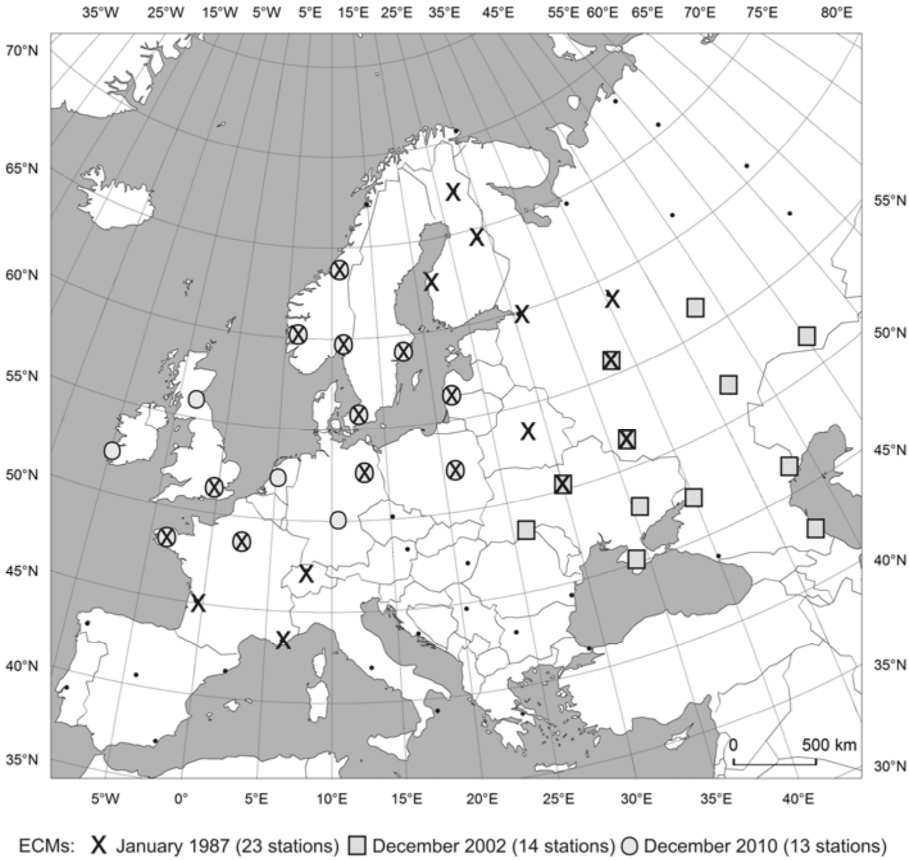


Fig. 4. Stations with extremely cold months: January 1987 (23 stations), December 2002 (14 stations), and December 2010 (13 stations).

The ECM of January 1987 was one of the most extensive ECMs in the 1951-2010 period, stretching over 23 stations from the western coast of Europe to Moscow and the Scandinavian Peninsula (Fig. 4, Table 4). The temperature anomaly Δt in the westernmost coastal areas was approx. –4 – –5°C (average temperature t above zero). In the interior of the continent Δt was ~–7 to –10°C with the highest, $\Delta t = -13.0$°C, in Vologda (the latter also recorded the lowest average temperature, $t = -24.5$°C) and Vaasa $\Delta t = -12.0$°C. In the central part of the area covered by the ECM, it was the coldest month in the 60-year period.

The last decade in the 20th century saw 7 ECMs, but none of them covered an area with more than 10% of the stations.

In the first decade of the 21st century the frequency of ECMs started to increase slightly again. There were 9 of them, including 3 with a greater extent than in the previous decade.

The ECM of December 2001 was recorded by 9 stations in south-central Europe. The temperature anomaly Δt in the southern stations (Athens, Split) was ~ –3°C (with temperature t above 5°C), and inside the continent it was approx. –5°C, with the largest anomaly in Sofia, $\Delta t = -6.3$°C. The lowest average temperature in Kiev was $t = -7.3$°C.

The next ECM of December 2002 spanned a vast area of southeastern Europe with 14 stations (Fig. 4, Table 4). Temperature anomaly Δt in the west of the area was between –5 and –6°C, in the east it was between –8 and –9°C, with the greatest anomaly occurring in Orenburg, $\Delta t = -9.7$°C, where the lowest average monthly temperature was also recorded at $t = -19.5$°C. In Makhachkala and Astrakhan the deviation of average monthly temperature t from the multi-annual average exceeded 3 standard deviations. At most of the stations (except for the westernmost areas), it was the coldest December in the six decades.

The last month of the 60-year period, December 2010 was an ECM at 13 stations in northwestern Europe (Fig. 4, Table 4). The temperature anomaly Δt in the west was approx. –3 to –4°C (with an average temperature t above zero) up to from –6 to –7°C in the Scandinavian Peninsula, with the greatest anomaly, $\Delta t = -7.3$°C, in Trondheim. In Valentia, Edinburgh and Copenhagen the deviation of average monthly temperature t from the 60-year average was higher than 3 standard deviations. The lowest temperature was $t = -9.2$°C in Oslo. For nearly all of the stations, it was the coldest December in the 60 years.

After 1986, all the ECMs were individual months, not related to extremely cold winter periods.

5. CONCLUSIONS

Extremely cold months (ECMs), understood as months with an average air temperature lower than the respective multi-annual average by at least 2 standard deviations, are quite frequent in Europe. During the 60-year period (1951-2010) there were 67 such months across the continent. However, most of the ECMs were limited in territorial extent to 1-3 stations out of the 60 stations included in the study, an extent that was recorded in as many as 40 ECMs. ECMs covering larger areas, more than 6 stations (*i.e.*, over 10% of the stations), only occurred in 20 months.

Overall, the greatest frequency of ECMs was recorded in the first two decades: half of all the ECMs (33) and over half of the cases (198 out of 387) were recorded before the 1969/70 season. The same years saw half of the ECMs with an extent of more than 10% of the stations, including the 3

ECMs with the greatest extent comprising over 20 stations: February 1954 (22 stations), February 1956 (36 stations), and January 1963 (25 stations). In the years that followed, ECMs occurred half as frequently, but their territorial extent was still significant in the decade of 1980/81−1989/90: January 1987 (23 stations) and January and February 1985 (19 stations each). ECMs were far less frequent in the last decade of the 20th century – only 7 months and 17 cases, but their frequency and range was on the rise again in the 21st century – in December 2001, 10 stations recorded the first ECM after 15 years, and for many stations it was the coldest December in the 60-year period. At the same time, the early 21st century experienced the highest frequency of extremely mild winter months in the entire six-decade period. Therefore, it can be concluded that despite the generally observable climatic warming, extremely cold winter months are still possible.

At the same time, it must be stressed that this applies to single months, and not whole winter periods, since only 32% of all the ECMs occurred during extremely cold winters, and the number of ECMs exceeded the number of such winters 3.75 times. Such relatively short-lasting and large drops in air temperature in winter are recorded across Europe, but are most frequent in central and northern parts of the continent. The greatest negative temperature anomalies at such times are recorded in eastern, and even more so in northeastern Europe where the average monthly temperature during ECMs may be lower than the average multi-annual temperature by 11-14°C.

Acknowledgements. We thank Mr. Paweł Pilch and Dr. Martin Cahn for reviewing the English.

References

Bardin, M.Yu. (2007), Anticyclonic quasi-stationary circulation and its effect on air temperature anomalies and extremes over western Russia, *Russ. Meteorol. Hydrol.* **32**, 2, 75-84, DOI: 10.3103/S106837390702001X.

Baur, F. (1954), Die Bestätigung bisheriger Ergebnisse der Gro\wetterforschung durch den Winter 1953/54, *Archiv. Met. Geoph. Biokl. A* **7**, 1, 188-198, DOI: 10.1007/BF02277915 (in German).

Błażejczyk, K., and G. McGregor (2007), Bio-thermal conditions and mortality in selected European agglomerations, *Prz. Geogr*. **79**, 3-4, 401-423 (in Polish).

Brönnimann, S. (2005), The global climate anomaly, 1940-1942, *Weather* **60**, 12, 336-342, DOI: 10.1256/wea.248.04.

Buchan, J., J.J.-M. Hirschi, A.T. Blaker, and B. Sinha (2014) North Atlantic SST anomalies and the cold North European weather events of winter 2009/10 and December 2010, *Mon. Wea. Rev.* **142**, 922-932, DOI: http://dx.doi.org/10.1175/MWR-D-13-00104.1.

Cattiaux, J., R. Vautard, C. Cassou, P. Yiou, V. Masson-Delmotte, and F. Codron (2010), Winter 2010 in Europe: A cold extreme in a warming climate, *Geophys. Res. Lett.* **37**, L20704, DOI: 10.1029/2010GL044613.

Cony, M., E. Hernández, and T. Del Teso (2008), Influence of synoptic scale in the generation of extremely cold days in Europe, *Atmósfera* **21**, 4, 389-401.

Girguś, R., and W. Strupczewski (1965), A selection from historical sources of unusual hydrological and meteorological phenomena on Polish territories in the Xth to XVIth century, PIHM, Instrukcje i Podręczniki, No. 87, WKiŁ, Warszawa (in Polish).

Graham, R.J., C. Gordon, M.R. Huddleston, M. Davey, W. Norton , A. Colman, A.A. Scaife, A. Brookshaw, B. Ingleby, P. McLean, S. Cusack, E. McCallum, W. Elliot, K. Groves, D. Cotgrove, and D. Robinson (2006), The 2005/06 winter in Europe and the United Kingdom: Part 1 – How the MetOffice forecast was produced and communicated, *Weather* **61**, 12, 327-336, DOI: 10.1256/wea.181.06.

Gumiński, R. (1931), Winter 1928/29 in Poland, *Prz. Geogr.* **11**, 119-127.

Hansen, J., M. Sato, and R. Ruedy (2012), Perception of climate change, *Proc. Natl. Acad. Sci. U.S.A.* **109**, E2415-E2423, DOI: 10.1073/pnas.1205276109.

Hirschi, J.J.-M. (2008), Unusual North Atlantic temperature dipole during the winter of 2006/2007, *Weather* **63**, 1, 4-11, DOI: 10.1002/wea.120.

Hirschi, J.J.-M., and B. Sinha (2007), Negative NAO and cold Eurasian winters: how exceptional was the winter of 1962/1963? *Weather* **62**, 2, 43-48, DOI: 10.1002/wea.34.

Isayev, A.A., and B.G. Sherstyukov (2008), Mean and extreme characteristics of Moscow climate at the end of the 20th century, *Russ. Meteorol. Hydrol.* **33**, 3, 151-158, DOI: 10.3103/S1068373908030035.

Jaagus, J. (2006), Climatic changes in Estonia during the second half of the 20th century in relationship with changes in large-scale atmospheric circulation, *Theor. Appl. Climatol.* **83**, 1-4, 77-88, DOI: 10.1007/s00704-005-0161-0.

Klein Tank, A.M.G., J.B. Wijngaard, G.P. Können, R. Böhm, G. Demarée, A. Gocheva, M. Mileta, S. Pashiardis, L. Hejkrlik, C. Kern-Hansen, R. Heino, P. Bessemoulin, G. Müller-Westermeier, M. Tzanakou, S. Szalai, T. Pálsdóttir, D. Fitzgerald, S. Rubin, M. Capaldo, M. Maugeri, A. Leitass, A. Bukantis, R. Aberfeld, A.F.V. van Engelen, E. Forland, M. Mietus, F. Coelho, C. Mares, V. Razuvaev, E. Nieplova, T. Cegnar, J. Antonio López, B. Dahlström, A. Moberg, W. Kirchhofer, A. Ceylan, O. Pachaliuk, L.V. Alexander, and P. Petrovic (2002), Daily dataset of 20th-century surfaceair temperature and precipitation series for the European Climate Assessment, *Int. J. Climatol.* **22**, 12, 441-1453, DOI: 10.1002/joc.773.

Kossowska-Cezak, U. (1997), Monthly thermal and precipitation conditions and their dependence on atmospheric circulation, *Prace Stud. Geograf.* **20**, 125-144 (in Polish).

Kossowska-Cezak, U., S. Pełech, and R. Twardosz (2016), Exceptionally cold winter months in Europe (1951-2010), *Prz. Geofiz*. **61**, 1-2, 45-72 (in Polish).

Maignan, F., F.M. Bréon, E. Vermote, P. Ciais, and N. Viovy (2008), Mild winter and spring 2007 over western Europe led to a widespread early vegetation onset, *Geophys. Res. Lett.* **35**, 2, L02404, DOI: 10.1029/2007GL032472, 2008.

Moberg, A., P.D. Jones, D. Lister, A. Walther, M. Brunet, J. Jacobeit, L.V. Alexander, P.M. Della-Marta, J. Luterbacher, P. Yiou, D. Chen, A.M.G. Klein Tank, O. Saladié, J. Sigró, E. Aguilar, H. Alexandersson, C. Almarza, I. Auer, M. Barriendos, M. Begert, H. Bergström, R. Böhm, C. J. Butler, J. Caesar, A. Drebs, D. Founda, F.-W. Gerstengarbe, G. Micela, M. Maugeri, H. Österle, K. Pandzic, M. Petrakis, L. Srnec, R. Tolasz, H. Tuomenvirta, P.C. Werner, H. Linderholm, A. Philipp, H. Wanner, and E. Xoplaki (2006), Indices for daily temperature and precipitation extremes in Europe analyzed for the period 1901-2000, *J. Geophys. Res*. **111**, D22, D22106, DOI: 10.1029/2006JD007103.

Ouzeau, G., J. Cattiaux, H. Douville, A. Ribes, and D. Saint-Martin (2011), European cold winter 2009–2010: How unusual in the instrumental record and how reproducible in the ARPEGE-Climat model? *Geophys. Res. Lett.* **38**, 11, L11706, DOI: 10.1029/2011GL047667.

Sidorenkov, N.S., and I.A. Orlov (2008), Atmospheric circulation epochs and climate changes, *Russ. Meteorol Hydrol.* **33**, 9, 553-559, DOI: 10.3103/S1068373908090021.

Twardosz, R., and U. Kossowska-Cezak (2013a), Exceptionally hot summers in Central and Eastern Europe (1951-2010), *Theor. Appl. Climatol.* **112**, 3-4, 617-628, DOI: 10.1007/s00704-012-0757-0.

Twardosz, R., and U. Kossowska-Cezak (2013b), Exceptionally hot summers months in Central and Eastern Europe during the years 1951-2010. **In:** I. Dincer, C. Ozgur Colpan, and F. Kaglioglu (eds.), *Causes, Impacts and Solutions to Global Warming*, Springer, New York, 17-35, DOI: 10.1007/978-1-4614-7588-0_2.

Twardosz, R., and U. Kossowska-Cezak (2015a), Extremely cold summers months in Central and Eastern Europe, 1951-2010, *Nat. Haz.* **75**, 2, 2013-2026, DOI: 10.1007/s11069-014-1411-1.

Twardosz, R., and U. Kossowska-Cezak (2015b), Exceptionally hot and cold summers in Europe (1951-2010), *Acta Geophys.* **63**, 1, 275-300, DOI: 10.2478/s11600-014-0261-2.

Twardosz, R., and U. Kossowska-Cezak (2016), Exceptionally cold and mild winters in Europe (1951-2010), *Theor. Appl. Climatol*. **125**, 1-2, 399-411, DOI: 10.1007/s00704-015-1524-9.

Ugryumov, A.I., and N.V. Khar'kova (2008), Modern changes in St. Petersburg climate, *Russ Meteorol Hydrol.* **33**, 1, 15-19.

Van den Besselaar, E.J.M., A.M.G. Klein Tank, and G. van der Schrier (2010), Influence of circulation types on temperature extremes in Europe, *Theor. Appl. Climatol.* **99**, 3-4, 431-439, DOI: 10.1007/s00704-009-0153-6.

Wang, C., H. Liu, and S.K. Lee (2010), The record-breaking cold temperatures during the winter of 2009/2010 in the Northern Hemisphere, *Atmos. Sci. Lett.* **11**, 3, 161-168, DOI: AG-D-16-00077.

Wijngaard, J.B., A.M.G. Klein Tank, and G.P. Können (2003), Homogeneity of 20th century European daily temperature and precipitation series, *Int. J. Climatol.* **23**, 679-692, DOI: 10.1002/joc.906.

Received 17 March 2016
Received in revised form 3 August 2016
Accepted 23 August 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2630-2649
DOI: 10.1515/acgeo-2016-0096

First Measurements of the Earth's Electric Field at the Arctowski Antarctic Station, King George Island, by the New Polish Atmospheric Electricity Observation Network

Marek KUBICKI[1], Anna ODZIMEK[1], Mariusz NESKA[1], Jerzy BERLIŃSKI[2*], and Stanisław MICHNOWSKI[1*]

[1]Institute of Geophysics, Polish Academy of Sciences, Warszawa, Poland; e-mail: swider@igf.edu.pl (corresponding author)

[2]Faculty of Electronics and Information Technology, Warsaw University of Technology, Poland
*emeritus

Abstract

Atmospheric electricity measurements are performed all over the globe for getting a better understanding of the processes and phenomena operating in the Earth's electric atmosphere, ionosphere and magnetosphere. Over recent years, we have established coordinated observations of atmospheric electricity, mainly of the vertical component of the Earth's atmospheric electric field, from Polish observation stations: Stanisław Kalinowski Geophysical Observatory in Świder, Poland, Stanisław Siedlecki Polar Station in Hornsund, Svalbard, Norway, and, for the first time, the Henryk Arctowski Antarctic Station in King George Island. The organisation of this network is presented here as well as a preliminary summary of geophysical conditions at Arctowski, important from the point of view of atmospheric electricity observations. In particular, we refer to the geomagnetic observations made at Arctowski station in 1978-1995. We also present the average fair-weather diurnal variation

Ownership: Institute of Geophysics, Polish Academy of Sciences

of the atmospheric electric field based on observations made so far between 2013 and 2015.

Key words: atmospheric electricity, observation network, Arctowski Antarctic Station, geomagnetic activity, fair weather conditions.

1. INTRODUCTION

Polar regions, the Arctic and the Antarctic, have unique locations near the geographic and geomagnetic poles and due to the characteristics of their atmosphere and climate have become an important platform for atmospheric and solar-terrestrial studies (Lessard *et al.* 2014). International scientific events such as International Polar Years, especially the first in 1882/1883 and second in 1932/1933, and the International Geophysical Year (IGY) in 1957/58 had large influence on the development of research activities in these regions by many countries, explorers and scientists. The solar-terrestrial research activities in Arctowski Antarctic Station, which was open in 1977 (Rakusa-Suszczewski 1977, 1980) and named after Henryk Arctowski, the geoscientist in the Belgian Antarctic Expedition on "Belgica" ship in 1897-1899 (Machowski 1998a, b), started with geomagnetic field observations. The magnetic observation station was established at Arctowski station during the 2nd Antarctic Expedition of Polish Academy of Sciences in 1977-1978. The construction of two magnetic pavilions began in December 1977 and observations started in March 1978 (Szymański 1980). The magnetic observations have been stopped at Arctowski in 1995.

Atmospheric electricity (AE) observations in Antarctic started around the time of IGY and since then have always been part of the Antarctic scientific programme. Measurements have been made at Roi Baudoin Base, South Pole station (Cobb 1976, Bering *et al.* 1991, Reddell *et al.* 2004), Vostok (Park 1976a, Frank-Kamenetsky *et al.* 1999, 2001), Davis (Burns *et al.* 1995, Tinsley *et al.* 1998), and Maitri (*e.g.*, Deshpande and Kamra 2001).

2. ATMOSPHERIC ELECTRICITY OBSERVATIONS RUN BY IG PAS

The Institute of Geophysics, PAS (IG PAS) and the Stanisław Kalinowski Geophysical Observatory in Świder have long traditions in the observations of atmospheric electricity. Electric field measurements were made at the Magnetic Observatory in Świder, Poland (21.25°E, 52.12°N, geomagnetic latitude ~50.5°N) as early as the 1920s. As none of the pre-World War II data survived, the existing series of electric field measurements at Świder starts in 1958 (Dziembowska 2009). Nowadays, the observatory uses a wide spectrum of experimental techniques and equipment for comprehensive meteorological, aerosol and atmospheric electricity observations (*e.g.*, Kubicki *et al.* 2003, 2007, 2016). The atmospheric electricity measurements include:

the vertical component of the ground-level DC electric field using a radioactive collector and field mill method, vertical component of the density of the DC electric current (the air-Earth current) by a Wilson antenna and positive and negative air conductivity in the range of mobility of small ions, measured by a Gerdien condenser. At the Stanisław Siedlecki Polish Polar Station in Hornsund, Spitsbergen, Norway (15.50°E, 77.00°N, geomagnetic latitude ~73.9°N) atmospheric electricity observations started in 1989. They include measurements of the vertical component of the DC electric field and, sporadically, the DC electric current density. The electric field is measured by a rotating dipole field-mill (Berliński *et al.* 2007) and, independently, using a radioactive collector method. Alongside the atmospheric electricity measurements, meteorological observations are also performed automatically or by traditional methods both at Świder and Hornsund.

Land observation stations, such as Świder, especially in highly-populated areas, often suffer from high and variable pollution (aerosol), the effects of which shift the local electric field from its global diurnal pattern. The air turbulence which drives atmospheric convection current also affects the electric field (the most at local mid-day) and makes the global atmospheric electricity monitoring more difficult (Kubicki *et al.* 2014). On the other hand, in the polar regions, especially in the broad vicinity of geomagnetic poles, the ground-level atmospheric electric field is affected by the ionospheric convection which results from the interaction of solar wind and the magnetosphere (Dungey 1961, Park 1976b). Atmospheric observations at Hornsund, owing to its location, have been primarily set to investigate such effects, in addition to the geomagnetic observations performed there continuously since the 1980s (Michnowski 1998, Michnowski *et al.* 1991, 1996, 2003, 2007; Kleimenova *et al.* 2010, 2011, 2012, 2013; Kozyreva *et al.* 2007; Odzimek *et al.* 2011; Frank-Kamenetskii *et al.* 2012).

Henryk Arctowski Polish Antarctic station, King George Island, South Shetland Islands (58.47°W, 62.16°S, geomagnetic latitude ~51°S), where atmospheric electricity measurements have not been made so far, offered an interesting possibility of investigating the DC electric field in the polar location which is not affected by anthropogenic aerosol and strong effects of atmospheric turbulence and which is sufficiently distant from the effects of ionospheric convection in the Southern Hemisphere – a valuable location for the monitoring the main (tropospheric) generator of atmospheric electricity. The advantage of this site, apart from its location in the Antarctic, was that meteorological observations have been performed there, and, besides, the level of magnetospheric disturbances could be estimated using the geomagnetic observations made at the station in the 1970-1990s.

In December 2011 we started a two-year research project set to measure and investigate diurnal variations of the atmospheric electric field and aerosol at ground-level in the spatial scale extending from middle latitudes in central Europe to the Arctic and Antarctic (Świder, Hornsund, and Arctowski stations). Besides the scientific goals, our motivation for the undertaking of this project was to improve the observational capacity of the Polish observatories and to systematise and standardise the atmospheric electricity observations, starting with strengthening and extending the existing observations and transforming it into a professional scientific observational network. The created network is described in the following section.

3. POLISH ATMOSPHERIC ELECTRICITY OBSERVATION NETWORK

A three-site, internet-linked network of atmospheric electricity observatories has been set up, which includes the Polish Polar Station in Hornsund (HRN) in the Arctic, Arctowski Station in the Antarctic (ARC), and the Świder Geophysical Observatory (SWI) in central Poland – Fig. 1a.

The instruments at each station include:

– A system for measurements of the electric field strength near ground level. In addition to the older radioactive collector type sensors at Świder and Hornsund, new sensors have been installed. They are rotating-dipole type sensors of frequency bandwidth from 0 to 10 Hz, and sensitivity of 2 V/m, amplitude range of ±1500, ±6000 V/m and output voltage ±3.5 V. The system requires 12 V buffer power supply (Berliński *et al.* 2007). Sensors are calibrated in a laboratory using a high-voltage power supply, and once a month it is short-circuit at the measurement site is in order to control the reference level.

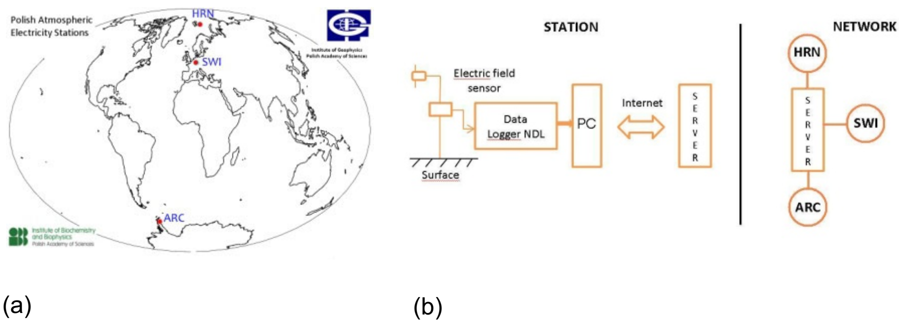


Fig. 1: (a) World map with the locations of observation stations of the Polish atmospheric electricity network: HRN − Hornsund, SWI − Świder, ARC – Arctowski; (b) Schematic diagram of the organisation of the atmospheric electricity observation network and observational data flow. Left side: organization of a single station. Right side: three-station network (HRN, SWI, ARC).

– A data logger (NDL) with a GPS clock for the precise timing and synchronisation of the measurements in the three observation sites. The sampling frequency can be changed within the range from 0.1 to 10 s; at present it is at 1 s. Data are recorded on a 32 GB flash memory card. The logger has a unique IP address, allowing communication and control via internet and data download as well as current data visualisation. The data logger and buffer power supply are installed in a portable container (0.5 m × 1.0 m × 0.6 m dimensions) which makes the installation and field measurements easier (Figs. 1b and 3).

– In addition, for the purpose of the project, portable sensors for the measurements of aerosol concentration for aerosol size from 10 nm to 1 μm have been used at Hornsund and Arctowski where the aerosol concentration measurements have been made for the first time. For the aerosol concentration measurements at Arctowski, the TSI 8525 model aerosol counter has been used, which measures particles of the size range from 20 nm to 1 μm (Kubicki *et al.* 2016).

A dedicated software package has been designed and created for the data upload from the three observation sites to a database on a server at IG PAS. The software enables visualisation of atmospheric electricity measurements taken at the three sites. The design of the database allows incorporating data from new observation sites. The schematic diagram of the network is shown in Fig. 1b.

4. ATMOSPHERIC AND GEOPHYSICAL CONDITIONS AT ARCTOWSKI STATION

4.1 Fair-weather conditions

Weather is an important factor that affects atmospheric electricity and has to be taken into account when analysing electric field data. The fair weather conditions used in our analysis mean there is no rain, drizzle, snow, hail, fog, no local or distant thunderstorms, the lower cloudiness is less than 4/8 (8/8 is overcast), and the wind velocity is less than 6 m/s. Below we summarise the main characteristics of weather conditions at Arctowski during the atmospheric electricity observation period and we present the analysis of the number of days with fair-weather conditions at the station.

The number of fair-weather days at Arctowski is small, on average 3.4 days per month. In 2013 the number of fair-weather days was 32, and 60 fair-weather days occurred in 2015. Over 2014 only 12 fair-weather days are registered but standard meteorological measurements have been made only during half of the year. Most fair-weather days occur in the local (Southern Hemisphere) winter season. This coincides with the time of the maximum thunderstorm activity in the Northern Hemisphere, and may be important for the analysis of the global (universal) signal in the electric field, since the station is more distant from the activity than the Świder and Hornsund stations are. A small number of days with little cloud cover was discussed in articles by Cygan (1981): in the period 1979/1980 (December 1979 – 17 March 1980) there was no cloudless days (cloudiness <1/8). There was only one day with cloudiness of 2/8 (2 in okta scale 0-8). The cloudiness on most days (*i.e.*, 60 days) was 6/8-8/8. According to Angiel *et al.* (2010), in 2006 (January-December) there was no clear sky days, and 11 days with cloudiness less than 2/8. Low clouds are the dominant cloud types, mostly with stratocumulus and stratus. Measurements during the cloudy days can be used for the studies of the cloud generator in the global circuit (Odzimek *et al.* 2014).

4.2 Aerosol conditions

An analysis of aerosol concentrations at Arctowski station is described in detail in Kubicki *et al.* (2016). The main and important conclusion concerning Arctowski is that natural aerosol concentrations (on most days below 1000 cm^{-3}, average fair-weather median value 535 cm^{-3}) do not have effects on the ground-level electric field, the same as in Hornsund, and contrary to Świder (where the effect of anthropogenic aerosol on the electric field takes place – see Kubicki *et al.* (2016).

4.3 Geomagnetic activity

In this section we present the results of comparison of the level of magnetic activity at Hornsund, Arctowski, and Belsk – another observatory of IG PAS. We assume that the activity at Świder is close to that at Belsk (BEL – 20.80° E, 51.83° N, geomagnetic latitude ~50.4°N) since the stations are at the distance of 50 km. The magnetic latitude of Arctowski is close to the magnetic latitude of Belsk and Świder, where the influence of magnetospheric generator in the ground-based electric field is usually absent. Some effects can be present during geomagnetic storms (Nikiforova *et al.* 2005, Kleimenova *et al.* 2008, 2009, 2013; Michnowski *et al.* 2014) and therefore we estimate the possible level of this disturbance at ARC by comparing geomagnetic activity levels at the three sites.

The magnetic activity at Arctowski was measured over the period of April 1978 to December 1995. Szymański *et al.* (1980) describe the setup of the magnetic observations at the station. We use available geomagnetic data from the three stations over that period, published in Publications of the Institute of Geophysics Polish Academy of Sciences, Series C – "Geomagnetism" in the tables "Three-hour-range indices K and magnetic character figures C". The indices have been calculated as described in

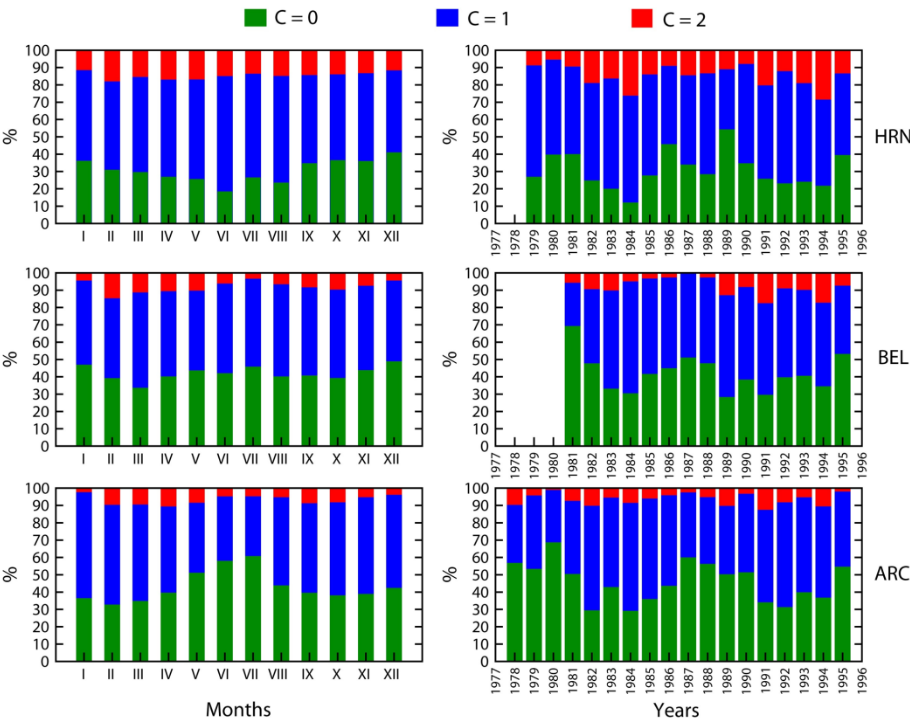


Fig. 2. Comparison of the geomagnetic activity at HRN, BEL, and ARC stations, based on the magnetic character figure C = 0 (weak), 1 (medium), 2 (strong). Left panels: Average variation of percentage of days characterized by C = 0, 1, 2 during a year. Right panels: percentage of days characterized by C = 0, 1, 2 over the period of magnetic observations at ARC (1979-1995). Month averages are based on 1981-1995 data.

Nowożynski *et al.* (1991). We perform simple comparison of the geomagnetic activity in terms of the magnetic character figure C (Fig. 2). We look at the average geomagnetic activity during a year (in each month), and the magnetic activity in each year of the considered period. A seasonal variation in the magnetic activity is always expected but the average levels are different.

In Arctowski the magnetic activity is overall comparable to Belsk station, although the Belsk station suffers from local artificial disturbance caused by railway which affects the number of magnetically quiet days (Neska *et al.* 2013). There is a small difference in the percentage of magnetically quiet days (C = 0), *i.e.*, 45% at ARC *versus* 42% at BEL, and the percentage of strongly disturbed days is slightly greater, 8% at ARC *versus* 7% at BEL). On the other hand, Hornsund (magnetic latitude ~73.8 S) is most magnetically disturbed, with only 30% of magnetically quiet days and 15% of highly disturbed days. In historical reports we have not found evidence of aurora sightings directly from Arctowski; however, such events cannot be excluded. By private communications (Julita Biszczuk-Jakubowska and the station crew taking care of AE observations) we have established that, at least by eye, the aurora has not been observed in 2006 and between 2011 and 2015, but it is only from 2013 that the crew has been alerted to notify about such events if they were able to see the aurora. We conclude that the magnetic activity at Arctowski station is comparable to that of Świder, but Arctowski is more suitable for monitoring the global atmospheric electric circuit − GEC (Rycroft *et al.* 2000) than Świder, because, as opposed to Świder, the effects of aerosol and atmospheric turbulence at ARC can be neglected, as demonstrated by Kubicki *et al.* (2016). On the other hand, the weather conditions are not very favourable and the atmospheric electricity fair-weather criteria are rarely met. Nevertheless, Arctowski station can provide valuable recordings of fair-weather electric field for comparison with Świder observations, as presented in Section 5.

4.4 Radioactivity

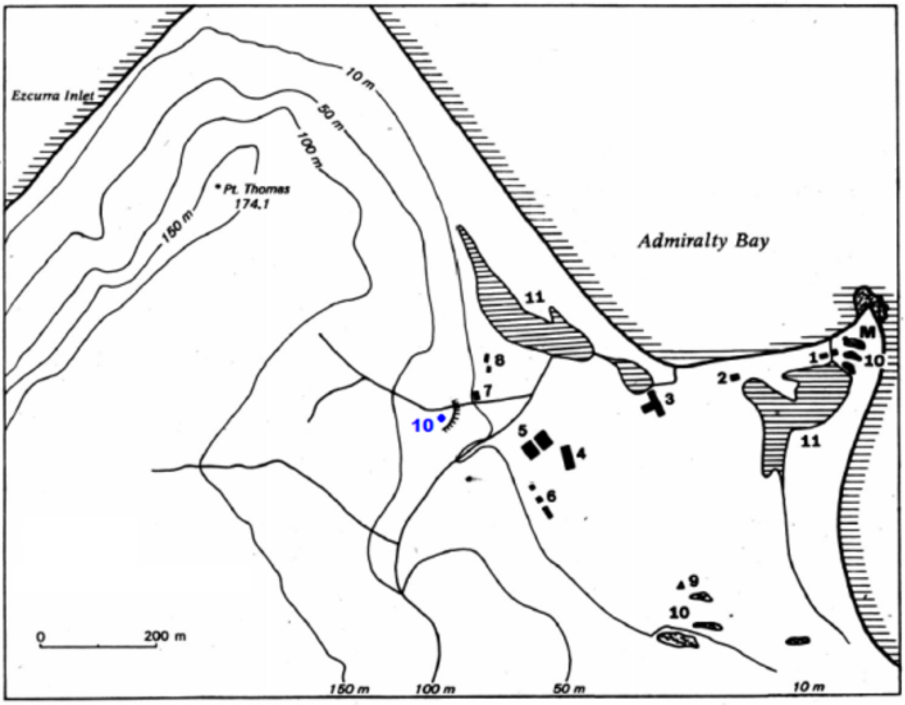
According to geological research (Birkenmajer 1980) the South Shetland islands are of volcanic origin and therefore may contain radioactive elements that cause increased ground-level air ionisation. This may affect the measured electric field due to the formation of the so-called electrode layer. The present electric field measurements show that there is no effect of the electrode layer on the electric field; however, professional measurements of the radioactivity of the soil are planned in future.

5. ATMOSPHERIC ELECTRIC FIELD AT ARCTOWSKI STATION

Atmospheric electricity observations, *i.e.*, of the atmospheric electric field have been performed at Arctowski Station for the first time. Figure 3 shows the electric field sensor and the container containing the data logger and power supply installed there in December 2012. Figure 4 shows the sensor's location, added to the schematic map of the station infrastructure from the paper on magnetic observations by Szymański (1980).

The electric field at the ground level is the superposition of electrical effects of global (main generators of the GEC) and local origin (meteorological phenomena, natural and anthropogenic aerosols) (Bennett and Harrison 2007). In Fig. 5 we present examples of electric field variations at Arctowski station in different weather conditions. In this paper we use the convention that the downward directed is positive – this is the direction the atmospheric electric field in fair-weather (atmospheric electricity convention) –

Fig. 3. Photographs of the electric field sensor and data acquisition system at Arctowski station, installed there in December 2012. Photography by Sylwia Łukawska.



Situation of geophysical laboratory and magnetic pavilions in Polish Antarctic Station Arctowski
1 — Biological laboratory, 2 — Meteorology — laboratory, 3 — Living house, 4 — Power station, 5 — Magazines, 6 — Summer laboratories, 7 — Geophysical laboratory, 8 — Magnetic pavillions, 9 — Geodethic pillar, M — Azimuth mark
10 - Electric field sensor

Fig. 4. Schematic plan of the Arctowski station with added location of the electric field sensor (10 – blue point on the map). Adapted from Szymański (1980, Fig. 1).

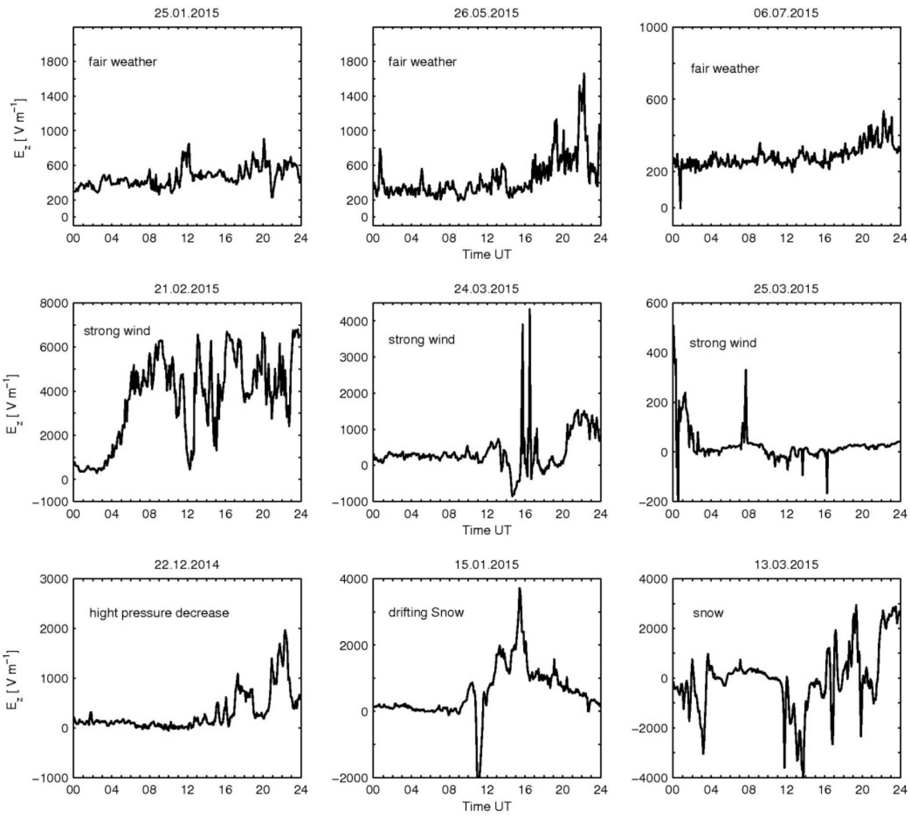


Fig. 5. Examples of diurnal variations of the ground-level electric field *Ez* measured at Arctowski station in 2013-2015. Values shown are not reduced to their ground-level values.

physically the values we show are equivalent to the values of the vertical atmospheric electric potential gradient (physics sign convention). The criterion of fair weather allows to select data with minimal local effects (Fig. 5, upper row). Strong winds can cause transportation of space charge accumulated in the surface layer associated with the effect of electrode layer and subsequently a change in the electric field (Fig. 5, middle row). During rain or snow, the wind can cause the turbulence mixing of precipitation particles on which electric charge can be stored and this may result in positive or negative spikes in the measured field. The effect of low clouds, precipitation or fog may appear in the variation of the vertical electric field as mild or sudden change of the sign or amplitude or a pulse in the atmospheric electric field (Fig. 5, bottom row).

5.1 Diurnal fair-weather electric field variations

In atmospheric electricity of great interest are diurnal variations of the electric field or electric current density in the conditions of fair weather. They are usually compared to the Carnegie curve which is the historical, average diurnal variation of the fair-weather atmospheric electric field derived from ocean measurements (Parkinson and Torreson 1931), and in the view of the global atmospheric circuit hypothesis (Wilson 1920) reflects the diurnal variation of electricity (from electrified clouds) in the atmosphere. It is believed that the average of the ocean measurements is free from factors which usually affect the field, depending on where it is measured and at what time, such as atmospheric convection, effect of aerosols, additional regional and local current generators. From this point of view the local conditions at Arctowski are moderate − the site is far from world thunderstorm activity centers and far from the normal range of the magnetospheric generator (as discussed in Section 4), atmospheric convection can sometimes occur there in the summer, for example on days with high difference of temperature between the ground and air – the temperatures at Arctowski station have been investigated by *e.g.* Kejna (1999), Kejna and Láska (1999), and more recently by Araźny *et al.* (2013).

Below we present an updated diurnal variation of the fair-weather electric field at Arctowski station. The new curve has been calculated on the basis of whole day diurnal variation of the electric field (00-24 UT) from 38 selected days of fair weather, from the three years of observations between 2013 and 2015*, i.e.*, from the beginning of the electric field measurements (Table 1). We show the diurnal variation of hourly average mean (Fig. 6a) and medians (Fig. 6b) and the Carnegie curve (Fig. 6c) for comparison (values taken from Isräel 1970). The errors displayed in Fig. 6a are the standard deviations, in Fig. 6b – the interquartile range determined by the first quartile from the bottom and the third quartile from the top. The difference between the third quartile and the median is usually larger than the difference between the median and first quartile which means that there are more values higher than the median value. This is not true particularly for 10 UT when lower values of the electric field occur more often at this hour – this is reflected in the mean average curve. Both curves have maximum at 19-20 UT which usually corresponds with the maximum activity of the American thunderstorm centre. The correlation coefficient of the mean curve with the Carnegie curve is 0.75, and of the median curve it is 0.61. Compared with the Carnegie curve the deviation of the curve's maximum from the average value of the curve (300 V/m – this value is not reduced to the value at ground level) is 14.2%, while for the Carnegie curve it is 40%. The average mean curve has a minimum between 10 and 13 UT, the median curve has a

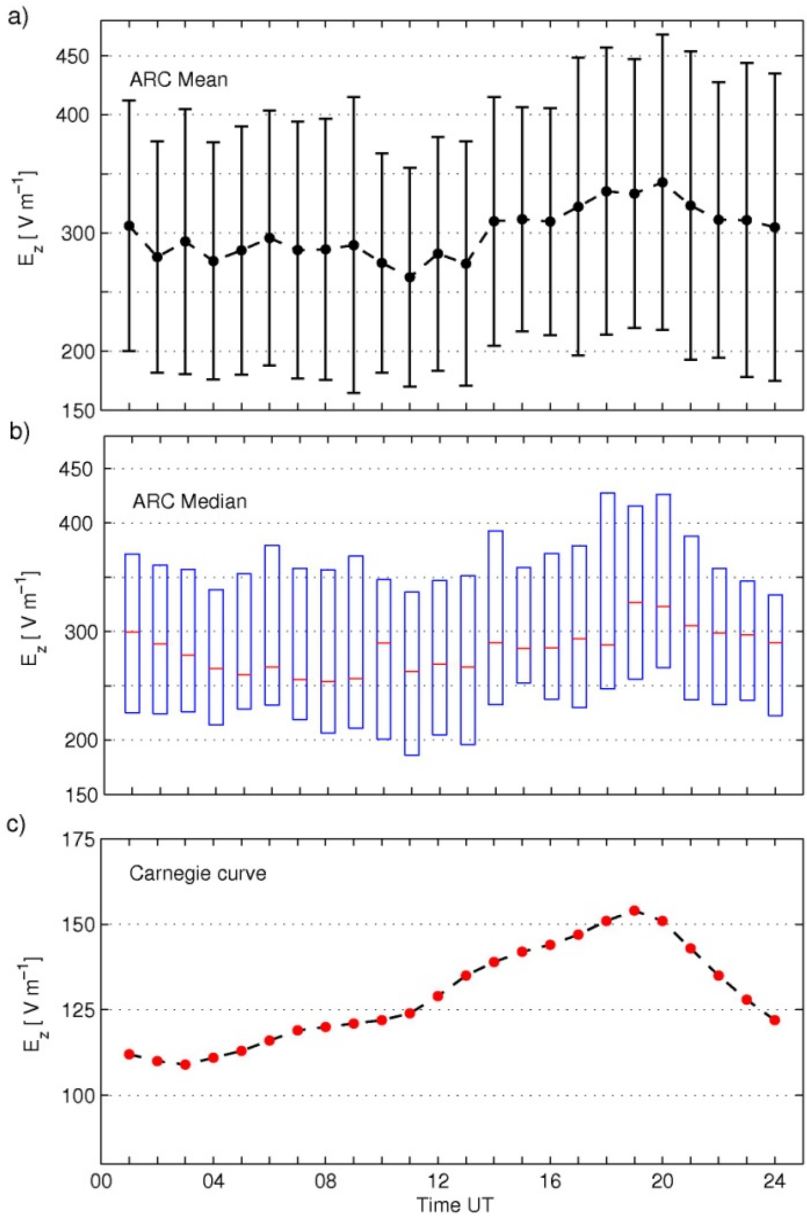


Fig. 6. Diurnal variation of the fair-weather atmospheric electric field E_z at Arctowski station based on measurements from 2013-2015. Values shown are not reduced to their ground-level values. The field is positive in agreement with the atmospheric electricity sign convention. (a) Average mean curve with error bars as standard deviation, (b) Median curve and the quartile bar errors, (c) The Carnegie curve (Isräel 1970).

Table 1

Analysis of atmospheric electricity fair-weather conditions at Arctowski station over the observation period from December 2013 to December 2015. The number of days fulfilling the criteria of fair weather (for at least half of the day) is given in the first column, days with high cloudiness (6-8), snow and strong wind − in the second, third, and fourth column, respectively.

Month	Fair weather	Cloudiness 6-8/8	Snow, drifting snow	Strong wind diurnal mean >10 m/s
Jan 2013	2	27	2	2
Feb 2013	3	24	-	3
Mar 2013	4	20	-	4
Apr 2013	2	25	-	8
May 2013	2	28	-	2
Jun 2013	4	20	-	5
Jul 2013	3	25	3	3
Aug 2013	4	24	1	5
Sep 2013	3	26	5	2
Oct 2013	2	27	-	2
Nov 2013	2	28	-	No data
Dec 2013	1	25	-	1
2013 Total	**32**			
Jan 2014	1	28	-	1
Feb 2014	0	26	-	1
Mar 2014	1	30	-	No data
Apr 2014	4	24	-	5
May 2014	2	13	-	3
Jun 2014	No data	No data	No data	3
Jul 2014	No data	No data	No data	7
Aug 2014	No data	No data	No data	3
Sep 2014	No data	No data	No data	2
Oct 2014	No data	No data	No data	No data
Nov 2014	2	13	-	2
Dec 2014	2	18	1	No data
2014 Total	**12**			
Jan 2015	2	20	3	2
Feb 2015	3	23	4	4
Mar 2015	5	15	7	6

to be continued

Table 1 (continuation)

Month	Fair weather	Cloudiness 6-8/8	Snow, drifting snow	Strong wind diurnal mean >10 m/s
Apr 2015	3	21	2	1
May 2015	4	26	4	-
Jun 2015	5	22	2	-
Jul 2015	12	19	2	-
Aug 2015	9	18	-	-
Sep 2015	3	23	-	-
Oct 2015	4	20	-	-
Nov 2015	6	No data	No data	No data
Dec 2015	4	13	No data	No data
2015 Total	**60**			
May 2013	2	28	-	2

minimum between 07 and 09 UT while the minimum of the Carnegie is at 03 UT. When compared with the diurnal variation from Świder (Kubicki *et al.* 2007) the Arctowski electric field curve has a minimum later than that of Świder (03-04 UT), and also different from the minimum of the Carnegie curve. At present it is difficult to judge whether this is because of local effects or global effects, which manifested themselves in this particular way.

6. SUMMARY AND CONCLUSION

Polish network of atmospheric electricity stations in three scientific observatories at PAS stations: Stanislaw Kalinowski Geophysical Observatory in Świder, Stanisław Siedlecki Polar Station in Hornsund and Henryk Arctowski Antarctic Station has been set up. Data from the network observation sites are organised in a database. The observation equipment is portable and allows easy installation of atmospheric electricity observations at other locations in future, especially at other sites in polar regions.

The atmospheric electricity observation programme is important for atmospheric electricity research and atmospheric physicists, particularly concerned with the Earth's global atmospheric electric circuit as it provides tools for the global circuit monitoring (Williams 2009). To considerable extent it can also be relevant in the studies of near-Earth environment and in meteorological studies, as the GEC is coupled to extra-terrestrial influences and has direct connections with the Earth's weather. Polish Antarctic stations, the Arctowski, and the A.B. Dobrowolski station located closer to the South Geomagnetic Pole, 100.75°E, 66.27°S (Gregorczuk 1980) and closed at present, are situated in locations which provide interesting opportunities for the development of atmospheric and solar-terrestrial research. The variations of the atmospheric electric field at Arctowski during fair-weather is a valuable addition to the monitoring of diurnal variations in the atmospheric electric circuit and monitoring of its generators. Data from the network can be used for global atmospheric electric circuit (GEC) studies and its diurnal, seasonal and long-term variations, validation of GEC models such as EGATEC (Odzimek *et al.* 2010), investigations of the electricity of clouds (Odzimek *et al.* 2014) and disturbances in fair-weather atmospheric electric field during events such as geomagnetic storms, Forbush decreases or solar proton events on atmospheric electricity at ground level (*e.g.*, Kleimenova *et al.* 2009).

A c k n o w l e d g m e n t s. This work has been supported by Polish National Science Centre grant number NCN-2011/01/B/ST10/07118, awarded to the Institute of Geophysics of the Polish Academy of Sciences. Atmospheric electricity observations at Świder and Hornsund have been supported within the statutory activities of Institute of Geophysics PAS, grant No. 3841/E-41/S/2013 and No. 3841/E-41/S/2015 of the Ministry of Science and Higher Education of Poland. Credits are also to the Institute of Biochemistry and Biophysics of the Polish Academy of Sciences for providing access to the observation site and operating the measurements of the atmospheric ground-level electric field at the Henryk Arctowski Polish Antarctic Station, and for providing meteorological and aerosol concentration data. The authors would like to thank Maciej Benedyk, Mirosław Szumny, Piotr Łepkowski who performed the measurements at the Polish Polar Station in Hornsund, and Sylwia Łukawska, Emil Kasprzyk and Dawid Gajownik who performed the measurements at the Polish Arctowski Antarctic Station.

For the analysis of geomagnetic activity we used data from the following volumes of *Publications of the Institute of Geophysics, Polish Academy of Sciences*: Arctowski: volumes C-21 (181), C-22 (182), C-28 (202), C-32 (212), C-35 (225), C-44 (244), C-54 (276), C-60 (292), C-63 (300); Belsk C-8 (133), C-9 (139), C-10 (144), C-13 (159), C-17 (166), C-20 (180), C-23 (187), C-26 (196), C-29 (205), C-34(208), C-37 (227), C-38 (228), C-40 (240), C-45 (250), C-49 (259), C-51 (267), C-55 (298), C-58 (287); Hornsund: C-14 (163), C-27 (196), C-31 (210), C-43 (243), C-47 (254), C-48 (256), C-53 (273), C-57 (286), C-64 (301). We thank Danuta Jasinkiewicz from the Geophysical Observatory in Świder for her help in digitalisation of the magnetic data from paper publications.

References

Angiel, P.J., M. Potocki, and J. Biszczuk-Jakubowska (2010), Weather condition characteristics at the H. Arctowski Station (South Shetland Islands, Antarctica) for 2006, in comparison with multi-year research results, *Misc. Geogr.* **14**, 79-89.

Araźny, A., M. Kejna, and I. Sobota (2013), Ground temperature at the Henryk Arctowski Station (King George Island, Antarctic) – Case study from the period January 2012 to February 2013, *Bull. Geogr., Phys. Geograph.* **6**, 1, 59-80, DOI: 10.2478/bgeo-2013-0004.

Bennett, A.J., and R.G. Harrison (2007), Atmospheric electricity in different weather conditions, *Weather* **62**, 10, 277-283, DOI: 10.1002/wea.97.

Bering III, E.A., J.R. Benbrook, G.J. Byrne, and A.A. Few (1991), Measurements of vertical atmospheric electric current at a network of sites in Antarctica including manned stations and automatic geophysical observations. **In:** Proc. the International Workshop on Global Atmospheric Electricity Measurements, S. Michnowski and L.H. Ruhnke (eds.), *Publs. Inst. Geophys. Pol. Acad. Sci.* **D-35**, 238, 119-135.

Berliński, J., G. Pankanin, G. and M. Kubicki (2007), Large scale monitoring of troposphere electric field. **In:** *Proc. 13th Int. Conf. Atmospheric Electricity, 13-18 August 2007, Beijing, China*, 124-126.

Birkenmajer, K. (1980), Report on geological investigations of King George Island, South Shetland Islands (West Antarctica) in 1978/79, *Stud. Geol. Pol.* **64**, 89-105 (in Polish).

Burns, G.B., M.H. Hessea, S.K. Parcella, S. Malachowski, and K.D. Coleb (1995), The geoelectric field at Davis station, Antarctica, *J. Atmos. Sol. Terr. Phys.* **57**, 14, 1783-1797, DOI: 10.1016/0021-9169(95)00098-M.

Cobb, W.E. (1976), Atmospheric Electric Measurements at the South Pole. **In:** H. Dolezalek, R. Reiter, and H. Landsberg (eds.), *Proc. 5th Int. Conf. on Atmospheric Electricity, Electrical Processes in Atmospheres, 2-7 September 1974, Garmisch-Partenkirchen, Germany*, 161-167, DOI: 10.1007/978-3-642-85294-7_23.

Cygan, B. (1981), Characteristics of meteorological conditions at the Arctowski Station during the summer season of 1979-1980, *Pol. Polar Res.* **2**, 3-4, 35-46.

Deshpande, C.G., and A.K. Kamra (2001), Diurnal variations of the atmospheric electric field and conductivity at Maitri, Antarctica, *J. Geophys. Res.* **106**, D13, 14207-14218, DOI: 10.1029/2000JD900675.

Dziembowska, A. (2009), Eighty years of fair-weather atmospheric electricity monitoring in Poland. **In:** P. Baranski and M. Kubicki (eds.), Recent developments in Atmospheric Electricity, *Publs. Inst. Geophys. Pol. Acad. Sci.* **D-73**, 412, 9-14.

Dungey, J.W. (1961), Interplanetary magnetic field and the auroral zones, *Phys. Rev. Lett.* **6**, 2, 47-48, DOI: 10.1103/PhysRevLett.6.47.

Frank-Kamenetsky, A.V., G.B. Burns, O.A. Troshichev, V.O. Papitashvili, E.A. Bering, and W.J.R. French (1999), The geoelectric field at Vostok, Antarctica: its relation to the interplanetary magnetic field and the cross polar cap potential difference, *J. Atmos. Sol. Terr. Phys.* **61**, 18, 13471-1356, DOI: 10.1016/S1364-6826(99)00089-9.

Frank-Kamenetsky, A.V., O.A. Troshichev, G.B. Burns, and V.O. Papitashvili (2001), Variations of the atmospheric electric field in the near-pole region related to interplanetary magnetic field, *J. Geophys. Res.* **106**, A1, 179-190, DOI: 10.1029/2000JA900058.

Frank-Kamenetskii, A.V., A.L. Kotikov, A.A. Kruglov, G.B. Burns, N.G. Kleimenova, O.V. Kozyreva, M. Kubitski, and A. Odzimek (2012), Variations in the near-surface atmospheric electric field at high latitudes and ionospheric potential during geomagnetic perturbations, *Geomagn. Aeron.* **5**, 52, 629-638, DOI: 10.1134/S0016793212050064.

Gregorczuk, M. (1980), Climate of Bunger Oasis (region of A.B. Dobrowolski Station, Antarctic), *Pol. Polar Res*, **1**, 4, 205-230.

Isräel, H. (1970), *Atmospheric Electricity*, Vol. 1, Fundamentals, Conductivity, Ions, Israel Program for Scientific Translations, Jerusalem, 317 pp.

Kejna, M. (1999), Air temperature on King George Island, South Shetland Islands, Antarctica, *Pol. Polar Res.* **20**, 3, 183-201.

Kejna, M., and K. Láska (1999), Spatial differentiation of ground temperature in the region of Arctowski Station, King George Island, Antarctica, *Pol. Polar Res.* **20**, 3, 221-241.

Kleimenova, N.G., O.V. Kozyreva, S. Michnowski, and M. Kubicki (2008), Effect of magnetic storms in variations in the atmospheric electric field at midlatitudes, *Geomagn. Aeron.* **48**, 5, 622-630, DOI: 10.1134/S0016793208050 071.

Kleimenova, N., O. Kozyreva, M. Kubicki, and S. Michnowski (2009), Variations of the mid-latitude atmospheric electric field (Ez) associated with geomagnetic disturbances and Forbush decreases of cosmic rays. **In:** P. Baranski and M. Kubicki (eds.), Recent developments in Atmospheric Electricity, *Publs. Inst. Geophys. Pol. Acad. Sci.* **D-73**, 412, 55-64.

Kleimenova, N.G., O.V. Kozyreva, M. Kubicki, and S. Michnowski (2010), Morning polar substorms and variations in the atmospheric electric field, *Geomagn. Aeron.* **50**, 1, 48-57, DOI: 10.1134/S0016793210010068.

Kleimenova, N.G., O.V. Kozyreva, M. Kubicki, and S. Michnowski (2011), Variations in the near-ground electric field at high latitudes and the potential drop across the polar cap during morning polar substorms, *Geomagn. Aeron.* **51**, 3, 394-401, DOI: 10.1134/S0016793211030091.

Kleimenova, N.G., O.V. Kozyreva, M. Kubicki, A. Odzimek, and L.M. Malysheva (2012), Effect of substorms in the Earth's nightside sector on variations in the surface atmospheric electric field at polar and equatorial latitudes, *Geomagn. Aeron.* **52**, 4, 467-473, DOI: 10.1134/S001679321204007X.

Kleimenova, N.G., O.V. Kozyreva, S. Michnowski, and M. Kubicki (2013), Influence of geomagnetic disturbances on atmospheric electric field (Ez) variations at high and middle latitudes, *J. Atmos. Sol.-Terr. Phys.* **99**, 117-122, DOI: 10.1134/S0016793212050064.

Kozyreva, O.V., N.N. Nikiforova, N.G. Kleimenova, S. Michnowski, and M. Kubicki (2007), Electric air-Earth vertical current pulsations at Hornsund during polar substorms. Case studies, **In:** *Proc. 13th Int. Conf. on Atmospheric Electricity, 13-18 August 2007, Beijing, China*, Vol. 1, 29-33.

Kubicki, M., S. Michnowski, B. Myslek-Laurikainen, and S. Warzecha (2003) Long term variations of some atmospheric electricity, aerosol, and extraterrestrial parameters at Świder Observatory. **In:** *Proc. 12th Int. Conf. on Atmospheric Electricity, 9-13 June 2003, Versailles, France*, Vol. 1, 291-294.

Kubicki, M., S. Michnowski, and B. Myslek-Laurikainen (2007), Seasonal and daily variations of atmospheric electricity parameters registered at the Geophysical Observatory at Świder (Poland) during 1965-2000. **In:** *Proc. 13th Int. Conf. on Atmospheric Electricity, 13-18 August 2007, Beijing, China*, 50-53.

Kubicki, M., A. Odzimek, N.G. Kleimenova, O.V. Kozyreva, and M. Neska (2014), Synchronisation of main Global Electric Circuit generators from ground-level electric field Ez at three distant locations on the globe at middle and high latitudes. **In:** *Proc. 15th Int. Conf. on Atmospheric Electricity, 15-20 June 2014, Norman, USA*, 9 pp.

Kubicki, M., A. Odzimek, and M. Neska (2016), Relationship of ground-level aerosol concentration and atmospheric electric field at three observation sites in the Arctic, Antarctic and Europe, *Atmos. Res.* **178-179**, 329-346, DOI: 10.1016/j.atmosres.2016.03.029.

Lessard, M.R., A.J. Gerrard, and A.T. Weatherwax (eds.) (2014), Solar-terrestrial research in polar regions: past, present, and future. **In:** *Proc. Polar Research Meeting, 12-13 November 2012, University of New Hampshire, USA*, 64 pp.

Machowski, J. (1998a), Henryk Arctowski (1871-1958), *Pol. Polar Res.* **19**, 1-2, 7-10.

Machowski, J. (1998b), Contribution of H. Arctowski and A. B. Dobrowolski to the Antarctic Expedition of Belgica (1897-1899), *Pol. Polar Res.* **19**, 1-2, 15-30.

Michnowski, S., A. Szymanski, and N.N. Nikiforova (1991), On simultaneous observations of geomagnetic and atmospheric electric field changes in Arctic station Hornsund, Spitsbergen. **In:** S. Michnowski and L. H. Ruhnke (eds.), Proc. Int. Workshop on Global Atmospheric Electricity Measurements, 10-16 September 1989, Mądralin, Poland, *Publ. Inst. Geophys. Pol. Acad. Sci.* **D-35**, 238, 83-96.

Michnowski, S., N.N. Nikiforova, and N.G. Kleimenova (1996), The response of the ground-level electric field at Hornsund to magnetospheric-ionospheric events. **In:** *Proc. 10th Int. Conf. on Atmospheric Electricity, 19-24 June 1996, Osaka, Japan*, 520-523.

Michnowski, S. (1998), Solar wind influences on atmospheric electricity variables in polar regions, *J. Geophys. Res.* **103**, D12, 13939-13948, DOI: 10.1029/98JD01312.

Michnowski, S., M. Kubicki, J. Drzewiecki, S. Israelsson, N. Kleimenova, N. Nikiforova, and O. Kozyreva (2003), Variations of atmospheric electricity elements in polar regions related to the solar wind changes, **In:** *Proc. 12th Int. Conf. in Atmospheric Electricity, 9-13 June 2003, Versailles, France*, Vol. 1, 295-296.

Michnowski, S., M. Kubicki, N.G. Kleimenova, N.N. Nikiforova, O.V. Kozyreva, and S. Israelsson (2007), The polar ground-level electric field and current variations in relation to solar wind changes. **In:** *Proc. 13th Int. Conf. on Atmospheric Electricity, 13-18 August 2007, Beijing, China*, Vol. 1, 9-12.

Michnowski, S., A. Odzimek, N.G. Kleimenova, O.V. Kozyreva, M. Kubicki, and N.N. Nikiforova (2014), Review of examples of solar wind lower atmosphere coupling observed in the electric field (Ez) variations at the Earth's surface during magnetic storms (2014). **In:** *Proc. 15th Int. Conf. on Atmospheric Electricity, 15-20 June 2014, Norman, USA*, 8 pp.

Neska, A., J. Reda, M. Neska, and Y. Sumaruk (2013), On the influence of DC railway noise on variation data from Belsk and Lviv Geomagnetic Observatories, *Acta Geophys.* **61**, 2, 385-403, DOI: 10.2478/s11600-012-0058-0.

Nikiforova, N.N., N.G. Kleimenova, O.V. Kozyreva, M. Kubicki, and S. Michnowski (2005), Unusual variations in the atmospheric electric field during the main phase of the strong magnetic storm of October 30, 2003, at Świder Polish midlatitude Observatory, *Geomagn. Aeron.* **45**, 1, 140-144.

Nowożynski, K., T. Ernst, and J. Jankowski (1991), Adaptive smoothing method for computer derivation of K-indices, *Geophys. J. Int.* **104**, 1, 85-93, DOI: 10.1111/j.1365-246X.1991.tb02495.x.

Odzimek, A., and M. Lester (2009), Modelling the Earth's global atmospheric electric circuit – development, challenges and directions. **In:** P. Baranski and M. Kubicki (eds.), Recent developments in Atmospheric Electricity, *Publs. Inst. Geophys. Pol. Acad. Sc.* **D-73**, 214, 37-54.

Odzimek, A., M. Lester, and M. Kubicki (2010), EGATEC: A new high-resolution engineering model of the global atmospheric electric circuit. 1. Currents in the lower atmosphere, *J. Geophys. Res.* **115**, D18207, DOI: 10.1029/2009JD013341.

Odzimek, A., M. Kubicki, M. Lester, and A. Grocott (2011), Relation between the SuperDARN ionospheric potential and ground electric field at polar station Hornsund. **In:** *Proc. 14th Int. Conf. on Atmospheric Electricity, 7-12 August 2011, Rio de Janeiro, Brazil*, 4 pp.

Odzimek, A., M. Kubicki, and P. Baranski (2014), Ground-level atmospheric electricity under low-level stratiform clouds. **In:** *Proc. 15th Int. Conf. on Atmospheric Electricity, 15-20 June 2014, Norman, USA*, 7 pp.

Park, C.G. (1976a), Solar magnetic sector effects on the vertical atmospheric electric field at Vostok, Antarctica, *J. Geophys. Res.* **3**, 8, 475-478, DOI: 10.1029/GL003i008p00475.

Park, C.G. (1976b), Downward mapping of high-latitude ionospheric electric fields to the ground, *J. Geophys. Res.* **81**, 1, 168-174, DOI: 10.1029/JA081i001p00168.

Parkinson, W.C., and O. Torreson (1931), The diurnal variation of the electric potential of the atmosphere over the ocean, *Int. Union Terrest. Magn. Electr. Bull.* **8**, 340-345.

Rakusa-Suszczewski, S. (1977), Henryk Arctowski Antarctic Station of Polish Academy of Sciences, *Nauka Polska* **11-12**, 103-114 (in Polish).

Rakusa-Suszczewski, S. (1980), III Antarctic Expedition to Arctowski Station (November 1978 – May 1979), *Pol. Polar Res.* **1**, 1, 127-146 (in Polish).

Reddell, B.D., J.R. Benbrook, E.A. Bering, E.N. Cleary, and A.A. Few (2004), Seasonal variations of atmospheric electricity measured at Amundsen-Scott South Pole Station, *J. Geophys. Res.* **109**, A09308, DOI: 10.1029/2004JA010536.

Rycroft, M.J., S. Israelsson, and C. Price (2000), The global atmospheric electric circuit, solar activity and climate change, *J. Atmos. Sol.-Terr. Phys.* **62**, 17-18, 1563-1576, DOI: 10.1016/S1364-6826(00)00112-7.

Szymański, A. (1980), The construction and work of magnetic observatory "Arctowski" (King George Island, South Shetland Islands), *Pol. Polar Res.* **1**, 2-3, 105-115.

Tinsley, B.A., W. Liu, R.P. Rohrbaugh, and M.W. Kirkland (1998), South Pole electric field responses to overhead ionospheric convection, *J. Geophys. Res.* **103**, D20, 26137-26146, DOI: 10.1029/98JD02646.

Wilson, C.T.R. (1920), Investigation on lightning discharges and on the electric field of thunderstorms, *Phil. Trans. A* **21**, 73-115.

Williams, E.R. (2009), The global electrical circuit: a review, *Atmos. Res.* **91**, 2-4, 140-152, DOI: 10.1016/j.atmosres.200805.018.

Received 24 February 2016
Received in revised form 1 July 2016
Accepted. 23 August 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2650-2676
DOI: 10.1515/acgeo-2016-00112

Classification of Aerosol over Central Europe by Cluster Analysis of Aerosol Columnar Optical Properties and Backward Trajectory Statistics

Artur SZKOP, Aleksander PIETRUCZUK, and Michał POSYNIAK

Institute of Geophysics, Polish Academy of Sciences, Warsaw, Poland;
e-mails: aszkop@igf.edu.pl, alek@igf.edu.pl (corresponding author),
mpos@igf.edu.pl

Abstract

A cluster analysis is applied to the Aerosol Robotic Network (AERONET) data obtained at Belsk, Poland, as well as three nearby Central European stations (Leipzig, Minsk and Moldova) for estimation of atmospheric aerosol types. Absorption Ångstrom exponent (AAE), aerosol optical thickness (AOT) and extinction Ångstrom exponent (EAE) parameters are used. Clustering in both 2D (AOT, EAE) and 3D (AOT, EAE, AAE) is investigated. A method of air mass backward trajectory analysis is then proposed, with the receptor site at Belsk, to determine possible source regions for each cluster. Four dominant aerosol source regions are identified. The biomass burning aerosol source is localized in the vicinity of Belarusian-Ukrainian border. Slovakia and northern Hungary are found to be the source of urban/industrial pollutants. Western Poland and eastern Germany are the main sources of polluted continental aerosols. The most differentiated source region of Scandinavia, Baltic Sea and Northern Atlantic, associated with lowest values of AOT, corresponds to clean continental and possibly maritime type aerosols.

Key words: aerosol optical properties, aerosol classification, cluster analysis, backward air mass trajectory, aerosol's origin determination.

---

1. INTRODUCTION

Aerosols have been observed in the global scale for more than a decade by both ground-based (Holben *et al.* 1999, 2001) and satellite (King *et al.* 1999, Kaufman *et al.* 2002, Mishchenko *et al.* 2007) remote sensing instruments. In the recent years, several classification schemes have been introduced to determine aerosol's type (Omar *et al.* 2009, Pappalardo *et al.* 2013). These new classification methods allow for better understanding of aerosols' origin and their source regions as well as aerosol modification during long range atmospheric transport. Moreover, knowledge of the aerosol type allows for a validation and a parametrization of aerosol models and satellite data (*e.g.*, Ginoux *et al.* 2001, Levy *et al.* 2007, Remer *et al.* 2005).

Aerosols are classified using satellite data by backtracking plumes to their probable source regions and analyzing satellite data products like fire locations or variation of optical properties (Russell *et al.* 2014). Another approach utilizes classification of aerosols based on their optical properties measured by both satellite and/or ground based measurements (Torres *et al.* 2002, Kaufman *et al.* 2002, Chin *et al.* 2002). One of the most commonly used databases containing aerosol optical properties is the Aerosol RObotic NETwork (AERONET) (Holben *et al.* 2001). This database, containing long series of optical properties for worldwide locations, allows to calculate radiative properties of various aerosols based on the location of the measurements site (*e.g.*, Dubovik *et al.* 2002, Cattrall *et al.* 2005). Relations between measured optical properties are also used to classify aerosols. The most commonly used method of finding aerosol type is the analysis of aerosol optical thickness (AOT) and its spectral dependence described by extinction Ångstrom exponent (EAE) (*e.g.*, Kalapureddy *et al.* 2009, Boselli *et al.* 2012). Analysis of extinction Ångstrom exponent and its derivative with respect to the wavelength is also used to classify aerosols (Gobbi *et al.* 2007) as well as combination of this method with LIDAR retrievals (Perrone *et al.* 2014). Moreover, LIDAR based parameters like LIDAR ratio, color ratio or depolarization are factors applied for the aerosol classification (Burton *et al.* 2012, Groß *et al.* 2013, 2015). In recent years, aerosol absorption parameters also have become one of the most promising factors for the aerosol classification. Costabile *et al.* (2013) used aerosol size distribution and spectral dependence of aerosol scattering and absorption to identify aerosol types. Giles *et al.* (2012) have shown that spectral dependence of aerosol absorption and extinction is useful for distinguishing between urban/industrial, biomass burning and mineral dust aerosols. The aforementioned method was used by Valenzuela *et al.* (2015) to identify dust over Granada. Clustering of aerosol properties combined with air mass' backward trajectory analysis was also used to classify aerosols, *e.g.*, Kikas *et al.* (2008).

In this paper, statistical analysis of AERONET retrievals for the period of 2001-2012, taken at four Central European stations (Belsk, Leipzig, Minsk, and Moldova), is used to determine the dominant aerosol types in the region. The utilized optical parameters include Aerosol Optical Thickness (AOT), Extinction Ångstrom Exponent (EAE) and Absorption Ångstrom Exponent (AAE). Moreover, a statistical analysis of air mass' backward trajectories is utilized for the Belsk station to determine probable source regions of the retrieved aerosol types. Source regions obtained by 3D (AOT *versus* EAE *versus* AAE) and 2D (EAE *versus* AOT) clustering are compared to these obtained from the fixed thresholds in the domain of EAE *versus* AOT. This allows for a direct comparison of all the classification methods as well as a preliminary identification of the best one for identification of aerosols originating from identified source regions.

2. INSTRUMENTATION AND METHODS

CIMEL Sun and Sky scanning photometers are used worldwide in determining aerosol optical properties by the Aerosol Robotic Network. The instruments provide aerosol optical thickness at several wavelengths derived from the direct Sun measurements as well as other optical parameters, *e.g.*, size distribution, refractive index and single scattering albedo (SSA), derived from sky radiance measurements. Detailed description of the instrument, data products and data processing are given by Holben *et al.* (1999, 2006), Eck *et al.* (1999), and Smirnov *et al.* (2000). In this work daily averaged, quality assured AERONET Level 2.0 aerosol optical thicknesses and single scattering albedo as well as parameters describing their spectral dependence are used. Uncertainties of AOT are estimated to vary from ±0.01 in the visible and near infrared to ±0.02 in the UV range (Holben *et al.* 1999, Eck *et al.* 1999). Uncertainties of single scattering albedo were estimated to be ±0.03 for AOT greater than 0.4 at 440 nm (Dubovik *et al.* 2000). However, in the case of data level 2.0 this uncertainty could in fact be smaller, see discussion given by Giles *et al.* (2012).

Wavelength dependence of AOT is described by the extinction Ångstrom exponent (EAE) parameter:

$$\mathrm{EAE} = -\frac{d\ln\left(\mathrm{AOT}(\lambda)\right)}{d\ln(\lambda)} \tag{1}$$

which is constant in the visible range. Small values of EAE typically indicate coarse aerosols, while large values indicate the fine ones. Columnar aerosol absorption properties are described by the absorption aerosol optical thickness (AAOT) which could be calculated form AERONET products by:

$$\mathrm{AAOT}(\lambda) = \left(1 - \mathrm{SSA}(\lambda)\right)\mathrm{AOT}(\lambda) \tag{2}$$

Spectral dependence of AAOT is described by the absorption Ångstrom exponent parameter defined, similarly to the extinction one, as:

$$\mathrm{AAE} = -\frac{d\ln\left(\mathrm{AAOT}(\lambda)\right)}{d\ln(\lambda)} \tag{3}$$

In this work values of EAE and AAE are computed by linear regression of logarithm of proper optical thickness *versus* logarithm of wavelength λ. Spectral dependence of AAOT for fine homogenous black carbon (BC) spheroids (soot) is proportional to $1/\lambda$ which indicates AAE close to one for BC particles (Bergstrom *et al.* 2002) whiles for brown carbon (BrC) values of AAE could be much larger (Moosmüller *et al.* 2009). In the case of non-homogenous particles of varying sizes the AAOT spectral dependence is not significantly influenced by the size of the particles (Berry and Percival 1986) but it can be a function of the particles' state and physical properties (Fuller *et al.* 1999). For instance, for a BC core shelled by other material, values of AAE could be smaller than one according to the calculations performed by Lack and Cappa (2010). Values of AAE as low as ~0.3 were found by Russell *et al.* (2010). However, AERONET derived AAE values significantly below one could also be caused by uncertainties in estimation of SSA, as it was discussed by Giles *et al.* (2012). Authors of the aforementioned paper found, based on AERONET data, that AAE varies from 1.5 to 2.3 for dust aerosol and from 1.1 to 1.8 for both urban/industrial and biomass burning aerosols. In our study a threshold is set at AAE value of 0.75, below which the cases are considered to be caused mainly by errors in the estimation of SSA.

In this work a clustering method is applied to the daily averaged AERONET data level 2.0 for Belsk (51°49′N, 20°,48′E), Leipzig (51°N, 12°E), Minsk (53°N, 27°E), and Moldova (47°N, 28°E) stations situated in the Central Europe to distinguish between different types of aerosols. A typically used approach is a clustering of AOT and EAE (Boselli *et al.* 2012) or AAE and EAE (Giles *et al.* 2012). We propose to perform a 3D clustering of AOT (at 440 nm), EAE (440-870 nm), and AAE (440-870 nm) data set and compare this analysis to 2D clustering of AOT (at 440 nm) and EAE (440-870 nm). According to the t-test the mean values of AOT measured at Leipzig, Minsk and Moldova stations do not differ significantly from the Belsk mean. All that stations represent Central Europe and data collected by them are considered as one block. The total number of measurements in the dataset is 3954 and this dataset is used in 2D clustering. The SSA retrievals, required for AAE calculation in 3D clustering, are not always available in the level 2.0 data, for which the uncertainties must remain below a superimposed threshold (Holben *et al.* 2006). The uncertainties of SSA retrievals increase for lower AOT values. In the dataset studied here SSA retrievals were found only for 681 cases. Moreover, our calculations of AAE yielded values smaller than one in a number of cases. As indicated in the previous paragraph values slightly below unity may occur for composite aerosol with a BC core whilst values much lower than one are caused mainly by errors in estimation of SSA. A threshold for AAE at 0.75 was chosen with values below it deemed unphysical. In the end the total number of the remaining data records used in the 3D analysis is 613. It is worth noting that no SSA retrievals were found in the studied level 2.0 AERONET data in the cases of AOT below 0.24. On the other hand SSA values were available for all measurements with AOT above 0.4. In an intermediate region between the aforementioned AOT values, the SSA values are only available for a part of the dataset.

The clustering is performed based on the Lloyd's algorithm, also known as the k-means method (Lloyd 1982). A prechosen number of k clusters, defined by the randomly seeded centroids, are populated by n observation points. Each n-th observation point is assigned to exactly one cluster based on its lowest point to centroid distance. The iterative algorithm then repositions the centroids to the average values of the observation points in a given cluster until the centroids remain stationary or a limit of the iterations is reached. Procedure is performed ten times for randomly distributed seeds and a case with best separation of clusters is used in subsequent analysis. The number of clusters is determined by means of silhouette analysis which allows to check data separation for each number of clusters and choose the number for best data separation between clusters (Rousseeuw 1987, Kaufman and Rousseeuw 1990). The chosen clustering algorithm assumes the distribution of the data points to be close to normal. Since the AOT values are typically closer to a log-normal distribution the logarithm of the AOT values is used for the clustering. Separation of the aerosol spectra to coarse and fine modes is expected to be extracted from AOT *versus* EAE domain. Aerosols containing carbon and mineral dust are expected to form a cluster in the domain of AAE *versus* EAE. Clusters characterized by different values of AOT, if obtained, could indicate aerosols of different concentrations which may be caused by different origins and sources. In our previous work large AOT values were typically concurrent with an advection of air mass from the east and the south direction (Jarosławski and Pietruczuk 2010, Pietruczuk and Chaikovsky 2012). Pietruczuk (2013) also observed large increase in the AOT during advection episodes from the south direction. Both of these findings indicate the existence of possible aerosol source regions in the aforementioned directions.

A statistical analysis of air mass trajectories calculated with HYSPLIT model (Stein *et al.* 2015) is performed in an attempt to identify possible source regions of the aerosols assigned to each individual cluster of aerosol optical properties. The Belsk station is chosen as a receptor site and only the corresponding AERONET measurements are chosen from the clustered dataset for the further analysis. Trajectories are computed, for the location of Belsk only, with the use of HYSPLIT model supplied by GDAS meteorological archives. In the case of gaps in the GDAS archives the FNL data are used. We used meteorological ensemble trajectories to reduce uncertainties. Five-day trajectories, with endpoints separated by 6-hour long intervals (6:00, 12:00, and 18:00 UTC) and ending at 500 m above ground level over Belsk, are used similarly to our previous works (Pietruczuk and Jarosławski 2013, Pietruczuk 2013). This altitude is representative for boundary layer where most of aerosols are transported. A trajectory is chosen when its time of arrival is within a day of a clustered AERONET measurement. The trajectories are then cast onto a two-dimensional 0.5°×0.5° grid positioned over Europe. The value of each individual grid point becomes equal to the total number of hours spent over it by the trajectories belonging to the given cluster population. The length of each individual trajectory (measured along the trajectory) is used for the normalization to compensate for the decrease of trajectory density with the growing distance from the receptor. Such a decrease, in the case of a sheaf of straight lines on a two-dimensional surface, is inversely proportional to range.

This is an unorthodox approach and it should be emphasized that it presents a significant deviation from the usual methods for trajectory analysis, that may be found in the literature (*e.g.*, Seibert *et al.* 1994, Stohl *et al.* 1995, Robinson *et al.* 2011, Dvorská *et al.* 2009). In the typically used methods a function of the parameter values measured at a receptor site is added to the grid points laying on the receptor site back-trajectory. The sums are then normalized by the total number of trajectories passing over, thus creating the mean value for the grid point. In this way a spatial distribution of the aerosol source for a receptor site is revealed. Such approach applied to AOTs favors areas associated with the largest AOTs registered at receptor site, Belarus and Ukraine in case of Belsk (Kabashnikov *et al.* 2014) whilst sources of other types of aerosols are in fact not seen in that analysis. The method proposed in this work is aimed at recognizing both sources and characteristics of aerosols identified over the receptor site.

Satellite data are also used in discussion of possible sources of aerosols. We used AOTs measured by Moderate Resolution Imaging Spectroradiometer (MODIS) instrument, as well as MODVOLC thermal anomalies. Thermal anomalies are high-temperature MODIS pixels which were used primarily to monitor volcanic activity (Wright *et al.* 2004). However, it is a good tool for open fires monitoring. Such fires may have a natural or anthropogenic origin like an agricultural activity or exploitation of oil fields (fire torches burning excess natural gas). The latter also explains the presence of fire pixels over the North Sea as a high number of oil rigs are operating in this area.

3. RESULTS OF AERONET DATA ANALYSIS

3.1 3D clustering

Cluster analysis of AERONET level 2.0 data for four central European stations containing SSA indicated the existence of three distinguishable clusters. Statistical parameters of aerosol optical properties for each cluster are listed in Table 1. In addition, parameters for the Belsk station (the receptor site in the trajectory analysis) are calculated separately. The differences in the values of the means calculated for Belsk as well as the remaining stations do not exceed 5% and are within 1 standard deviation in all the cases. This indicates that Belsk is a typical station for European Plain and thus is affected by similar types of aerosols as the other above-mentioned sites. It allows us to use the clusters determined for all the stations (depicted in Fig. 1) for the analysis limited only to Belsk station. An example of such an analysis is presented in the next step.

Clusters 3 and 1, containing 385 and 159 cases, respectively, are characterized by similar mean values of both Ångstrom exponents and different magnitudes of AOT. Mean values of EAE around 1.60-1.70 are indicative of fine particles and AAE around 1.13-1.18 suggests that they contain carbon.

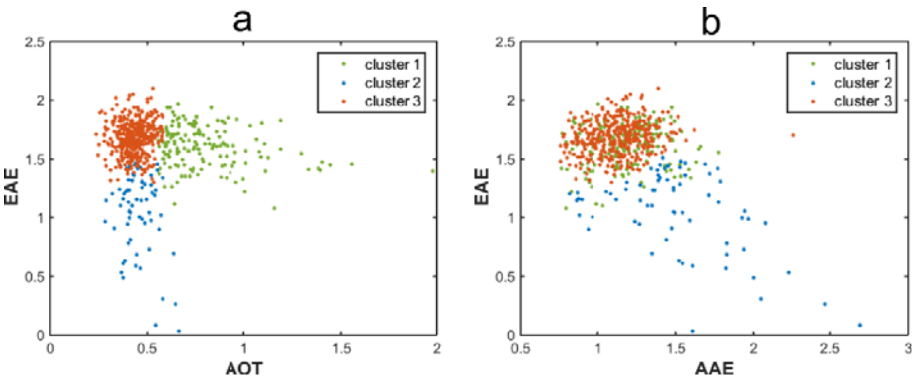


Fig. 1. The results obtained from the 3D clustering of AERONET data measured in the Central Europe (Belsk, Minsk, Moldova, and Leipzig). The results of the clustering, performed over 3-dimensional parameter space of AAE, EAE, and AOT, is projected on EAE *versus* AOT domain (a) and on EAE *versus* AAE domain (b).

Table 1

The results obtained from the 3D clustering. N is the cluster population size while Δ and δ are means and standard deviations, respectively

Dataset	Cluster	N	AOT at 440 nm		EAE		AAE	
			Δ	δ	Δ	δ	Δ	δ
All sites	1	159	0.80	0.26	1.60	0.18	1.18	0.21
	2	69	0.45	0.08	1.07	0.34	1.50	0.37
	3	385	0.43	0.07	1.69	0.14	1.13	0.18
Belsk	1	31	0.81	0.20	1.60	0.22	1.19	0.22
	2	9	0.46	0.09	1.23	0.19	1.49	0.30
	3	78	0.42	0.08	1.69	0.14	1.10	0.17

A little smaller AAE in cluster 3 indicates that the aerosol related to this cluster contains a larger carbon fraction. Simple classification scheme based on combination of AOT and EAE (*e.g.*, Barnaba and Gobbi 2004) suggests the presence of continental aerosols in both clusters characterized by different magnitudes of AOT which may indicate different types of continental aerosols. For example, Boselli *et al.* (2012), by performing 2D clustering, obtained a cluster of larger AOTs, around 0.21, and EAE around 1.61 which they identify as containing contaminated continental aerosol, *e.g.*, biomass burning or industrial. However, in our case, because of method's limitations, AOTs in clusters 3 and 1 are much higher (mean values of 0.43 and 0.8, respectively) than obtained by these authors, and significantly smaller than reported for biomass burning aerosols, where values larger than 1 are often seen (*e.g.*, Balis *et al.* 2003, Eck *et al.* 2003, Toledano *et al.* 2007, Hsu *et al.* 1999). In our case, cluster 1 could be related to biomass burning aerosol, while cluster 3 comprises continental aerosol of other types but also one containing carbon, like urban or industrial one. Cluster 2 of 69 cases is characterized by relatively small values of extinction Ångstrom exponent (mean EAE around 1), mean AOT around 0.45 and mean values of AAE around 1.5. This combination suggests the presence of mineral dust. The obtained vales of AOTs and EAE are typical for Saharan dust events observed in the Mediterranean (Kaskaoutis *et al.* 2007a, Kosmopoulos *et al.* 2008, Kaskaoutis *et al.* 2007b) and AAE for mineral dust (Giles *et al.* 2012).

In order to find the possible aerosol source regions and identify aerosol types, a statistical analysis of backward trajectories was performed. The results of this analysis are shown in Fig. 2. In this work, trajectories ending over Belsk are used. Please note that an analysis was not performed for

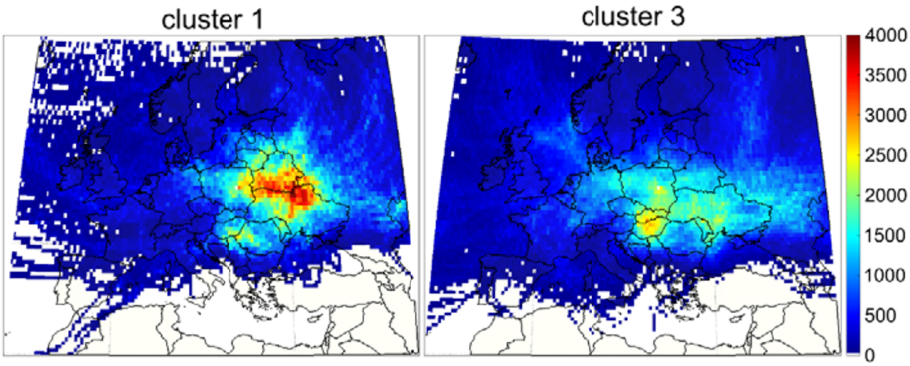


Fig. 2. The distribution of trajectories with endpoints (located at 500 m above the receptor) concurrent with the 3D clustered measurements for Belsk station. The values represent the number of hours spent by all the trajectories over a grid point, normalized by the lengths of individual trajectories (see Section 2). Clusters 1 and 3 are depicted. The population of cluster 2 is too small for this method of analysis, hence it is omitted here.

cluster 2, related to the mineral dust, because of an insufficient number of trajectories related to it (only 9 cases for the Belsk station dataset). Moreover, the chosen altitude of 500 m above the ground levels could be too small to investigate mineral dust over Belsk. For example, Saharan dust in the case of this station is rare and is registered typically in the free troposphere (Pietruczuk and Chaikovsky 2012). However, the number of records in clusters 1 and 3 is sufficient for a statistical analysis. Both of these clusters are related to fine, absorbing aerosols and the only difference is in the magnitude of AOT. Statistical analysis of the trajectories indicated different possible source regions of the aerosol, and thus probably different types of aerosols. In the case of cluster 1 with the largest AOT, typical for biomass burning events, the largest density of the trajectories corresponds to regions of seasonal biomass burning (Barnaba *et al.* 2011). Thus, cluster 1 is related to biomass burning aerosol. In the case of cluster 3, the largest density of trajectories is located over Slovakia and Hungary and also the existence of hot spots is indicated over Ukraine and in direct vicinity of Belsk. According to this analysis and our previous findings (Pietruczuk 2013), cluster 3 should be related to urban/industrial aerosol. The following results indicate to an underrepresentation of northern and western origin in the trajectory population. This stands in conflict with wind statistics for the Central Europe as the western winds are the most common for this region. In fact, more than 50% of backward trajectories is related to western direction. As the performed 3D clustering method excludes the cases with lower AOT values (where no SSA value is available) these cases should include predominantly the underrepresented trajectory directions. Moreover, the limit of AOT renders the types of aerosols characterized by small values of AOT invisible in 3D clustering, *i.e.*, maritime, continental aerosol for clear conditions and even contaminated continental (Boselli *et al.* 2012). Such aerosols may still be analyzed through other methods, namely the 2D clustering and the fixed threshold approaches elaborated in the following paragraphs.

3.2 2D clustering

A 2D clustering is performed on all available daily averaged AERONET Level 2.0 data containing AOT and EAE measured at four stations used in the previous section. This analysis revealed 5 clusters; three of them correspond to that obtained by 3D clustering and two are related to smaller AOTs which are unavailable for 3D clustering. It is worth to note that the data used for 3D clustering are available only for AOT (at 440 nm) greater than 0.24. Moreover, not all of the data with AOTs between 0.24 and 0.4 fulfil conditions required for calculation of SSA Level 2.0 and because of that are not available for 3D clustering. The obtained clusters are depicted in Fig. 3 and the statistical parameters of aerosol optical properties related to the obtained clusters are listed in Table 2. The main differences, when comparing to the 3D clustering, are very sharp borders between the clusters which are well defined and may suggest good separation of the data. However, according to the performed silhouette analysis (Rousseeuw 1987, Kaufman and Rousseeuw 1990) separation of the clusters in this case is rather poor. Obtained borders seem to be artificial, which indicates a weak separation.

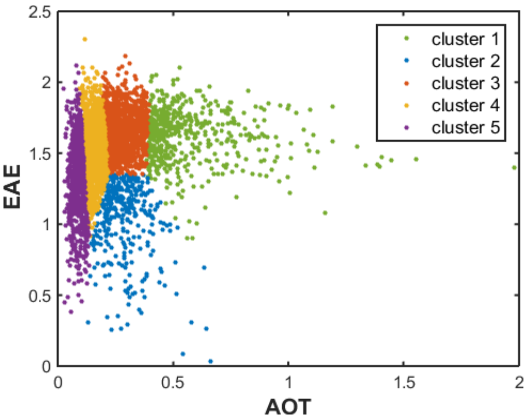


Fig. 3. The results obtained from the 2D clustering of AERONET data obtained at four locations in the Central Europe (Belsk, Minsk, Moldova, and Leipzig). The clustering was performed over a 2-dimensional parameter space of EAE and AOT.

Table 2

The results obtained from the 2D clustering. N is the cluster population size while Δ and δ are means and standard deviations, respectively

Dataset	Cluster	N	AOT		EAE	
			Δ	δ	Δ	δ
All sites	1	520	0.58	0.23	1.62	0.18
	2	336	0.29	0.08	1.04	0.26
	3	1100	0.28	0.05	1.66	0.15
	4	1248	0.16	0.03	1.53	0.20
	5	750	0.08	0.02	1.32	0.28
Belsk	1	100	0.58	0.19	1.61	0.20
	2	52	0.27	0.06	1.12	0.88
	3	246	0.28	0.05	1.66	0.15
	4	279	0.16	0.03	1.52	0.19
	5	146	0.09	0.02	1.29	0.26

Therefore, some points associated with a certain cluster may in fact belong to another one. This is probably caused by the use of only two parameters for clustering, which may not be enough to distinguish between types of aerosol characterized by similar optical properties. Another difference is a shift of cluster 1 into smaller AOTs direction and the presence of smaller AOTs in cluster 2.

The clusters related to smaller AOTs are numbered 4 and 5. AOTs related to cluster 5 are the smallest, with a mean value of 0.08±0.02, whilst EAE extends to the almost entire range with a mean value of 1.32±0.28, where assigned uncertainties are standard deviations for cluster. This cluster may contain two types of aerosol, depending on measured EAE. Aerosols characterized by small AOTs and large values of EAE are typically described as continental ones, whilst those characterized by small values of EAE are described by maritime ones. Mean AOT related to cluster 4 is equal or a little larger than 0.16±0.03, while the mean value of EAE is 1.53±0.20. Such a kind of aerosol may be described as polluted continental (Boselli *et al.* 2012) or mixed one (Russell *et al.* 2010). Cluster 3 is related to fine aerosol with comparatively larger AOTs of about 0.28±0.05 which is much smaller than the one found in the 3D clustering. This, with the mean EAE values in the vicinity of 1.66±0.15, indicates that this cluster contains mainly cases of contaminated continental aerosol (Boselli *et al.* 2012) and industrial aerosol. This classification is also supported by the findings of Dubovik *et al.* (2002) where similar values (AOT of 0.26 and EAE between 1.2 and 2.3) were found for a station located in Paris and classified as "urban-industrial and mixed" aerosol. According to other studies, *e.g*., of Kalapureddy *et al.* (2009) or Pawar *et al.* (2015), this aerosol should be classified as urban/industrial. cluster 1, because of a possible misclassification of the data close to a sharp border, may contain also urban/industrial aerosol. However, as discussed in the previous section, larger values of AOT in this cluster should be related to biomass-burning aerosol. According to an analysis performed in the case of the 3D clustering, aerosol optical parameters related to cluster 2 indicate predominance of desert dust in this cluster.

As in the case of the 3D clustering, statistical analysis of backward trajectories was performed in order to find possible aerosol source regions and support identification of aerosol types. Trajectory densities related to each cluster are depicted in Fig. 4. In this case, many more trajectories related to cluster 2 were obtained. It allowed to perform an analysis for this cluster. The largest densities, obtained for the air mass coming from the southern and western directions, should not be equated to dust source. A transport of aerosol from this direction is rather accompanied by a transport of desert dust from northern Africa. Moreover, endpoints of trajectories at 500 m above ground level do not reflect typical altitude for dust observed over Belsk. In the case of cluster 1, the area with the largest trajectory density is similar to that obtained in case of the 3D clustering for both biomass-burning and ur - ban/industrial aerosol. It should be expected because cluster 1 in the case of the 2D clustering is shifted towards smaller AOTs and contains some points that were associated with 3D cluster 3 described as continental aerosol containing carbon, industrial one. 2D cluster 3 is related to an advection of aerosol from the western direction rather than from the southern one. It is caused by shifting of this cluster to smaller AOTs as well as by presence of a large number of cases not present in the 3D clustering for which no SSA was available in the AERONET Level 2.0 data. Aerosols with smaller AOTs, related to this cluster and advection from the western direction (northern Germany and Benelux), should be classified mainly as mixed or polluted continental, while those with larger AOTs and advected from the southern direction should be classified as urban/industrial. High density of trajectories in the vicinity of the receptor site suggests that it is locally contaminated. However, in this case, the concept of "locality" may be extended onto western Poland and eastern Germany. In the case of cluster 4 no distinct source region is indicated. Latvia, Lithuania, and Estonia as well as advection from Atlantic are only slightly marked. According to analysis of aerosol optical properties, this cluster should be described rather as mixed one than polluted continental. Analysis of the trajectories related to cluster 5 indicates the northern part of the Baltic Sea and the North Sea as the most probable sources of aerosol. This cluster is not pure, and probably comprises more than one type of aerosol. Besides the potentially maritime aerosol (coarse

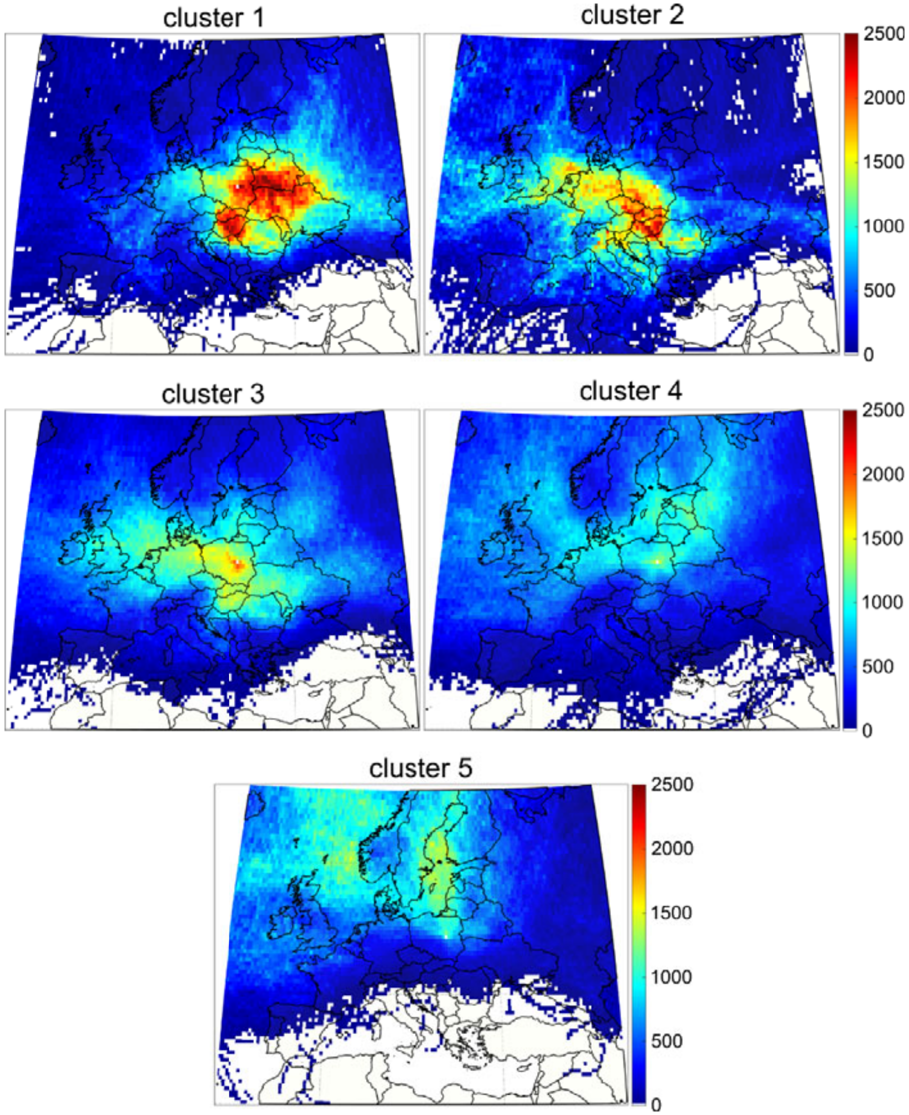


Fig. 4. The distribution of trajectories with endpoints (located 500 m above the receptor) concurrent with the 2D clustered measurements for Belsk station. The values represent the number of hours spent by all the trajectories over a grid point, normalized by the lengths of individual trajectories (see Section 2). Note different color scale than in Fig. 2.

particles with small AOTs), it contains continental aerosols (large values of EAE). This could be related to the advection of clear air mass, even cleaner than in the case of cluster 4 (see smaller AOTs). Maritime aerosol, if present, should not be interpreted as salty water droplets but rather as an aerosol containing sea salt, as all the used trajectories are 5 day long and the mean lifetime of sea salt particles in the atmosphere is in the order of 5-6 days (Junge and Gustafson 1957, Koch *et al.* 2006). In the case of an absence of precipitation on the path from sea salt source to the receptor site, it is reasonable to assume that the coarse aerosol originating from mentioned regions is in fact a maritime aerosol containing sea salt. Additional trajectory analysis based on a fixed threshold of EAE is performed in the next section.

3.3 Threshold method

The differences between source regions obtained in the two previous subsections, as well as the desire to check source of coarse aerosol with small AOT, prompted us to perform additional analysis of trajectories supported by aerosol typing based on predefined levels of AOT and EAE. Such a typing may be found in the literature (Eck *et al.* 1999, Pace *et al.* 2006, Kaskaoutis *et al.* 2007a, b; Kaskaoutis *et al.* 2009, Kalapureddy *et al.* 2009, Pawar *et al.* 2015). We decided to use thresholds of AOT and EAE listed in Table 3 for the predefined types of aerosols; thresholds used by other authors are, for example, listed in work of Pawar *et al.* (2015). Synergy of the 2D and the 3D clustering with the analysis of backward trajectories allowed us to set thresholds in a way to distinguish between urban and biomass burning aerosols and obtain, as separate as possible, source regions in trajectory analysis. Such a differentiation between biomass burning and urban/industrial aerosol was not found in the aforementioned analyses based on fixed thresholds.

Table 3

The classification of the typical aerosol types occurring over Central Europe based on their optical properties

Cluster	AOT	EAE
Biomass burning	over 0.55	over 1.3
Urban/industrial	0.3 – 0.55	over 1.3
Saharan dust	over 0.3	up to 1
Continental	0 – 0.2	over 1.3
Maritime	0 – 0.2	up to 1

Results of the preformed trajectory analysis are shown in Fig. 5. Thanks to the carefully chosen threshold, distinguishing between biomass burning and urban/industrial aerosol, possible source regions of such aerosols are separated. The largest trajectory densities, related to mineral dust, should not be

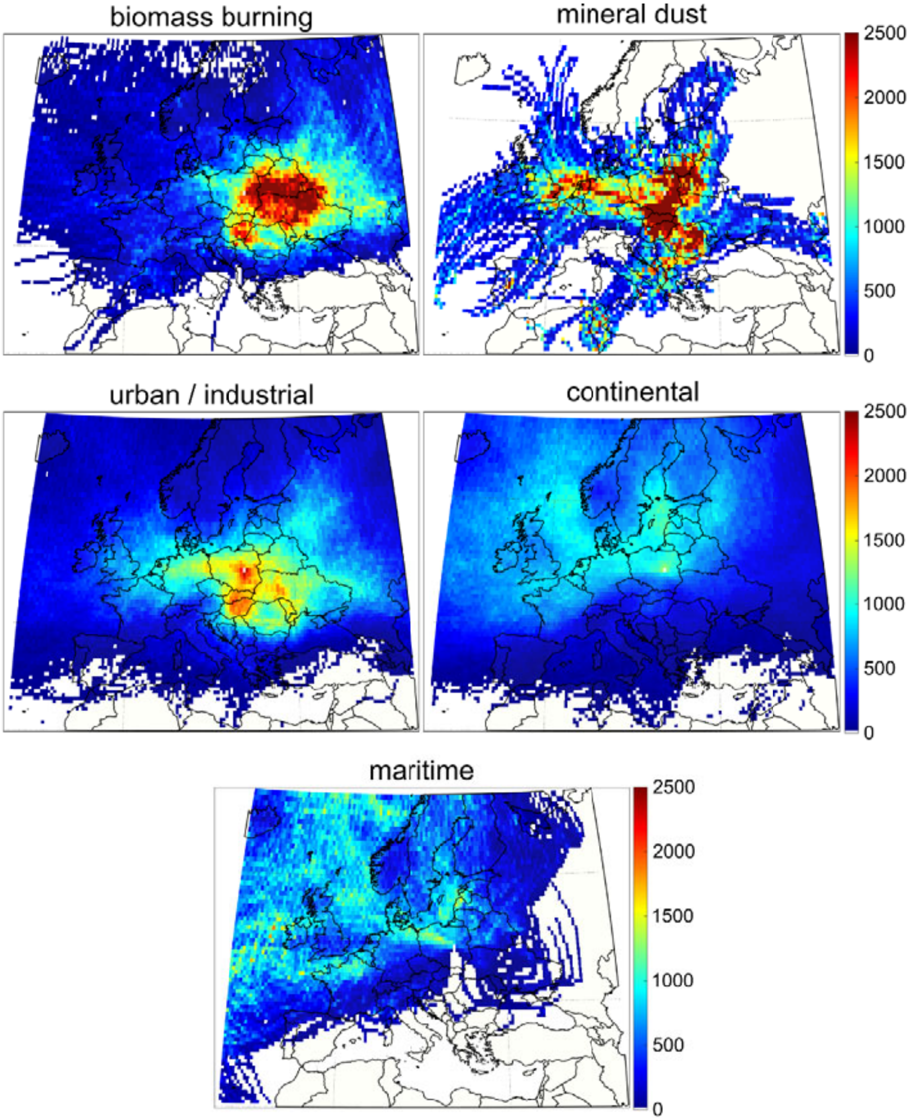


Fig. 5. Spatial density of the trajectories with endpoints (located 500 m above the receptor) concurrent with the fixed thresholds for AOT and EAE. Note different color scale than in Fig. 1.

interpreted as possible source regions as they correspond to a long-range transport, predominantly from northern Africa over typical paths of transport of mineral dust to Central Europe. Possible source regions obtained for continental aerosol are similar to those found for cluster 4 in the 2D clustering with the addition of North Sea and English Channel regions. This may indicate that aerosol, which is in this case attributed to the continental type, is in fact partly related to a fast moving air mass deriving from North Sea/Northern Atlantic. Results of the trajectory analysis related to aerosol described as maritime one are the most noisy, however the concentration of trajectories is clearly seen over North Sea and Northern Atlantic, west of Ireland.

4. SOURCE REGIONS AND THE CORRESPONDING AEROSOL TYPES

The proposed methods for identification of potential source regions of various aerosol types were applied to data gathered over Belsk, Poland. The analysis yielded several potential source regions, including: (a) a region comprising the vicinity of Belarus-Ukraine border, (b) Slovakia and Hungary, (c) western Poland and eastern Germany, and (d) Scandinavia, Baltic Sea and widely understood Northern Atlantic, including Northern Sea. It must be stressed here that these identified source regions are not always strictly restricted to a single cluster/threshold originating from the aforementioned analysis methods but they rather represent a dominant source of aerosol with a given set of optical parameters.

Aerosol transported from the region (a) is characterized by the largest observed daily mean AOTs (>0.5), large mean values of EAE (~1.6), and AAE mean values of 1 which suggests significant load of fine, strongly absorbing aerosol. Advection of air mass from this direction is usually concurrent with significantly increased concentrations of PM10 registered in central Poland (Pietruczuk and Jaroslawski 2013) and largest AOTs (Jaroslawski and Pietruczuk 2010, Kabashnikov *et al.* 2014). It should be mentioned that the correlation between AOT and PM10 is found to be statistically insignificant for both Belsk and Warsaw (Zawadzka *et al.* 2013). Moreover, significant increase of AOTs are observed during transport of aerosol from the vicinity of Minsk to Belsk (Pietruczuk 2013), showing regions situated east of Poland to be significant sources of aerosol. Detailed analysis of MODVOLC fire pixels, thermal anomalies registered by MODIS instrument (Wright *et al.* 2004) presented in Fig. 6, indicate Ukraine as a region affected by wild fires of both natural and anthropogenic origin (seasonal burning of barren vegetation). MODVOLC pixel density in northern Ukraine and southern Belarus is not as high as in the case of south eastern Ukraine but this region is still strongly affected by the wildfires of similar mixed origin.

The source region (b) is surrounded by the Carpathian arch from the north and east. This orography, together with the dominance of western wind circulation, may favor accumulation of pollutions in this region. A number of possible sources of aerosol may be found at this region including the strongly urbanized and industrialized areas located next to Danube river, *i.e.*,

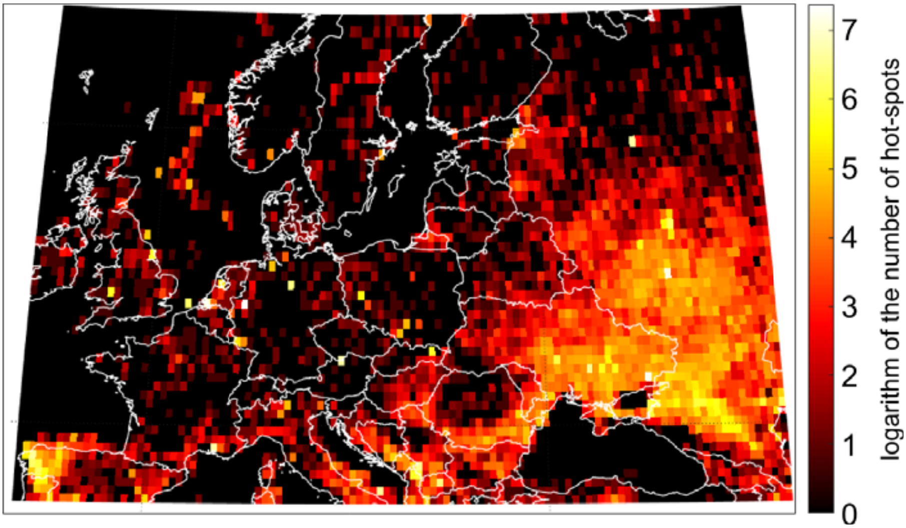


Fig. 6. Thermal anomalies over Europe obtained from MODVOLC algorithm; hot-spots at land are related to open fires and at the sea related to exploitation of oil fields. The data originates from MODIS (NASA) instrument for the 2000-2010 period.

Vienna, Bratislava, Budapest and heavy industry in the Kosice vicinity. Moreover, according to MODVOLC, biomass burning products may be transported from this region to central Poland. All that potential sources are clearly seen in Fig. 7 as hot spots of increased mean values of AOT taken by MODIS instrument and hot spots of MODVOLC thermal anomalies in Fig. 6. Advection of air mass from the south is typically related to increased PM10 concentrations and increased AOTs registered at Belsk. Although this region is a possible source of pollutants, the aerosol itself may be accumulated gradually as the air moves towards the receptor site. For example, a significant increase of AOT was observed during transport of aerosol from Vienna and Bratislava to Belsk over heavy industry dominated region situated near the Polish-Czech-Slovakian border (Pietruczuk 2013). It suggests that the industry south of central Poland, mainly in Upper Silesia and Ostrava regions, may be a significant aerosol source. This is supported by both of these regions being hot spots of exceeded permissible levels of particulate matter (EEA 2015).

The widely understood vicinity of the receptor site (c) should be also considered as a potential aerosol source. Hot spots of increased AOTs are clearly seen in western Poland and Saxony in eastern Germany; see Fig. 7. These are the densely populated areas surrounding Poznan, Lower Silesia region in Poland and Leipzig in Germany. Advection of relatively slow moving air mass from western direction is accompanied by observation of slight-

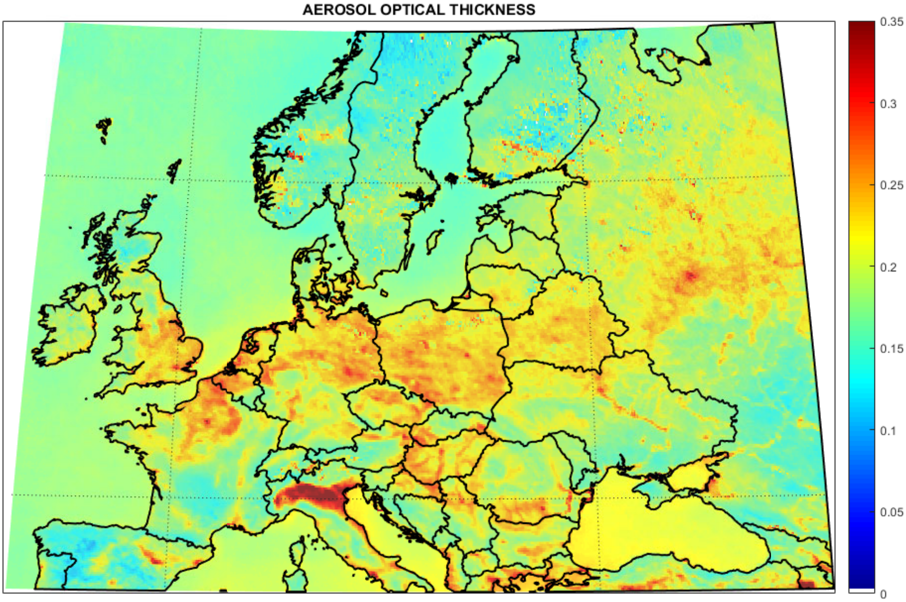


Fig. 7. Mean AOT at 500 nm over Europe obtained from the MODIS instrument (collection 6, L2), 2004-2015 period.

ly increased AOTs and PM10 (Pietruczuk and Jaroslawski 2013). Such kind of aerosol represented by cluster 3 in the 2D analysis was named polluted continental. Its mean AOT is 0.28 and mean EAE is 1.66. It is worth to note that this aerosol was not found in the Section 3.3 because it is mixed-type aerosol with optical properties placed between urban/industrial and continental one. Advection of fast moving air mass from the west is accompanied by observation of small AOTs related to cluster 5 in 2D analysis or continental and maritime aerosols analyzed in Section 3.3.

The last source region (d) is the widest and most differentiated. The pattern for maritime aerosols partially overlaps that obtained for aerosol described as a continental one. For example, in the cluster 5 (2D analysis) a large portions of Baltic and North Sea are visible alongside the continental Europe. This indicates that both marine and continental type observations are related to advection of clear, fast moving air masses from western and northern directions. This makes distinguishing between maritime sources over Baltic Sea and Northern Atlantic and continental sources over Scandinavia and British Isles difficult. It should also be noted that the lifetime of maritime aerosol is relatively short. Model simulations presented by Jaegle *et al.* (2011) suggest that a lifetime of sea salt below 4 μm radius is in the 12-25 h range. In principle, this is a sufficient time, in the case of fast moving south-bound trajectories, for aerosol to be transported above Belsk from the Baltic region. However, the mention of maritime aerosol in central or southern Poland may not be easily found in literature. Chemical analysis of $PM_{2.5}$ performed by Rogula-Kozłowska *et al.* (2012) in Zabrze (southern Poland) indicates that NaCl may constitute up to 5.8% of the ambient airborne dust during winter season and about 1.5% in the summer. The larger values of the NaCl fraction are expected in winter because chloride salts (both NaCl and KCl) are used in Poland to melt snow and ice on roads. However, the presence of NaCl in the summer $PM_{2.5}$ measurements may suggest that some sea salt is being transported as far as to the southern Poland. Moreover, the presence of the maritime aerosol should not be ruled out since the EAE values in cluster 5 are predominantly small, which is characteristic for this type of aerosol.

5. DISCUSSION AND CONCLUSIONS

The aforementioned analysis shows the existence of specific regions as potential sources of certain types of aerosols. However, aerosol typing based on statistical analysis of its optical properties alone may be insufficient to credibly distinguish the aerosol types for central Europe. Some of the obtained clusters may contain a mixture of different aerosol types. Such a cluster mixing occurs predominantly in the 2D clustering described in Section 3.2, where for example cluster 1 contains both biomass burning and urban/industrial aerosols. The cluster analysis applied on the AOT *versus* EAE space does not sufficiently separate aerosol types. This implies potential misclassification of aerosols with similar optical properties and consequently a poor separation of the potential source regions in the statistical analysis of the trajectories. The main drawback in this clustering approach is the lack of additional parameters for differentiating aerosols.

The significant non-zero skewness of AOT distribution may also be a limiting factor. Although the characteristic of this distribution is typically close to being log-normal, the low values dominate the data even when the logarithm of the AOT is considered. This causes the separation of aerosols characterized by small AOTs to be significantly better than in the case of aerosols with large AOT values. Therefore, we propose the use of an additional parameter (AAE) for differentiating aerosol types. Unfortunately, the retrieval of AAE is only available for large AOT cases and cloud-free conditions, whilst EAE is available for smaller AOTs and also in the case of broken clouds. Despite this limitation, a good separation of biomass burning, urban/industrial and mineral dust was obtained in Section 3.1, showing usefulness of this approach. A possible solution to the poor separation of aerosol types in both the 2D and the 3D clustering may be the use of fixed thresholds, as presented in Section 3.3. An introduction of disjoint, artificially implemented, ranges of aerosol optical properties may overcome the problem associated with the misclassification of aerosols characterized by similar properties. A proper appointment of the ranges proves difficult, however, and requires additional knowledge of the aerosol types and properties within the studied dataset. In this analysis the results obtained from the 2D and, more importantly, the 3D clustering were used for a proper definition of the thresholds for the fixed range approach.

The aforementioned analysis was supported by a statistical analysis of backward air mass trajectories in order to determine the possible source regions and validate the identification of aerosol types. The obtained results are in a good agreement with our previous findings where the clustering was applied to the backward trajectories. However, clustering of the trajectories enabled us to find directions from which aerosols characterized by a certain set of optical properties were advected. In this work we define probable source regions of predominant aerosol types. Estimated source regions lay in the directions found in our previous works. The largest AOTs are related to an advection of air mass from the eastern and the southern directions (Jarosławski and Pietruczuk 2010, Pietruczuk 2013). Moreover, largest PM10 concentrations were registered at Belsk during advection episodes from the same directions (Pietruczuk and Jarosławski 2013). This agrees with the finding of this work that aerosols characterized by largest AOTs, biomass burning aerosols, have their possible source of origin between Belarus and Ukraine, which are located directly east of Belsk. Urban/industrial aerosols characterized by a little smaller AOTs, however still large, originate from Slovakia and Hungary located south of Belsk. According to our previous findings, the cleanest air episodes are typically related to an advection from Scandinavia and from over the Northern Atlantic. This work clarifies that most probable source of such kind of aerosols are Baltic and Northern Seas. Slow moving trajectories, arriving from the western direction, are related to accumulation of aerosol within the air mass traveling over heavily urbanized and industrialized regions in Western Europe (Pietruczuk and Jarosławski 2013, Pietruczuk 2013). This may be clearly seen in the analysis of backward trajectories related to cluster 4 as an increased density of the trajectories over Western Poland end Eastern Germany.

To summarize, three methods for identification of aerosol type based on its optical parameters were studied. It was found that, although applicable in some cases, these methods prove to be insufficient for a reliable identification scheme for all of the aerosol types occurring over Belsk, Poland. The use of an aerosol absorption parameter could possibly allow for a significantly better cluster separation and therefore for a more precise aerosol identification in most of the cases, provided the AAE parameter was available for all the AERONET measurements, including the low AOT ones. The proposed auxiliary statistical analysis of backward air mass trajectories can provide the additional information regarding the aerosol in the form of its approximate source region or at least the general direction of the concurrent air advection. In the further studies of aerosols over Belsk, a synergy of statistical analysis of aerosol properties together with statistical analysis of backward trajectories should be used for aerosol classification.

Acknowledgments. This work was supported by the Polish National Science Centre under grant NCN 2013/09/B/ST10/03553. Authors would like to acknowledge AERONET station PIs: Albert Ansmann, Anatoli Chaikovsky, Piotr Sobolewski and Brent Holben for the use of the data from the AERONET stations. Authors also acknowledge the MODIS mission scientists and associated NASA personnel for the production of the data used in this research effort. The MODIS data were obtained from the Level-1 and Atmosphere Archive & Distribution System (LAADS).

References

Balis, D.S., V. Amiridis, C. Zerefos, E. Gerasopoulos, M. Andreae,P. Zanis, A. Kazantzidis, S. Kazadzis, and A. Papayannis (2003), Raman lidar and sunphotometric measurements of aerosol optical properties over Thessaloniki, Greece during a biomass burning episode, *Atmos Environ.* **37**, 32, 4529-4538, DOI: 10.1016/S1352-2310(03)00581-8.

Barnaba, F., and G.P. Gobbi (2004), Aerosol seasonal variability over the Mediterranean region and relative impact of maritime, continental and Saharan dust particles over the basin from MODIS data in the year 2001, *Atmos. Chem. Phys.* **4**, 9/10, 2367-2391, DOI: 10.5194/acp-4-2367-2004.

Barnaba, F., F. Angelini, G. Curci, and G.P. Gobbi (2011), An important fingerprint of wildfires on the European aerosol load, *Atmos. Chem. Phys.* **11**, 20, 10487-10501, DOI: 10.5194/acp-11-10487-2011.

Bergstrom, R.W., P.B. Russell, and P. Hignett (2002), Wavelength dependence of the absorption of black carbon particles: Predictions and results from the TARFOX experiment and implications for the aerosol single scattering albedo, *J. Atmos. Sci.* **59**, 3, 567-577, DOI: 10.1175/1520-0469(2002)059 <0567:WDOTAO>2.0.CO;2.

Berry, M.V., and I.C. Percival (1986), Optics of fractal clusters such as smoke, *J. Modern Optics* **33**, 5, 577-591, DOI: 10.1080/713821987.

Boselli, A., R. Caggiano, C. Cornacchia, F. Madonna, L. Mona, M. Macchiato, G. Pappalardo, and S. Trippetta (2012), Multi year sun-photometer measurements for aerosol characterization in a Central Mediterranean site, *Atmos. Res.* **104**, 98-110, DOI: 10.1016/j.atmosres.2011.08.002.

Burton, S.P., R.A. Ferrare, C.A. Hostetler, J.W. Hair, R.R. Rogers, M.D. Obland, C.F. Butler, A.L. Cook, D.B. Harper, and K.D. Froyd (2012), Aerosol classification using airborne High Spectral Resolution Lidar measurements-methodology and examples, *Atmos. Meas. Tech.* **5**, 1, 73-98, DOI: 10.5194/amt-5-73-2012.

Cattrall, C., J. Reagan, K. Thome, and O. Dubovik (2005), Variability of aerosol and spectral lidar and backscatter and extinction ratios of key aerosol types derived from selected Aerosol Robotic Network locations, *J. Geophys. Res.* **110**, D10, DOI: 10.1029/2004JD005124.

Chin, M., P. Ginoux, S. Kinne, O. Torres, B.N. Holben, B.N. Duncan, RV. Martin, J.A. Logan, A. Higurashi, and T. Nakajima (2002), Tropospheric aerosol optical thickness from the GOCART model and comparisons with satellite and Sun photometer measurements, *J. Atmos. Sci.* **59**, 3, 461-483, DOI: 10.1175/1520-0469(2002)059<0461:TAOTFT>2.0.CO;2.

Costabile, F., F. Barnaba, F. Angelini, and G.P. Gobbi (2013), Identification of key aerosol populations through their size and composition resolved spectral scattering and absorption, *Atmos. Chem. Phys.* **13**, 5, 2455-2470, DOI: 10.5194/acp-13-2455-2013.

Dubovik, O., A. Smirnov, B.N. Holben, M.D. King, Y.J. Kaufman, T.F. Eck, and I. Slutsker (2000), Accuracy assessments of aerosol optical properties retrieved from Aerosol Robotic Network (AERONET) Sun and sky radiance measurements, *J. Geophys. Res.* **105**, D8, 9791-9806, DOI: 10.1029/2000JD900040.

Dubovik, O., B. Holben, T.F. Eck, A. Smirnov, Y.J. Kaufman, M.D. King, D. Idier, T. Anre, and I. Slutsker (2002), Variability of absorption and optical properties of key aerosol types observed in worldwide locations. *J. Atmos. Sci.* **59**, 3, 590-608, DOI: 10.1175/1520-0469(2002)059<0590:VOAAOP>2.0.CO;2.

Dvorská, A., G. Lammel, and I. Holoubek (2009), Recent trends of persistent organic pollutants in air in central Europe-Air monitoring in combination with air mass trajectory statistics as a tool to study the effectivity of regional chemical policy. *Atmos. Environ.* **43**, 6, 1280-1287, DOI: 10.1029/2000JD900040.

Eck, T.F., B.N. Holben, J.S. Reid, O. Dubovik, A. Smirnov, N.T., O'Neill, I. Slutske and S. Kinne (1999), Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols, *J. Geophys. Res.* **104**, D24, 31333-31349, DOI: 10.1029/1999JD900923.

Eck, T.F.,B.N. Holben, J.S. Reid, N.T. O'Neill, J.S. Schafer, O. Dubovik, A. Smirnov, M.A. Yamasoe, and P. Artaxo (2003), High aerosol optical depth biomass burning events: A comparison of optical properties for different source regions, *Geophys. Res. Lett.* **30**, 20, DOI: 10.1029/2003GL01786.

EEA (2015), Air quality in Europe — 2015 report, European Environment Agency, DOI: 10.2800/62459.

Fuller, K.A., W.C. Malm, and S.M. Kreidenweis (1999), Effects of mixing on extinction by carbonaceous particles, *J. Geophys. Res.* **104**, D13, 15941-15954, DOI: 10.1029/1998JD100069.

Giles, D.M., B.N. Holben, T.F. Eck, A. Sinyuk, A. Smirnov, I. Slutsker, R.R. Dickerson, A.M. Thompson, and J.S. Schafer (2012), An analysis of AERONET aerosol absorption properties and classifications representative of aerosol source regions, *J. Geophys. Res.* **117**, D17, DOI: 10.1029/2012 JD018127.

Ginoux, P., M. Chin, I. Tegen, J.M. Prospero, B. Holben, O. Dubovik, and S.J. Lin (2001), Sources and distributions of dust aerosols simulated with the GOCART model, *J. Geophys. Res.* **106**, D17, 20255-20273, DOI: 10.1029/ 2000JD000053.

Gobbi, G.P., Y.J. Kaufman, I. Koren, and T.F. Eck (2007), Classification of aerosol properties derived from AERONET direct sun data, *Atmos. Chem. Phys.* **7**, 2, 453-458, DOI: 10.5194/acp-7-453-2007.

Groß, S., M. Esselborn, B. Weinzierl, M. Wirth, A. Fix, and A. Petzold (2013), Aerosol classification by airborne high spectral resolution lidar observations, *Atmos. Chem. Phys.* **13**, 5, 2487-2505, DOI: 10.5194/acp-13-2487-2013.

Groß, S., V. Freudenthaler, M. Wirth, and B. Weinzierl (2015), Towards an aerosol classification scheme for future EarthCARE lidar observations and implications for research needs, *Atmos. Sci. Lett.* **16**, 1, 77-82, DOI: 10.1002/asl2. 524.

Holben, B.N., D. Tanré, A. Smirnov, T.F. Eck, I. Slutsker, O. Dubovik, F. Lavenu, N. Abuhassen, and B. Chatenet (1999), Optical properties of aerosols from long term ground-based aeronet measurements. **In:** *Proc. ALPS99, 17-23 January 1999, Meribel, France*, WK1-O-19.

Holben, B.N., D. Tanre, A. Smirnov, T.F. Eck, I. Slutsker, N. Abuhassan, and G. Zibordi (2001), An emerging ground-based aerosol climatology: Aerosol optical depth from AERONET, *J. Geophys. Res.* **106**, D11, 12067-12097, DOI: 10.1029/2001JD900014.

Holben, B.N., T.F. Eck, I. Slutsker, A. Smirnov, A. Sinyuk, J. Schafer, D. Giles, and O. Dubovik (2006), AERONET's version 2.0 quality assurance criteria. **In:** *Asia-Pacific Remote Sensing Symp.*, International Society for Optics and Photonics, 64080Q-64080Q, DOI: 10.1117/12.706524.

Hsu, N.C., J.R. Herman, O. Torres, B.N. Holben, D. Tanre, T.F. Eck, A. Smirnov, B. Chatenet, and F. Lavenu (1999), Comparisons of the TOMS aerosol index with Sun-photometer aerosol optical thickness: Results and applications, *J. Geophys. Res.* **104**, D6, 6269-6279, DOI:10.1029/1998JD200086.

Jaeglé, L., P.K. Quinn, T.S. Bates, B. Alexander, and J.T. Lin (2011), Global distribution of sea salt aerosols: new constraints from in situ and remote sensing observations, *Atmos. Chem. Phys.* **11**, 7, 3137-3157, DOI: 10.5194/acp-11-3137-2011.

Jarosławski, J., and A. Pietruczuk (2010), On the origin of seasonal variation of aerosol optical thickness in UV range over Belsk, Poland, *Acta Geophys.* **58**, 6, 1134-1146, DOI: 10.2478/s11600-010-0019-4.

Junge, C.E., and P.E. Gustafson (1957), On the distribution of sea salt over the United States and its removal by precipitation, *Tellus* **9**, 2, 164-173, DOI: 10.1111/j.2153-3490.1957.tb01869.x.

Kabashnikov, V., G. Milinevsky, A. Chaikovsky, N. Miatselskaya, V. Danylevsky, A. Aculinin, D. Kalinskaya, E. Korchemkina, A. Bovchaliuk, A. Pietruczuk, P. Sobolewsky, and V. Bovchaliuk (2014), Localization of aerosol sources in East-European region by back-trajectory statistics, *Int. J. Remote Sens.* **35**, 19, 6993-7006, DOI: 10.1080/01431161.2014.960621.

Kalapureddy, M.C.R., D.G. Kaskaoutis, P. Ernest Raj, P.C.S. Devara, H.D. Kambezidis, P.G. Kosmopoulos, and P.T. Nastos (2009), Identification of aerosol type over the Arabian Sea in the premonsoon season during the Integrated Campaign for Aerosols, Gases and Radiation Budget (ICARB), *J. Geophys. Res.* **114**, D17, DOI: 10.1029/2009JD011826.

Kaskaoutis, D.G., P. Kosmopoulos, H.D. Kambezidis, and P.T. Nastos (2007a), Aerosol climatology and discrimination of different types over Athens, Greece, based on MODIS data, *Atmos. Environ.* **41**, 34, 7315-7329, DOI: 10.1016/j.atmosenv.2007.05.017.

Kaskaoutis, D.G., H.D. Kambezidis, N. Hatzianastassiou, P.G. Kosmopoulos, and K.V.S. Badarinath (2007b), Aerosol climatology: on the discrimination of aerosol types over four AERONET sites, *Atmos. Chem. Phys.* **7**, 3, 6357-6411, DOI: 10.5194/acpd-7-6357-2007.

Kaskaoutis, D.G., K.V.S. Badarinath, S. Kumar Kharol, A. Rani Sharma, and H.D. Kambezidis (2009), Variations in the aerosol optical properties and types over the tropical urban site of Hyderabad, India, *J. Geophys. Res.* **114**, D22, DOI: 10.1029/2009JD012423.

Kaufman, L., and P.J. Rousseeuw (1990), *Finding Groups in Data: An Introduction to Cluster Analysis*, John Wiley and Sons Inc., Hoboken, DOI: 10.1002/9780470316801.

Kaufman, Y.J., D. Tanré, and O. Boucher (2002), A satellite view of aerosols in the climate system, *Nature* **419**, 6903, 215-223, DOI: 10.1038/nature01091.

Kikas, U., A. Reinart, A. Pugatshova, E. Tamm, and V. Ulevicius (2008), Microphysical, chemical and optical aerosol properties in the Baltic Sea region, *Atmos. Res.* **90**, 2, 211-222, DOI: 10.1016/j.atmosres.2008.02.009.

King, M.D., Y.J. Kaufman, D. Tanré, and T. Nakajima (1999), Remote sensing of tropospheric aerosols from space: Past, present, and future, *Bull. Am. Meteorol. Soc.* **80**, 11, 2229-2259, DOI: 10.1175/1520-0477(1999)080<2229:RSOTAF>2.0.CO;2.

Koch, D., G.A. Schmidt, and C.V. Field (2006), Sulfur, sea salt, and radionuclide aerosols in GISS ModelE, *J. Geophys. Res.* **111**, D6, DOI: 10.1029/2004JD005550.

Kosmopoulos, P.G., D.G. Kaskaoutis, P.T. Nastos, and H.D. Kambezidis (2008), Seasonal variation of columnar aerosol optical properties over Athens, Greece, based on MODIS data, *Remote Sens. Environ.* **112**, 5, 2354-2366, DOI: 10.1007/s11270-013-1605-2.

Lack, D.A., and C.D. Cappa (2010), Impact of brown and clear carbon on light absorption enhancement, single scatter albedo and absorption wavelength dependence of black carbon, *Atmos. Chem. Phys.* **10**, 9, 4207-4220, DOI: 10.5194/acp-10-4207-2010.

Levy, R.C., L.A. Remer, and O. Dubovik (2007), Global aerosol optical properties and application to Moderate Resolution Imaging Spectroradiometer aerosol retrieval over land, *J. Geophys. Res.* **112**, D13, DOI: 10.1029/2006JD007815.

Lloyd, S. (1982), Least squares quantization in PCM, *IEEE Trans Inform. Theor*. **28**, 2, 129-137, DOI: 10.1029/2006JD0078.

Mishchenko, M.I., I.V. Geogdzhayev, B. Cairns, B.E. Carlson, J. Chowdhary, A.A. Lacis, and L.D. Travis(2007), Past, present, and future of global aerosol climatologies derived from satellite observations: A perspective, *J. Quant. Spectrosc. Rad. Trans.* **106**, 1, 325-347, DOI: 10.1016/j.jqsrt.2007.01.007.

Moosmüller, H., R.K. Chakrabarty, and W.P. Arnott (2009), Aerosol light absorption and its measurement: A review, *J. Quant. Spectrosc. Rad. Trans.* **110**, 11, 844-878, DOI: 10.1016/j.jqsrt.2009.02.035.

Omar, A.H., D.M. Winker, M.A. Vaughan, Y. Hu, C.R. Trepte, R.A. Ferrare, K.-P. Lee, C.A. Hostetler, Ch. Kittaka, R.R. Rogers, R.E. Kuehn, and Z. Liu (2009), The CALIPSO automated aerosol classification and lidar ratio selection algorithm, *J. Atmos. Ocean Technol.* **26**, 10, 1994-2014, DOI: 10.1175/2009JTECHA1231.1.

Pace, G., A.D. Sarra, D. Meloni, S. Piacentino, and P. Chamard (2006), Aerosol optical properties at Lampedusa (Central Mediterranean). 1. Influence of transport and identification of different aerosol types, *Atmos. Chem. Phys*. **6**, 3, 697-713, DOI: 10.5194/acp-6-697-2006.

Pappalardo, G., L. Mona, G. D'Amico, U. Wandinger, M. Adam, A. Amodeo, and A. Boselli (2013), Four-dimensional distribution of the 2010 Eyjafjallajökull volcanic cloud over Europe observed by EARLINET, *Atmos. Chem. Phys*. **13**, 8, 4429-4450, DOI: 10.5194/acp-13-4429-2013.

Pawar, G.V., P.C.S. Devara, and G.R. Aher (2015), Identification of aerosol types over an urban site based on air-mass trajectory classification, *Atmos. Res*. **164**, 142-155, DOI: 10.1016/j.atmosres.2015.04.022.

Perrone, M.R., F. De Tomasi, and G.P. Gobbi (2014), Vertically resolved aerosol properties by multi-wavelength lidar measurements, *Atmos. Chem. Phys.* **14**, 3, 1185-1204, DOI: 10.5194/acp-14-1185-2014.

Pietruczuk, A. (2013), Short term variability of aerosol optical thickness at Belsk for the period 2002-2010, *Atmos. Environ.* **79**, 744-750, DOI: 10.1016/j.atmosenv.2013.07.054.

Pietruczuk, A., and A. Chaikovsky (2012), Variability of aerosol properties during the 2007-2010 spring seasons over central Europe, *Acta Geophys.* **60**, 5, 1338-1358, DOI: 10.2478/s11600-012-0017-9.

Pietruczuk, A., and J. Jarosławski (2013), Analysis of particulate matter concentrations in Mazovia region, central Poland, based on 2007-2010 data, *Acta Geophys.* **61**, 2, 445-462, DOI: 10.2478/s11600-012-0069-x.

Remer, L.A., Y.J. Kaufman, D. Tanré, S. Mattoo, D.A. Chu, J.V. Martins, and B.N. Holben (2005), The MODIS aerosol algorithm, products, and validation, *J. Atmos. Sci.* **62**, 4, 947-973, DOI: 10.1175/JAS3385.1.

Robinson, N.H., H.M. Newton, J.D. Allan, M. Irwin, J.F. Hamilton, M. Flynn, and H. Coe (2011), Source attribution of Bornean air masses by back trajectory analysis during the OP3 project, *Atmos. Chem. Phys.* **11**, 18, 9605-9630, DOI: 10.5194/acp-11-9605-2011.

Rogula-Kozłowska, W., K. Klejnowski, P. Rogula-Kopiec, B. Mathews, and S. Szopa (2012), A study on the seasonal mass closure of ambient fine and coarse dusts in Zabrze, Poland, *Bull. Environ. Contam. Toxic.* **88**, 5, 722-729, DOI: 10.1007/s00128-012-0533-y.

Rousseeuw, P.J. (1987), Silhouettes: a graphical aid to the interpretation and validation of cluster analysis, *Comput. Appl. Math.* **20**, 53-65, DOI: 10.1016/0377-0427(87)90125-7.

Russell, P.B., R.W. Bergstrom, Y. Shinozuka, A.D. Clarke, P.F. De Carlo, J.L. Jimenez, J.M. Livingston, J. Redemann, O. Dubovik, and A. Strawa (2010), Absorption Angstrom exponent in AERONET and related data as an indicator of aerosol composition, *Atmos. Chem. Phys.* **10**, 3, 1155-1169, DOI: 10.5194/acp-10-1155-2010.

Russell, P.B., M. Kacenelenbogen, J.M. Livingston, O.P. Hasekamp, S.P. Burton, G.L. Schuster, M.S. Johnson, K.D. Knobelspiesse, J. Redemann, S. Ramachandran, and B. Holben (2014), A multiparameter aerosol classification method and its application to retrievals from spaceborne polarimetry, *J. Geophys. Res.* **119**, 16, 9838-9863, DOI: 10.1002/2013JD021411.

Seibert, P., H. Kromp-Kolb, U. Baltensperger, D.T. Jost, M. Schwikowski, A. Kasper, and H. Puxbaum (1994), Trajectory analysis of aerosol measurements at high alpine sites. **In:** *Proc. 4th EUROTRAC Symp. "Transport and Transformation of Pollutants in the Troposphere", 25-29 March 1996, Garmisch-Partenkirchen, Germany*, 689-693.

Smirnov, A., B.N. Holben, T.F. Eck, O. Dubovik, and I. Slutsker (2000), Cloud-screening and quality control algorithms for the AERONET database, *Remote Sens. Environ.* **73**, 3, 337-349, DOI: 10.1016/S0034-4257(00)00109-7.

Stein, A.F., R.R. Draxler, G.D. Rolph, B.J.B. Stunder, M.D. Cohen, and F. Ngan (2015), NOAA's HYSPLIT atmospheric transport and dispersion modeling system, *Bull. Am. Meteorol. Soc.* **96**, 12, 2059-2077, DOI: 10.1175/BAMS-D-14-00110.1.

Stohl, A., G. Wotawa, P. Seibert, and H. Kromp-Kolb (1995), Interpolation errors in wind fields as a function of spatial and temporal resolution and their impact on different types of kinematic trajectories, *J. Appl. Meteorol.* **34**, 10, 2149-2165, DOI: 10.1175/1520-0450(1995)034<2149:IEIWFA>2.0.CO;2.

Toledano, C., V.E. Cachorro, A.M. De Frutos, M. Sorribas, N. Prats, and B.A. De la Morena (2007), Inventory of African desert dust events over the southwestern Iberian Peninsula in 2000-2005 with an AERONET Cimel Sun photometer, *J. Geophys. Res.* **112**, D21, DOI: 10.1029/2006JD008307.

Torres, O., P.K. Bhartia, J.R. Herman, A. Sinyuk, P. Ginoux, and B. Holben (2002), A long-term record of aerosol optical depth from TOMS observations and comparison to AERONET measurements, *J. Atmos. Sci.* **59**, 3, 398-413, DOI: 10.1175/1520-0469(2002)059<0398:ALTROA>2.0.CO;2.

Wright, R., L.P. Flynn, H. Garbeil, A.J.L. Harris, and E. Pilger (2004), MODVOLC: near-real-time thermal monitoring of global volcanism, *J. Volcanol. Geothermal Res*. **135**, 29-49, DOI: 10.1016/j.jvolgeores.2003.12.008.

Valenzuela, A., F.J. Olmo, H. Lyamani, M. Antón, G. Titos, A. Cazorla, and L. Alados-Arboledas (2015), Aerosol scattering and absorption Angström exponents as indicators of dust and dust-free days over Granada (Spain), *Atmos. Res.* **154**, 1-13, DOI: 10.1016/j.atmosres.2014.10.015.

Zawadzka, O., K.M. Markowicz, A. Pietruczuk, T. Zielinski, and J Jaroslawski (2013), Impact of urban pollution emitted in Warsaw on aerosol properties. *Atmos. Environ*. **69**, 15-28, DOI: 10.1016/j.atmosenv.2012.11.065.

Received 8 March 2016
Received in revised form 26 September 2016
Accepted 18 October 2015

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2677-2716
DOI: 10.1515/acgeo-2016-0064

Early Thermal History of Rhea: The Role of Serpentinization and Liquid State Convection

Leszek CZECHOWSKI[1] and Anna ŁOSIAK[2,3]

[1]Institute of Geophysics, Faculty of Physics, University of Warsaw, Warszawa, Poland; e-mail: lczecho@op.pl

[2]Institute of Geological Sciences, Polish Academy of Sciences in Wrocław, Wrocław, Poland

[3]Department of Lithospheric Research, University of Vienna, Vienna, Austria

Abstract

Early thermal history of Rhea is investigated. The role of the following parameters of the model is investigated: time of beginning of accretion, t_{ini}, duration of accretion, t_{ac}, viscosity of ice close to the melting point, η_0, activation energy in the formula for viscosity, E, thermal conductivity of silicate component, k_{sil}, ammonia content, X_{NH3}, and energy of serpentinization, c_{serp}. We found that t_{ini} and t_{ac} are crucial for evolution. All other parameters are also important, but no dramatic differences are found for realistic values. The process of differentiation is also investigated. It is found that liquid state convection could delay the differentiation for hundreds of My. The results are confronted with observational data from Cassini spacecraft. It is possible that differentiation is fully completed but the density of formed core is close to the mean density. If this interpretation is correct, then Rhea could have accreted any time before 3-4 My after formation of CAI.

Key words: medium-sized satellites, thermal evolution, gravitational differentiation, serpentinization, Rhea.

1. INTRODUCTION

The group of medium-sized icy satellite (MIS) of Saturn consists of 6 bodies. The largest is Rhea with a radius of 764 km, and the smallest is Mimas with a radius of 199 km. All of MIS are spherical bodies. They consist of mixtures of rocks and ices.

The impact craters are the most common structures found on the surface of Rhea (*e.g.*, Plescia 1985, De Pater and Lissauer 2001, p. 203); therefore, it was classified as a dead body by Rothery (1992). However, the Cassini mission revealed presence of younger structures of tectonic and volcanic origin (*e.g.*, Jaumann *et al.* 2009).

For the present research (as well as for Czechowski 2012) the data concerning gravitational field are especially important. Analysis of the Doppler data acquired by the Cassini spacecraft yields the mass of Rhea and the quadrupole moments of its gravity field with high accuracy. Eventually, Iess *et al.* (2007) conclude: "the data exclude fully differentiated models in which the core would be composed of unhydrated silicates and the mantle would be composed of pure ice. […] The one model that fits the gravity data and is self-consistent when energy transport and ice melting are qualitatively considered is an "almost undifferentiated" Rhea, in which a very large uniform core is surrounded by a relatively thin ice shell containing no rock at all". This conclusion gives some bounds for our investigation.

The radar observations of Rhea give information about chemical composition of the surface layer. The layer consists mostly of the water ice surface with organic polymers and carbon dioxide present in small amounts. The ammonia bearing compounds are absent on the surface but radar indicates an increase in ammonia with depth. Anyway, the total content of ammonia in Rhea is probably low (Ostro *et al.* 2006, Prentice 2006).

A few papers considered thermal evolution of Rhea. Summaries of previous investigations of thermal history of Rhea (and other MIS) are given in Schubert *et al.* (2010), Matson *et al.* (2009), and Czechowski (2012). Rhea is probably the best (of all MIS of Saturn) object for modeling because there is no tidal heating. Matson *et al.* (2009) provides a limited information concerning thermal evolution of Rhea; this study stated that the temperature increase due to despinning could be up to ~20 K, and due to accretion ~90 K. Moreover, according to Barr and Canup (2008), Rhea was formed no earlier than 4 My after the CAI (Calcium-Aluminium-rich Inclusions in meteorites) condensation, if Rhea is indeed undifferentiated.

We used numerical methods similar to Czechowski (2012). However, we consider different problems. First, we discuss the role of meteorite data for research concerning Rhea. This part could be used also for other medium-sized icy satellites. Although the role of meteoritic data for models of MIS is known, it is rarely discussed in details. Second, the role of chemical composition and chemical reactions in Rhea is discussed. This problem has not been considered in Czechowski (2012) at all. Third, we investigate the role of different parameters important for thermal evolution. Moreover, we give more attention to details of the process of differentiation. A better model of grains' sinking is presented. Finally, the role of serpentinization for thermal evolution is calculated. To our best knowledge, it is the first thermal model of Rhea where serpentinization is included.

The paper is organized as follows. MIS and the basic properties of meteorite material are discussed in Section 2. Section 3 discusses heat sources and general properties of convection. The numerical model of heat transfer used for our calculations is presented in Section 4. The thermal history and the role of different parameters are discussed in Section 5. Conclusions are in the last section.

2. MATERIAL PROPERTIES OF MIS

2.1 Differentiation of the MIS

MIS consist of two types of materials characterized by very different physical properties: ices of volatiles (mostly H_2O with a small addition of CO_2 and NH_3) and a silicate-metal mixture. The average density of the icy component is relatively well constrained and varies within a narrow range of 930-1010 kg m^{-3}. The pressure in the center of Rhea is below 200 MPa, so only ices Ic, Ih, and (below temperature ~75 K) I XI are possible (see Fig. 1). Their densities are ~940 kg m^{-3} (Cogoni *et al.* 2011).

Densities used to represent the rocky component of MIS vary significantly, from 2500 kg m^{-3} (Schubert *et al.* 2007), through 2700 kg m^{-3} (Porco *et al.* 2006), up to as much as 3250 kg m^{-3} (Prialnik and Merk 2008) or even 3500 kg m^{-3} (Schubert *et al.* 2007, 2010). However, those highest values are not in agreement with the current understanding of the process of the formation and evolution of icy satellites (*e.g.*, Castillo-Rogez *et al.* 2007, Schubert *et al.* 2007, Zolotov 2007, Sohl *et al.* 2010).

The icy satellites of the giant planets are thought to be mostly formed by aggregation of thermally unprocessed, aqueously unaltered dust particles from the solar accretion disk (Canup and Ward 2009, Mosqueira *et al.* 2010, Coradini *et al.* 2010, Sohl *et al.* 2010). However, some addition of material by collisions with planetesimals is also possible (*e.g.*, Kargel *et al.* 2000). The constituents of the icy satellites' rocky fraction were anhydrous and reduced at the time of formation. They consisted mostly of Fe-Ni metal, FeS, Mg-, and Ca-silicates, Ca-Al oxides and organic matter. The initial chemical composition of the silicate phase was probably similar to either carbona-

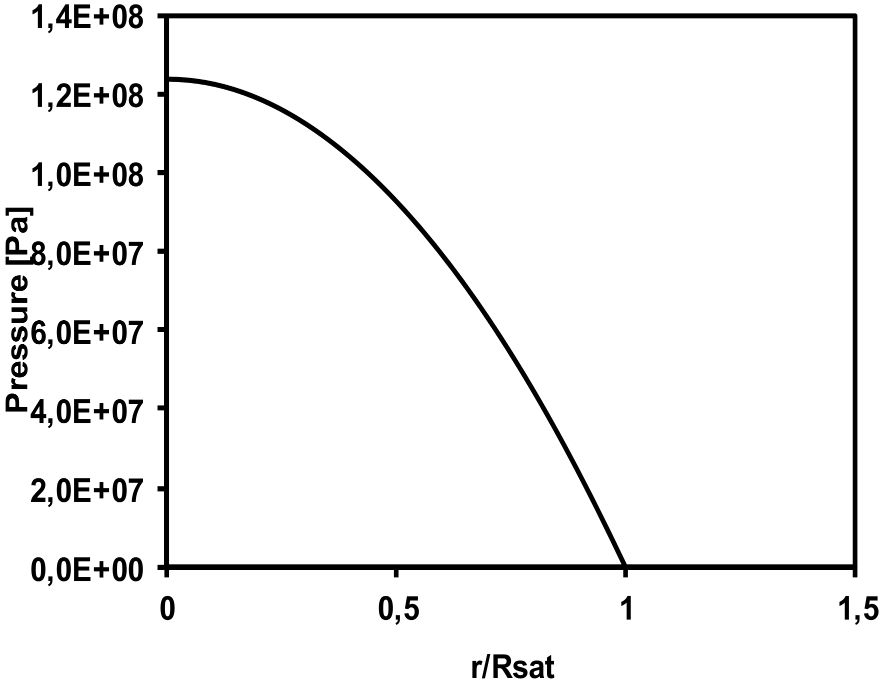


Fig. 1. Pressure distribution in Rhea. Note that the pressure in the center is below 200 MPa, so only ices Ic, Ih, and I XI are possible (Cogoni *et al.* 2011).

ceous chondrites (Kargel *et al.* 2000, Zolotov and Shock 2001, McKinnon and Zolensky 2003) or L/LL chondrites (Kuskov and Kronrod 2005).

After accretion, energy from the decay of radionuclides increased temperature of the satellites. Since most of radionuclides were enclosed within the silicate grains, the rocky particles were behaving as local heat sources and could melt the adhering ices. The first melts could have been formed under relatively low temperatures (Sohl *et al.* 2010) if the water ice was contaminated; *e.g.*, an eutectic brine of NH_3-H_2O can form at 175 K (Desch *et al.* 2009) and HCl hydrates melt at 186 K (Zolotov and Mironenko 2007). With the increase of temperature, the size of melt pockets increases and the liquid composition became more diluted (Zolotov and Mironenko 2007). The other possible sources of energy are: gravitational energy from silicate-ice differentiation (core forming), gravitational energy from contraction, kinetic energy from impacts, mineral hydration reactions, and tidal energy (*e.g.*, Schubert *et al.* 2007, Hussmann *et al.* 2010). For MIS, the energy of differentiation is not important. Moreover for Rhea the tidal energy is probably not important either (unless Rhea was in an orbit-orbit resonance in the past). The energy of impacts heats only upper layers, so its role for global processes is rather limited, except in a case of very catastrophic impacts.

If temperature was high for a long enough time, then melt pockets formed around silicate grains converged. If, moreover, the liquid volume is significant (~50%) then silicate grains are moving towards the center of mass forming a differentiated body with a rocky core and icy mantle (*e.g.*, Schubert *et al.* 2007, Sohl *et al.* 2010, Travis *et al.* 2012). If the temperature did not reach the melting temperature, the satellite remained undifferentiated. In the case of differentiated bodies, the temperature could continue to increase. If it approached 775-875 K, previously hydrated minerals started to disintegrate and their bounded water was removed from the mineral structure (*e.g.*, Peacock 2001, Llana-Funez *et al.* 2007), although subsequent hydration, after dropping the temperature, is very probable, especially since dehydration increases porosity (*e.g.*, Rutter *et al.* 2009). If the temperature rose to more than approximately 1275 K, the silicate-metal core could have differentiated (*e.g.*, Showman and Malhotra 1999, Sohl *et al.* 2002, Schubert *et al.* 2010). This situation probably has not occurred in MIS.

2.2 Post-differentiation hydrothermal activity

The rock-water interaction continued also after the process of MIS body differentiation was completed because the silicate cores were permeable and in contact with the ocean. Terrestrial rocks are permeable up to the pressure of 300 MPa (Walsh and Brice 1984); the permeability is probably of the order of $\sim 10^{-17}$ m^2 (*e.g.*, Stein *et al.* 1995, Sinha and Evans 2004, Sohl *et al.* 2010). It means that even large icy satellites (such as Europa) have probably 20-30 km thick layer of upper rocky core permeable for water.

In smaller bodies (such as Rhea), the pressure and its gradient are low; therefore, the entire silicate core should be permeable for water (Fig. 1). It means that if the lower part of the volatile-rich mantle is melted (forming an underground ocean at the core-mantle boundary) then the silicate-water interaction could continue after core formation (Sohl *et al.* 2010). This could have led to the formation of extensive hydrothermal circulation system within MIS bodies (Travis *et al.* 2012) that would influence the thermal history of Rhea by enabling serpentanization to occur over long time periods. The duration of extensive hydrothermal circulation depends on many factors. Note that the permeability could decrease as a result of a high pressure gradient and crystallization of secondary minerals within pores. These processes and the cooling of the satellite could reduce the Rayleigh number below its critical value (*e.g.*, Czechowski and Kossacki 2012). In the largest icy satellites such as Titan, Ganymede and Triton, the silicate-water interaction could have been limited by the formation of high pressure ices on the bottom of the sea (Sohl *et al.* 2010).

The two most important consequences of the water-rock interaction are: (i) dissolution of water-soluble elements in the liquid phase which significantly decreases freezing temperature, and (ii) formation of hydrated and oxidized minerals which replace anhydrous particles (*e.g.*, Hutchison 2004, Weisberg *et al.* 2006, Brearley 2006, Bland *et al.* 2009). Hydration of minerals can significantly change their physical properties, *e.g.*, it decreases the average density of material: during oxidation, metallic Fe (~7500 kg m^{-3}) transforms to much less dense magnetite (5120 kg m^{-3}), and hydration of olivine (3300 kg m^{-3}) leads to the formation of serpentine (~2500 kg m^{-3}). Additionally, the serpentinization process can serve as an additional source of heat which can be a driving force of the hydrothermal systems, especially in the later stages of satellite evolution (Vance *et al.* 2007).

The interaction of the rocky and icy phases based on the model described above is consistent not only with the thermochemical equilibria models (*e.g.*, Zolotov 2007), but also with (i) the analogue studies on meteorites from hydrated parent bodies (Brearley 2006), (ii) terrestrial low temperature weathering on Antarctica (Jull *et al.* 1988, Velbel *et al.* 1991, Losiak and Velbel 2011), (iii) the direct Cassini observations of the properties of water vapor plume and ice particles emerging from Enceladus (*e.g.*, Postberg *et al.* 2011), and (iv) the discovery of multiple non-icy materials (abundant hydrated sulfates: $MgSO_4 * nH_2O$, $Na_2SO_4 * nH_2O$, and $H_2SO_4*nH_2O$) on the surface of icy Galilean satellites of Jupiter (Dalton *et al.* 2005).

2.3 Meteorite properties

Remote sensing and numerical modeling show that non-volatile component of MIS has the same composition as hydrated chondritic meteorites (*e.g.*, Kargel *et al.* 2000, Zolotov and Shock 2001, McKinnon and Zolensky 2003, Kuskov and Kronrod 2005) such as Ivuna-type carbonaceous chondrite (CI) and Mighei-type carbonaceous chondrite (CM) (*e.g.*, Brearley 2006, Bland *et al.* 2009). All of the CI and CM meteorites, including those that were collected immediately after their fall on Earth, are extensively aqueously altered (*e.g.*, Hutchison 2004). Aqueous alteration resulted in the partial (CMs) or complete (CIs) disintegration of the chondrules, and formation of a phyllosilicate matrix. The contents of the bound (structural) water within those carbonaceous chondrites can be up to 22% (*e.g.*, Hutchison 2004, Weisberg *et al.* 2006).

The process of aqueous alteration on the CI and CM parent bodies occurred in relatively static water with ices constituting about 50% of volume (*e.g.*, Young *et al.* 2003); the metasomatic elemental fractionation occurred only over very short distances (*e.g.*, Young *et al.* 2003, Brearley 2004, 2006; Rubin *et al.* 2007, Bland *et al.* 2009). This statement is based on survival of highly water-reactive phases until terrestrial exposure (*e.g.*, Jull *et al.* 1988, Gounelle and Zolensky 2001, Losiak and Velbel 2011), and the fact that water-soluble element abundances are nearly uniform in bulk samples of all CMs. In the case of CM meteorites, the extent of silicate hydration was limited by the water availability (Velbel *et al.* 2012), while in the icy satellites the hydration was probably more restricted by the temperature.

The CI chondrites' matrix consists of serpentine interlayered with saponite and other phases embedded within them, such as magnetite, sulfides (pyrrhotite, pentlandite, cubanite), carbonates (calcite, dolomite, beunnerite, siderite), and sulfates (gypsum, epsomite, bloedite, Ni bloedite) (*e.g.*, Tomeoka and Buseck 1988, Brearley 2006). The measurements of physical properties of CI meteorites are rarely undertaken because they are: (i) very rare (only 5 separate falls are known), (ii) very valuable because of their pristine chemical composition, and (iii) extremely fragile. The Orgueil CI chondrite from the Vatican collection has a grain density of 2430 kg m^{-3} and a bulk density of 1580 kg m^{-3} (with porosity of 35%) (Consolmagno and Britt 1998, Macke *et al.* 2011). No data on the thermal conductivity of the CI meteorites is available.

The CM meteorites are less aqueously altered than CIs; they still have recognizable chondrules (about 20% of volume), calcium-aluminum-rich inclusions (about 5%) and even some remnants of metal grains (Brearley and Jones 1998). However, they are mainly composed of the matrix rich in phyllosilicates and other aqueously altered phases such as magnetite, various sulfides, sulfates, carbonates, halides (halite and sylvite), hydroxides (brucite) as well as silicates such as chlorite saponite and vermiculite (*e.g.*, Brearley 2006, Velbel *et al.* 2012). Since CMs are much more abundant, more measurements are available. The average grain density is 2920 kg m^{-3} (min 2740 kg m^{-3} and max 3260 kg m^{-3}) (Macke *et al.* 2011). The average bulk density of the CM type is 2200 kg m^{-3} (from 1880 kg m^{-3} to 2470 kg m^{-3}) (Macke *et al.* 2011), although it can be as low as 1662 kg m^{-3} (Opeil *et al.* 2010). The average porosity of CM chondrites is around 24.7% (from 15.0% to 36.7%.) (Macke *et al.* 2011). The thermal conductivity of the CM2 meteorite (Cold Bokkeveld, density 1662 kg m^{-3}) measured at 200 K is very low: 0.5 W m^{-1} K^{-1}, specific heat c_p = 500 J kg^{-1} K^{-1}, thermal diffusivity $\kappa = 6.02\times10^{7}$ m^2 s^{-1} and inverse thermal inertia $\Gamma = 15.51\times10^4$ m^2 $s^{1/2}$ kJ^{-1} (Opeil *et al.* 2010). The thermal conductivity k for the range of temperature between 6 K and 300 K increases with temperature and can be approximated by the equation (Opeil *et al.* 2010):

$$k = -0.0254 + 0.00563T - 2.07\times10^{-5}\,T^2 + 3.11\times10^{-8}\,T^3 .$$

Low values of thermal conductivity of those meteorites are mainly a result of their high porosity and bad contact between the grains. In MIS the pressure is higher, then the contact is better and possible pores are filled with ice (or water), that significantly changes the bulk conductivity. Therefore, we mainly used data for grains instead of the bulk values for porous meteorites.

3. HEAT SOURCES AND GENERAL PROPERTIES OF CONVECTION

3.1 Heat sources

Endogenic activity needs a source of heat. In the present paper we consider the following heat sources: decay of SLI (Short-Lived radioactive Isotopes, *i.e.*, Al^{26}, Fe^{60}, Mn^{53}), decay of LLI (Long-Lived radioactive Isotopes, *i.e.*, K^{40}, Th^{232}, U^{238}, U^{235}), the heat of serpentinization, and the heat of accretion (*e.g.*, Prialnik *et al.* 1987). The initial heat rates generated by a given isotope just after formation of CAI are given in Table 1. For any given time t the heat rate per 1 kg is:

$$Q(t) = f_m \sum_{i=1}^{7} Q_{oi} \exp\left(-C_i t\right), \qquad (1)$$

where f_m is a silicate mass fraction in the medium and time t is measured from the formation of CAI. Other symbols are explained in Table 1.

It is known that the tidal heating is an important source of heat in some satellites (*e.g.*, Czechowski 2006, Castillo-Rogez 2006) but for Rhea it could be neglected, because the tidal parameter ψ for Rhea is low; ψ is defined as:

Table 1

Data concerning radioactive isotopes considered in the paper (after Robuchon *et al.* 2010, modified)

Isotope	Half life $t_{1/2}$ [My]	Initial rate of heating per unit mass of element [W kg^{-1}]	Isotopic concentration *ppb* [10^{-9}]	Half life $t_{1/2}$ [s]	Decay constant C_i [s^{-1}]	Initial rate of heating per unit mass of silicate Q_{0i} [W kg^{-1}]
U238	4468	9.47E-05	26.2	1.409E+17	4.916E-18	2.48E-12
U235	703.81	5.69E-04	8.2	2.221E+16	3.120E-17	4.66E-12
Th232	14030	2.64E-05	53.8	4.427E+17	1.565E-18	1.42E-12
K40	1277	2.92E-05	1104	4.029E+16	1.720E-17	3.22E-11
AL28	0.716	3.41E-01	600	2.259E+13	3.067E-14	2.05E-07
Fe60	1.5	7.10E-02	200	4.733E+13	1.464E-14	1.42E-08
Mn53	3.7	2.70E-02	25.7	1.167E+14	5.936E-15	6.94E-10

$$\psi = MR_{\mathrm{sat}}^{\ 3}/(ma^3)\ , \tag{2}$$

where R_{sat} is the radius of satellite, m is its mass, M is the mass of Saturn, and a is the semimajor axis of the satellite orbit. For Rhea $\psi = 0.000748$ only. For more details, see discussion in Czechowski (2012).

The heat of accretion results from kinetic energy of the accreting matter. The temperature of an accreting layer T_{ac} could be calculated from the formula (*e.g.*, Multhaup and Spohn 2007):

$$T_{\mathrm{ac}} = (4/3\ c_{\mathrm{p}})\ \xi\ \pi\ G\ \rho\ R^2(t)\ (1 + 1/2e) + T_{\mathrm{nebula}}\ , \tag{3}$$

where $G = 6.67\ 10^{-11}\ \mathrm{N\ m^2\ kg^{-2}}$ is the gravity constant, ρ is the density, $R(t)$ is the current radius of the accreting body, c_p is the specific heat. We adopt also here constant $\xi = 0.4$ and the Safronov number $e = 2$ after Multhaup and Spohn (2007). The initial distribution of the temperature in our models is calculated using Eq. 3. The surface temperature on Rhea and the nebula temperature are assumed to be the same, namely $T_{\mathrm{s}} = T_{\mathrm{nebula}} = 75$ K. This is a typical value used in similar models (*e.g.*, Multhaup and Spohn 2007). It corresponds approximately to the present surface temperature of the Saturnian satellites. Some models use higher temperature of the nebula (*e.g.*, 250 K – see Multhaup and Spohn 2007). However, such high initial temperature of interior of Rhea could be presently excluded from the consideration because it leads to widespread melting and differentiation of the interior. The observation of Iess *et al.* (2007) suggests very limited differentiation only.

According to the considerations from previous section and discussion in Sections 2.1-2.3, we included the reaction of serpentinization as a source of heat. After Abramov and Mojzsis (2011) we consider the following reaction:

Mg_2SiO_4 (forsterite) + $MgSiO_3$ (enstatite) + $2H_2O$ --> $Mg_3Si_2O_5(OH)_4$ (antigorite) (4)

This reaction releases ~240 000 J per kg of serpentine produced (~69 kJ per mole). Similar reactions were considered by Malamud and Prialnik (2013). We are interested mainly in energy released during serpentinization; therefore, we do not discuss details of these chemical processes. The density of serpentine is 2470 kg $\mathrm{m^{-3}}$ but we assume that the silicates contain also 20% of nonreactive rock of density 3630 kg $\mathrm{m^{-3}}$ (Abramov and Mojzsis 2011). Although many values of density are used in models of MIS (see discussion in Section 2.1) we eventually have chosen the silicate density according to Abramov and Mojzsis (2011), *i.e.*, $2470 \times 80\% + 3630 \times 20\% = 2702$ kg $\mathrm{m^{-3}}$ and 240 000 J $\mathrm{kg^{-1}}$ for energy of serpentinization.

Table 2

Basic data and formulae used in the considered model

Symbol	Formula or value	Remarks	Reference
R_{Sat} [m]	764 300	Final radius of Rhea	Iess *et al.* (2007)
$R(t)$ [m]	$R(t) = A\,t$ for $t_{ini} < t < t_{ini} + t_{ac}$	Radius of Rhea during accretion	
ρ_{ice} [kg m^{-3}]	940	Density of ice	Iess *et al.* (2007)
ρ_{sil} [kg m^{-3}]	2470 80% + 3630 20%= 2702	Density of silicates	See text
ρ [kg m^{-3}]	1233	Density of Rhea	Iess *et al.* (2007)
f_V [1]	$f_V = (\rho - \rho_{ice})/(\rho_{sil} - \rho_{ice})$	Volume fraction of silicates	
f_m [1]	$f_m = \rho_{sil} f_V / \rho$	Mass fraction of silicates	
k_{ice} [W m^{-1} K^{-1}]	$0.4685 + 488.12/T$	Thermal conductivity of ice	Robuchon *et al.* (2010)
k_{sil} [W m^{-1} K^{-1}]	4.2 or 2.5	Thermal conductivity of silicates	See text
k [W m^{-1} K^{-1}]	$f_v\, k_{sil} + (1 - f_v)\, k_{ice}$	Thermal conductivity	
$c_{p\ ice}$ [J kg^{-1} K^{-1}]	$185 + 7.037\, T$	Specific heat of ice	Castillo-Rogez *et al.* (2007)
$c_{p\ sil}$ [J kg^{-1} K^{-1}]	920	Specific heat of silicates	Castillo-Rogez *et al.* (2007)
c_p [J kg^{-1} K^{-1}]	$f_m\, c_{sil} + (1 - f_m)\, c_{ice}$	Specific heat	
c_{serp} [J kg^{-1}]	241 000	Heat of serpentinization	Abramov and Mojzsis (2011)
$c_{p\ melts}$ [J kg^{-1}K^{-1}]	4180	Specific heat of water	
c_m [J kg^{-1}]	333×10^3	Latent heat of melting of ice	
T_m [K]	$273.15 - 7.95\times10^{-8}\, p - 9.6\times10^{-17}\, p^2 - 0.538\, X_{NH3} - 6.5\times10^2\, X_{NH3}^2 - 4.4\times10^8\, p\, X_{NH3}$	Temperature of melting (p is the pressure, X_{NH3} is the content of ammonia)	Leliwa-Kopystynski *et al.* (2002)
α [K^{-1}]	1.56×10^{-4}	Coefficient of thermal expansion	Hobbs (1974)
$g(r)$ [m s^{-2}]	$(4/3)\,\pi\, G\, \rho\, r$	Gravity at given r	
D [m]		Thickness of the considered spherical layer	

3.2 Solid state convection

Following Czechowski (2012) two types of large-scale motion in Rhea interior are considered in the present paper: solid state thermal convection (SSC) and thermal convection in liquid (LSC – Liquid State Convection). The SSC is a very slow convection observed in media that are solid (from microscopic point of view) but behave like liquid in specific conditions. Interiors of many planets and satellites could be treated as liquid in geological time. Some global tectonic structures on MIS could be interpreted as results of solid-state convection (*e.g.*, Czechowski and Leliwa-Kopystyński 2003, Schubert *et al.* 1986, Kargel and Pozio 1996). The very large values of the Prandtl number Pr are characteristic for this motion. Pr is defined as:

$$Pr = \eta \,/\, \rho\, \kappa \,, \tag{5}$$

where η [Pa s] denotes the effective viscosity, κ [$m^2\ s^{-1}$] is the thermal diffusivity, and ρ [$kg\ m^{-3}$] is the density. For solid state convection in Rhea $Pr = O(10^{20})$. It means that inertial terms in the equation of motion could be neglected (*e.g.*, Czechowski 1993).

The Rayleigh number Ra is used to characterize the intensity of convection. The intensity increases with Ra. The exact form of Ra depends on a specific physical situation. In situations considered here the heating for SSC is mainly from below, because the liquid water (in the core) is below its lower boundary, so the lower boundary could be treated as isothermal. Therefore, for SSC we use the following definition of the Rayleigh number Ra (*e.g.*, Schubert *et al.* 2001, p. 270):

$$Ra = \rho\, g_r\, \alpha\, \Delta T\, d^3 \,/\, (\kappa\, \eta) \,, \tag{6}$$

where g_r is the average gravity in the considered region [$m\ s^{-2}$], α is the coefficient of thermal volume expansion [K^{-1}], $\kappa = k/(\rho\, c_p)$ is the coefficient of temperature diffusion where c_p is the specific heat [$J\ kg^{-1}K^{-1}$], k is the thermal conductivity [$W\ m^{-1}\ K^{-1}$], ΔT is the temperature difference [K] across the considered region (layer), d is the thickness of this layer, and $\eta(T)$ is the effective (temperature dependent) viscosity [Pa s]. The $\eta(T)$ is evaluated for the average temperature T_{av} in the considered region (similar symbol T_{ave} is used here for the average temperature of the whole satellite). Other non-constant parameters (*e.g.*, κ, ρ) are also averaged over the considered region. The temperature difference $\Delta T = T_m - T_s$, where T_s is the surface temperature and T_m is the temperature of the molten region. If there is no melted core at all, then temperature in the center T_c is used instead of T_m. Convection starts if the Rayleigh number exceeds its critical value Ra_{cr}. Value of Ra_{cr} depends on the situation, but usually $Ra_{cr} = O(1000)$, (*e.g.*, Turcotte and Schubert 2002), so we choose $Ra_{cr} = 1000$. Note that sometimes Ra is defined in a different way; our value of Ra_{cr} corresponds to Ra defined using $\eta(T)$ evaluated for the average temperature T_{av} in the considered region (*e.g.*, Czechowski 1993, p. 221).

3.3 Liquid state convection

The second mode of thermal convection considered here (LSC) could take place in a molten region. The Prandtl number is $Pr = O(1)$. It means that inertial forces like Coriolis force could be important for this motion (*e.g.*, Czechowski 2012).

In the molten region the viscosity drops several orders, even down to $\eta = O(10^{-3})$ Pa s. It means that Ra increases several orders, up to 10^{14} or even more. The LSC could be very intensive, resulting in almost adiabatic temperature gradient. So in the molten core we assume the adiabatic gradient expressed by:

$$\frac{dT}{dr} = \frac{g\alpha_w T}{c_{pw}}, \tag{7}$$

where r is the radial distance (spherical coordinate), α_w is the thermal expansion coefficient, and c_{pw} is the specific heat for liquid water, g is the gravity in the core. Equation 7 indicates very low temperature gradient. Note also that coefficient of thermal expansion α_w of water is negative for $T < 277$ K at low and moderate pressures. For pressure corresponding to the central part of Rhea the anomaly of α_w disappears and $\alpha_w > 0$. For outer parts of Rhea, LSC requires temperature higher than 277 K. The serpentinization (starting after melting) could be a source of additional heat that leads to fast increase of the temperature of the molten region, up to $T > 277$ K, making the LSC possible just after the melting (without the need for the increase of temperature resulting from the radiogenic heating).

Czechowski (2014) indicates that velocity of LSC could be higher than the terminal velocity of typical grains (diameter ~1 mm) even in small Enceladus. The terminal velocity v_{term} is the velocity of a sinking grain where gravity is compensated by other forces. In Rhea, the velocity of LSC could be higher than in Enceladus. It means that (for some time) the mixing prevails over differentiation. The differentiation could occur only when LSC is slowing down (see Section 5.3 below). It means that for most of the time, LSC is getting heat from the inside (the grains with radioactive isotopes are suspended in the liquid), so we use Ra_{in} defined for internal heating as follows:

$$Ra_{in} = \alpha_m \rho_m^2 g d^5 Q(t)/(k_m \kappa_m \eta_m), \tag{8}$$

where subscript m denotes that given parameter is calculated for molten region of the satellite. Note that liquid water contains suspended solids, so the values of these parameters are not the same as for the water. The $\kappa_m = k_m / (\rho_m c_{pm})$ is the average value of coefficient of temperature diffusion, and k_m is average thermal conductivity [W m^{-1} K^{-1}], and $Q(t)$ is the rate of heating per 1 kg.

3.4 Viscosity of the solid mantle

The volume percentage of rocky component in Rhea is approximately ~17%. It means that rheological properties of its interior are determined by the icy component (Roscoe 1952, Schubert *et al.* 1986). According to Mangold *et al.* (2002) the creep is observed for the content of ice above 28%. However, for small content of the ice the viscosity is 10 to 50 times higher than for pure ice.

In general, the uppermost layer of the satellite ("lithosphere") is elastic for small deformation and brittle for large deformations. The medium below the lithosphere is also solid but for very slow geologic processes it behaves like a viscous fluid. There are no observational data concerning temperature-dependent viscosity $\eta(T)$ of the interior of MIS. Therefore, we should investigate the problem for some range of viscosity. After Czechowski (2012) we started from the formula:

$$\eta(T) = \eta_x \ \sigma^{(1-i)} \ \exp\left(\frac{E}{RT}\right), \tag{9}$$

where η_x is a constant, σ is the second invariant of the deviatoric stress tensor, i is the power law index ($i = 1$ corresponds to a Newtonian fluid), E is the activation energy of the dominant mechanism of deformation, and $R = 8.314$ [J K^{-1} mole^{-1}] is the universal gas constant (McKinnon 1998, Goldsby and Kohlstedt 1997, Durham *et al.* 1998, Forni *et al.* 1991). Parameters η_x, E and i depend on many factors; *e.g.*, size of ice crystal, content of gases, size of mineral grains *etc.*, most of them being essentially unknown. Durham *et al.* (1998) give two extreme values of E: 43 kJ mole^{-1} (for water ice I below 195 K) and 107.5 kJ mole^{-1} (for NH_3*$2H_2O$). These values correspond to E/R: 5172 K and 12 930 K, respectively. Rheology of both materials is highly nonlinear. Fortunately, non-Newtonian flow could be simulated by Newtonian flow with lower E (Christensen 1984, Dumoulin *et al.* 1999); therefore, we assume that $i = 1$ (*i.e.*, Newtonian flow) but we use $E = 50$ kJ kg^{-1} mole^{-1}.

The viscosity η_0 for the melting temperature T_m is used as a parameter in our model. It is equal to:

$$\eta_0 = \eta(T_m) = \eta_x \exp\left(\frac{E}{R\,T_m}\right). \tag{10}$$

However, we use here some results of Solomatov (1995) for the temperature-dependent viscosity convection. He considers the viscosity in the form:

$$\eta(T) = \eta_S \exp(-\gamma T), \tag{11}$$

where η_S and γ are constants (see also: Davaille and Jaupart 1993, Grasset and Parmentier 1998). To achieve the compatibility of Eqs. 9 and 11, we introduce:

$$\gamma = E/(R\,\zeta^2\,T_m^2), \tag{12}$$

where ζ is a parameter ($0 < \zeta \le 1$) and T_m denotes the melting temperature. The constant η_S is given by:

$$\eta_S = \eta_0 \exp\,[E\,(2/\zeta - 1)/R\,T_m]. \tag{13}$$

For such η_S, Eqs. 9 and 11 give the same value of viscosity for $T = T_m$ and $\varsigma = 1$, $i = 1$.

4. NUMERICAL MODEL

4.1 Basic equations of the model

The parameterized theory of convection used in Czechowski (2012) is chosen for the present research. It is based on the 1-dimensional equation of the heat transfer in spherical coordinates:

$$\rho\, c_p \frac{\partial T(r,t)}{\partial t} = \mathrm{div}(k(r,T)\,\mathrm{grad}\,T(r,t)) + q(r,T,t), \tag{14}$$

where r is the radial distance (spherical coordinate), ρ is the density [kg m^{-3}], c_p [J kg^{1} K^{-1}] is the specific heat, $q(r, T, t) = \rho\, Q$ [W m^{-3}] is the heat rate per unit volume, and k [W m^{-1} K^{-1}] is the thermal conductivity. Note that $Q(r, t)$ includes sources and sinks of the heat: radiogenic heat resulting from the decay of isotopes, latent heat of melting and latent heat of solidification. The heat of accretion is used to calculated temperature of the accreting material (see Eq. 3). This temperature is used as the boundary condition for temperature during accretion.

If the Rayleigh number in the considered layer exceeds its critical value Ra_{cr}, then the convection starts (*e.g.*, Czechowski 1993, p. 180; Czechowski and Kossacki 2009). The effect of convection can be described by dimensionless Nusselt number Nu (*e.g.*, Sharpe and Peltier 1978, Peltier and Jarvis 1982, Czechowski 2006, compare also Davies 2007). We use the following definition of Nu (*e.g.*, Czechowski 1993, p. 185; Czechowski and Kossacki 2012):

$$Nu = (\text{True total surface heat flow}) / (\text{Total heat flow without convection}) \; . \quad (15)$$

Now one can include the heat transport by convection simply multiplying the k in the considered spherical region:

$$k_{\mathrm{conv}} = Nu\, k \; , \quad (16)$$

where k_{conv} is an effective coefficient of the heat conduction. Of course, the temperature distribution obtained from 1D parameterized theory is different comparing to the true 3 D distribution. Fortunately, the details of temperature distribution are not important for our calculations. It is enough that the average temperature of the layer and total heat transfer through the layer will correspond to the true ones. This condition is satisfied.

For the Nusselt number, Czechowski (2012) uses formulae given by Solomatov (1995) for temperature-dependent viscosity. Solomatov (1995) considered 3 cases: the low viscosity contrast between hot interior and cold upper part of the convective layer, the medium viscosity contrast, and the high viscosity contrast. However, our preliminary calculations have shown that a simpler approach is satisfactory. Therefore, after Turcotte and Schubert (2002, p. 273) we use:

$$Nu = 1.04\, (Ra \,/\, Ra_{\mathrm{crit}})^{1/3} \; . \quad (17)$$

Equation 14 is solved for $R(t) \geq r \geq 0$, where $R(t)$ is the current radius of Rhea. During accretion the radius $R(t)$ increases in time, namely: $R(t) = A\, t$ (where A is a constant) for $t_{\mathrm{ini}} < t < t_{\mathrm{ini}} + t_{\mathrm{ac}}$. The constant A is chosen in such a way that $R(t) = R_{Sat} = 763$ km for $t = t_{\mathrm{ini}} + t_{\mathrm{ac}}$, *i.e.*, when accretion is completed. The linear function chosen for $R(t)$ means that the rate of mass accretion is increasing as $R(t)^3$ that agrees with the increasing geometrical and gravitational cross-sections of the accreting body. The linear function is also used by Merk *et al.* (2002). The temperature of accreted matter depends on the radius $R(t)$ (and consequently on the mass) of the proto-Rhea and is given by Eq. 3.

5. ROLE OF PARAMETERS FOR THERMAL EVOLUTION

5.1 Temperature limit

The upper limit for the temperature (*i.e.*, the potential temperature) inside Rhea can be calculated neglecting the heat transport. Graphs of the potential temperature (both upper lines) and potential melting (lower lines) for Rhea versus time of start of accretion t_{ini} (counted from forming CAI) are given in Fig. 2. There are included: the heat production from the radioactive decay during 20 My (after time t_{ini}), the heat of serpentinization (if melting has oc-

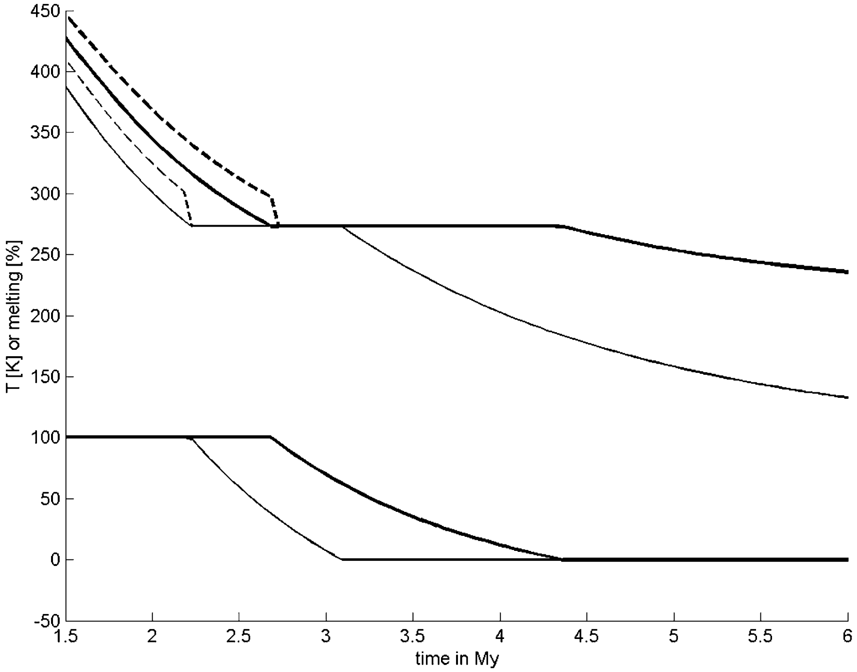


Fig. 2. Graphs of potential temperature (the solid and dashed upper lines) and potential melting (lower solid lines) for Rhea versus time of the beginning of accretion t_{ini}. The dashed lines give results for total heat of radioactive decay during 20 My (after time t_{ini}) and serpentinization. The thick lines included also the heat of accretion. The heat transfer is neglected so all heat is used for temperature increase and melting. The solid lines are for heating without serpentinization. The effect of serpentinization is seen as sharp steps on the upper lines. The value of 100% for lower lines (presenting potential melting) means that total latent heat of ice melting is supplied.

curred) and the heat of accretion. All heat is used for the temperature increase and/or melting (note that melting occurs only if the latent heat of melting is transferred to the medium). The solid lines give results for radiogenic heat and serpentinization. The dotted line is for heating without serpentinization. For early accretion (*i.e.*, t_{ini} =1 My) the maximum temperature is ~500 K. For more realistic t_{ini} = 2 My, the upper limit for T is ~400 K. The effect of serpentinization is seen as the sharp steps on the plots of temperatures (upper solid lines). The release of chemical energy is possible only in the molten medium. If heat of accretion is negligible then for t_{ini}> 2.2 My melting is not possible as well as differentiation and serpentinization. Therefore, the dashed lines representing temperature without serpentinization are below the solid lines. Note also that the temperature of accreting material

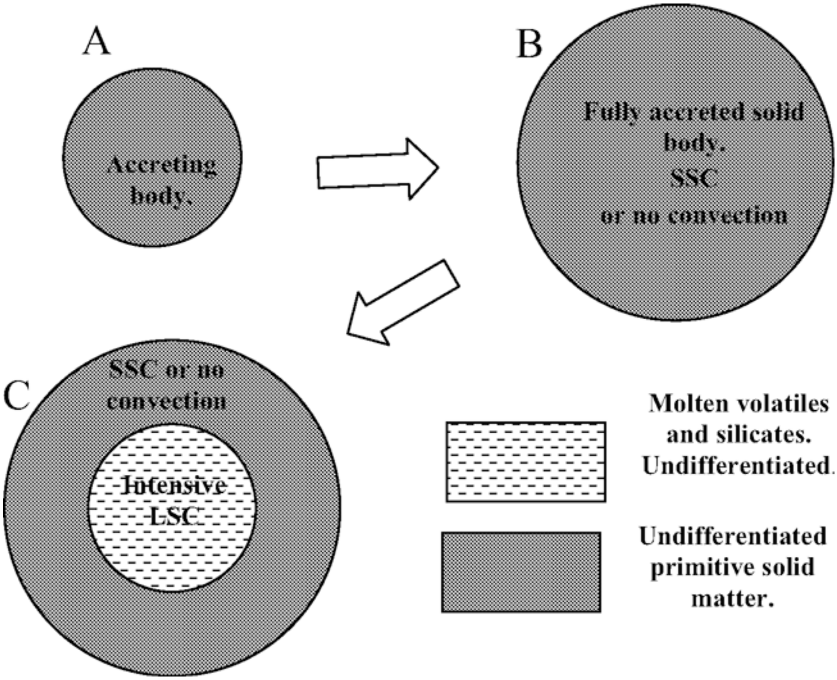


Fig. 3. Scheme of the considered evolution: A − accreting body, B − fully accreted solid body, C − if heating from SLI is intensive enough then part of the body is melted forming two layers body (melted: "core" and solid: "mantle"). After Czechowski (2012), modified.

(calculated according to Eq. 3) depends on the current radius of the body during accretion. Therefore, two cases are presented: for the center of the satellite (no heat of accretion – thick lines) and for the surface (where maximum of energy is released – thin lines). If the heat of accretion is included then melting (as well as differentiation and serpentinization) is possible even for t_{ini}~2.7 My.

The scheme of the considered evolution is given in Fig. 3. We start from accretion (Fig. 3A). The accreting body is solid, with initially inverse temperature gradient (*i.e.*, $\partial T/\partial r > 0$). The fully accreted solid body is also initially solid, but the negative temperature gradient is established in a few My, so SSC is possible for most of the time (Fig. 3B). The stage of evolution presented in Fig. 3C is reached only if heating from SLI is intensive enough to melt part of the satellite. Then two layers are formed, melted central sphere referred here as a "core" and solid layer referred here as a "mantle". In the core one could expect intensive LSC. SSC could operate in the solid mantle. The scheme is rather simple but interactions of the processes give quite complicated possible evolutionary paths, which are discussed below.

5.2 Role of model's parameters

The values of the most important parameters of our model are not known with a good accuracy. Therefore, we investigate thermal history of Rhea for different values of them to determine the role of each parameter. The role of following parameters is considered: time of beginning of accretion, t_{ini} , viscosity of ice close to the melting point, η_0, activation energy in the formula for viscosity, E, duration of accretion, t_{ac}, energy of serpentinization, c_{serp}, and content of ammonia, X_{NH3}. Figures 4-9 present evolution of globally averaged temperature T_{ave} (weighted average) and temperature in the center of the body T_c. T_{ave} and T_{c} represent the thermal state of Rhea and therefore are the best for our purpose. Horizontal axis gives $\log_{10}(t)$, where t is the time from CAI formation in My.

Figure 4 gives results for the following set of values of parameters: $\eta_0 = 10^{14}$ Pa s, $E = 50$ kJ kg^{-1} mole^{-1}, t_{ac}=0.1 My, $c_{\mathrm{serp}} = 240\,000$ J kg^{-1}, $X_{\mathrm{NH3}} = 0$, $k_{\mathrm{sil}} = 4.2$ Wm^{-1} K^{-1}. Moreover, 3 values of time of the beginning of accretion t_{ini} are used: 1.2, 2.4, and 4.2 My. These values will be referred here as the "basic" values and the model that uses them as the basic model. Comparing to Czechowski (2012) we use different values of t_{ini}, t_{ac} and η_0. The increase of η_0 is based on the results of Mangold *et al.* (2002) discussed in Section 3.4. We try also shorter $t_{ac} = 0.1$ My (it is justified by the results of Mosqueira *et al.* 2010, Merk *et al.* 2002). Later accretion characterized by larger t_{ini} is a consequence of shorter t_{ac}. Note that low t_{ini} means a large degree of differentiation.

Consider now the case for t_{ini}=1.2 My. The model is simple but one can see a few stages of evolution. They are (all given times are counted from the formation of CAI):

- ❑ From 1.2 to 1.3 My − accretion. In this case (short duration of accretion, $t_{ac} = 0.1$ My) the inverse gradient of temperature (*i.e.*, $\partial T/\partial r > 0$) is maintained to the end of accretion; consequently, SSC is not possible.
- ❑ From 1.3 to 2 My. SLI and LLI have supplied enough heat for establishing normal temperature gradient (*i.e.*, $\partial T/\partial r < 0$). It leads to an outburst of very intensive SSC (like an overturn, $Ra \approx 10^7\, Ra_{cr}$). Large loss of the heat results in substantial decreases of T_{ave}. This intensive SSC lasts from ~1.3 to ~1.9 My. During this time the centre of Rhea reaches the melting temperature. The latent heat stabilizes the temperature of the center T_{c} from ~1.5 to ~2 My.

- From 2 to 2.1 My − rapid increase of the molten region. The thickness of the solid "mantle" decreases from 100% of the radius of Rhea R_{sat} down to ~21% of R_{sat}. This results in a large decrease of intensity of SSC (see Eq. 6).
- From 2.1 to 7 My. Intensive SSC in the solid "mantle" with $Ra \approx 10^4 Ra_{cr}$ and intensive LSC in the molten "core" are developed. LSC maintains an adiabatic gradient of the temperature in the "core". The velocity of LSC is high and consequently LSC does not allow differentiation of the melted region. The thickness of the solid mantle decreases down to ~14% of R_{sat}. Maxima of T_c and T_{ave} are reached at ~7 My.
- From 7 to 400 My. In general, very slow cooling but the thickness of the solid mantle is still decreasing down to ~7% of R_{sat} at 70 My. The intensity of both LSC and SSC decreases because radiogenic heating by SLI is decreasing. However, even for 400 My (the end of our modeling) both modes of convection operate and the radius of molten region is still ~78% of R_{sat}.

Consider now the case $t_{ini} = 2.4$ My. SLI have supplied enough heat for establishing the normal temperature gradient (*i.e.*, $\partial T/\partial r < 0$) at ~2.7 My and SSC starts and its intensity increases. The maximum value of $Ra \approx 10^6 Ra_{cr}$ is reached at 5 My. However, there is no rapid outburst and there is no temporary decrease of T_{ave}. The maximum of T_{ave} is reached at $t = 3.2$ My. It is followed by slow decreasing of Ra and T_{ave}. T_c reaches T_m at 4.2 My and stabilizes to 220 My. Eventually no part of satellite is melted and the stage C from Fig. 3 is never reached. Also there is no heating from serpentinization. After 220 My, T_c starts to decrease.

Consider now the case $t_{ini} = 4.2$ My. The role of SSC is moderate for thermal evolution. SSC starts at 5.4 My and its intensity increases very slowly. Maximum $Ra \approx 10^2 Ra_{cr}$ is reached at 120 My. T_c reaches its maximum at 200 My.

Concluding, results for the "basic" model indicate the crucial role of t_{ini} for evolution. The melting, differentiation and forming core are possible only for the low value of t_{ini} (see also Fig. 3). The final (*i.e.*, at ~400 My) thermal states of models without melting (*i.e.*, here for $t_{ini} = 2.4$ and 4.2 My) are essentially the same because they differ only in the initial heat input. The final state of models with melting is not the same, because of different thermal properties of core, different distribution of LSI, *etc*.

Compare now models with different values of chosen parameters. The role of η_0 is presented in Fig. 5. The value of η_0 is increased from "basic" 10^{14} to 10^{16} Pa s. Graphs of two models are plotted together; the thinner lines are for the basic model and thick lines are for model with $\eta_0 = 10^{16}$ Pa s. Higher η_0 means lower Ra and consequently decreased efficiency of cooling.

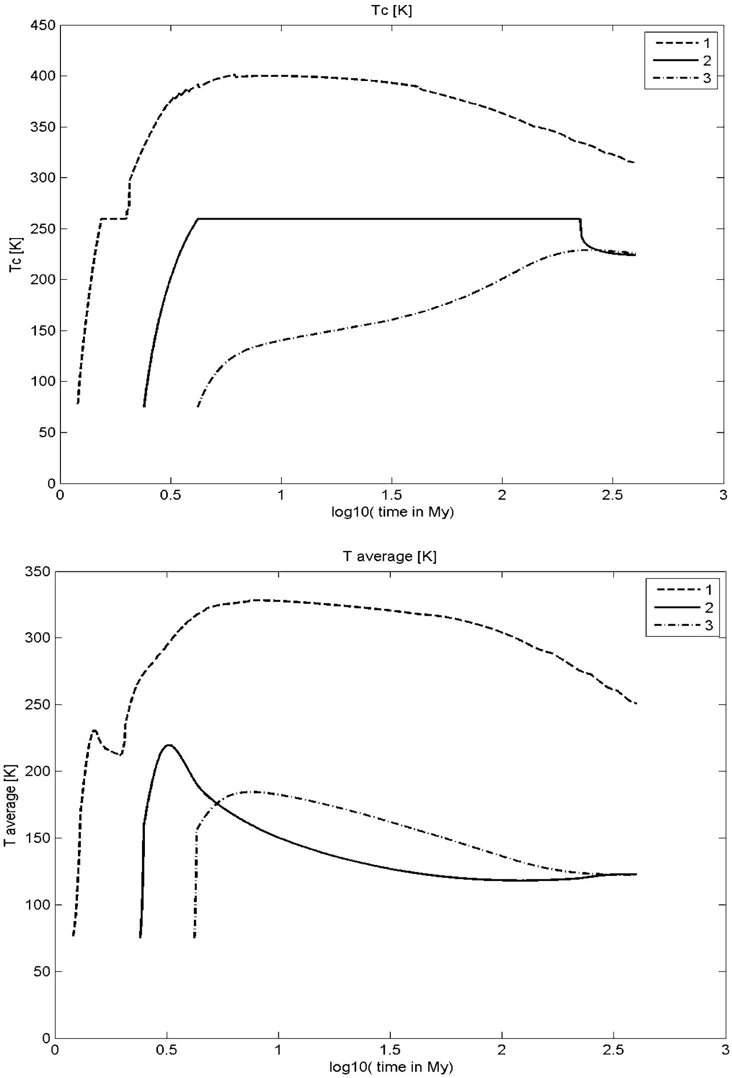


Fig. 4. The results for the "basic" model, *i.e.*, for the following values of parameters: viscosity parameter $\eta_0 = 10^{14}$ Pa s, the activation energy $E = 5\times10^4$ J mole^{-1}, the duration of accretion $t_{ac} = 0.1$ My, the energy for serpentinization $c_{serp} = 240\ 000$ J kg^{-1}, and the ammonia content $X_{NH3} = 0$. The upper panel presents the evolution of temperature in the center of the satellite T_c *versus* time t. The lower panel presents average temperature of the satellite T_{ave}. Horizontal axis gives $\log_{10}(t)$, where t is time from CAI formation in My. Evolutions for 3 different times of beginning of accretion t_{ini} are given. Lines from the left to the right correspond to: $t_{ini} = 1.2$ My (dashed line 1), 2.4 My (solid line 2), 4.2 My (dashed-dotted line 3), respectively. More explanation in the text.

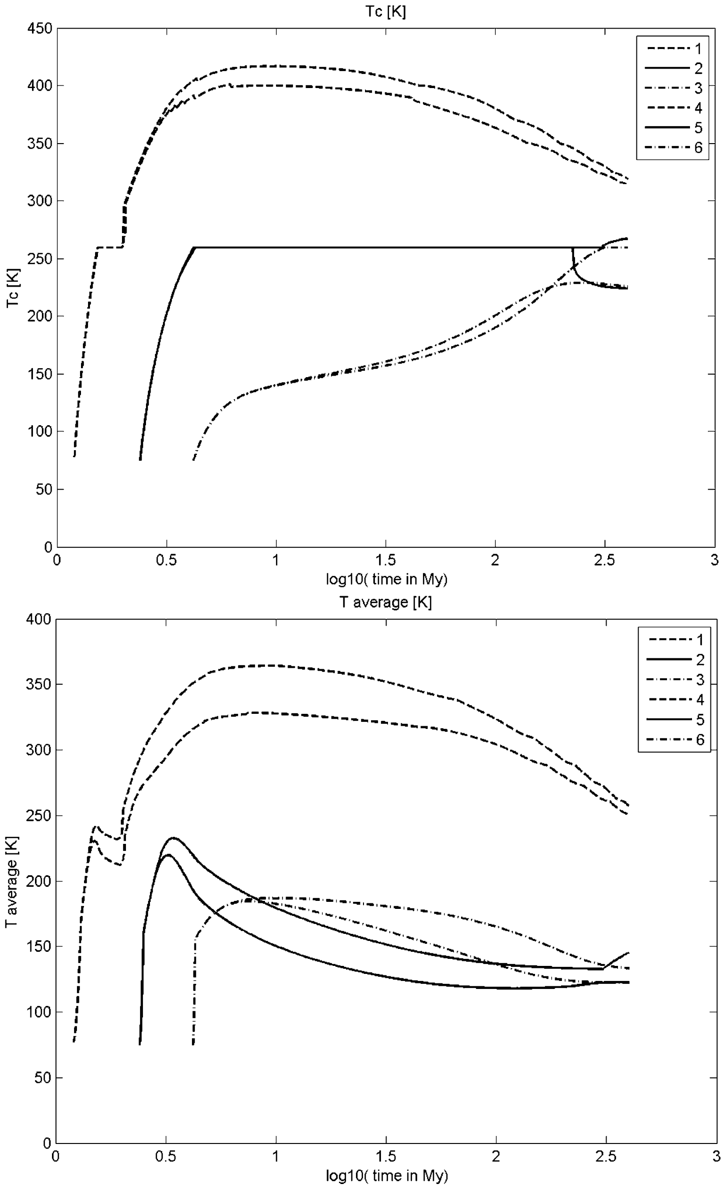


Fig. 5. The role of viscosity η_0 for thermal evolution. Two values of η_0 are used: 10^{14} Pa s (thin lines) and 10^{16} Pa s (thick lines). The higher value of η_0 reduces the role of SSC in the solid part of the satellite. The rest of parameters have the "basic" values, *i.e.*, $t_{ac} = 0.1$ My, $c_{serp} = 240\,000$ J kg^{-1}, and $E = 5\times10^4$ J mole^{-1}, $k_{sil} = 4.2$ W m^{-1} K^{-1}, $X_{NH3} = 0$. Lines from the left to the right correspond to $t_{ini} = 1.2$ My (dashed lines: 1 and 4), 2.4 My (solid lines: 2 and 5), 4.2 My (dashed-dotted lines: 3 and 6), respectively.

Therefore, generally temperatures for the "basic" model are below temperatures for $\eta_0 = 10^{16}$. Note also that for $\eta_0 = 10^{16}$, for the case of $t_{ini} = 2.4$ My (thick solid line in the upper panel) temperature in the center T_c exceeded melting temperature after ~380 My. This is possible because some heat from accretion released in upper layers is conducted to the center of the satellite (compare discussion in Section 5.1 and Fig. 2). Melting means also a release of serpentinization energy. Eventually both, T_c and T_{ave}, increase after ~380 My.

The case for $t_{ini} = 4.2$ My in Fig. 5 is interesting because the lines for the basic model are (for some time) slightly above lines for $\eta_0 = 10^{16}$ Pa s. Such nonlinear effects are common for evolution determined by SSC. Later, the lines intersect at ~180 My and eventually lines for basic model merge with the lines for $t_{ini} = 2.4$ My. The merging is expected because for 400 My SLI virtually does not exist and SSC is very weak. Therefore, the final thermal state of undifferentiated Rhea is determined by LLI and thermal conductivity only.

The role of energy of activation E is presented in Fig. 6. The value of E is decreased from the "basic" 5×10^4 J mole^{-1} to 3×10^4 J mole^{-1}. Lower E means lower viscosity. It increases Ra and consequently increases the efficiency of cooling; therefore, lines for the 'basic' model are generally above the lines for $E = 3\ 10^4$. Note a substantial difference of evolution of T_c for $t_{ini} = 2.4$ My (thick solid line in the upper panel) comparing to the basic model (thin solid line in the upper panel); the duration when $T_c = T_m$ is much shorter comparing to the basic model. For $t_{ini} = 4.2$ My both lines intersect similarly to the case presented in Fig. 5.

Figure 7 presents effects of longer duration of accretion t_{ac}. The longer t_{ac} means that more energy of accretion and radiogenic energy is lost because of small size of accreting body. For $t_{ini} = 1.2$ the effect is very substantial; T_c and T_{ave} are lower on ~100 K for ~300 My. It is also a result of SSC in accreting body (SSC starts at 1.5 My, *i.e.*, 0.3 My after start of accretion). For $t_{ini} = 2.4$ My the effect is weaker; SSC starts at 3.1 My, T_{ave} is lower on ~50K for a few My, the duration of $T_c = T_m$ is much shorter (the solid lines in the upper panel of Fig. 7). For $t_{ini} = 4.2$ My the effect is even weaker but still visible.

The role of energy of serpentinization c_{serp} for thermal evolution is given in Fig. 8. Two values of c_{serp} are used: 0 J kg^{-1} (thick lines) and the "basic" value 240 000 J kg^{-1} (thin lines). The effect is possible only if serpentinization have occurred, *i.e.*, if some part of the Rhea is melted. Therefore, the temperature drop on ~20 K is observed only for $t_{ini} = 1.2$ My.

Figure 9 presents the role of ammonia content X_{NH3}. Two values of X_{NH3} are used: $X_{NH3} = 0.1$ (thick lines) and $X_{NH3} = 0$ (thin lines). The main effect of NH_3 is a decrease of the melting temperature T_m (see Table 2). For

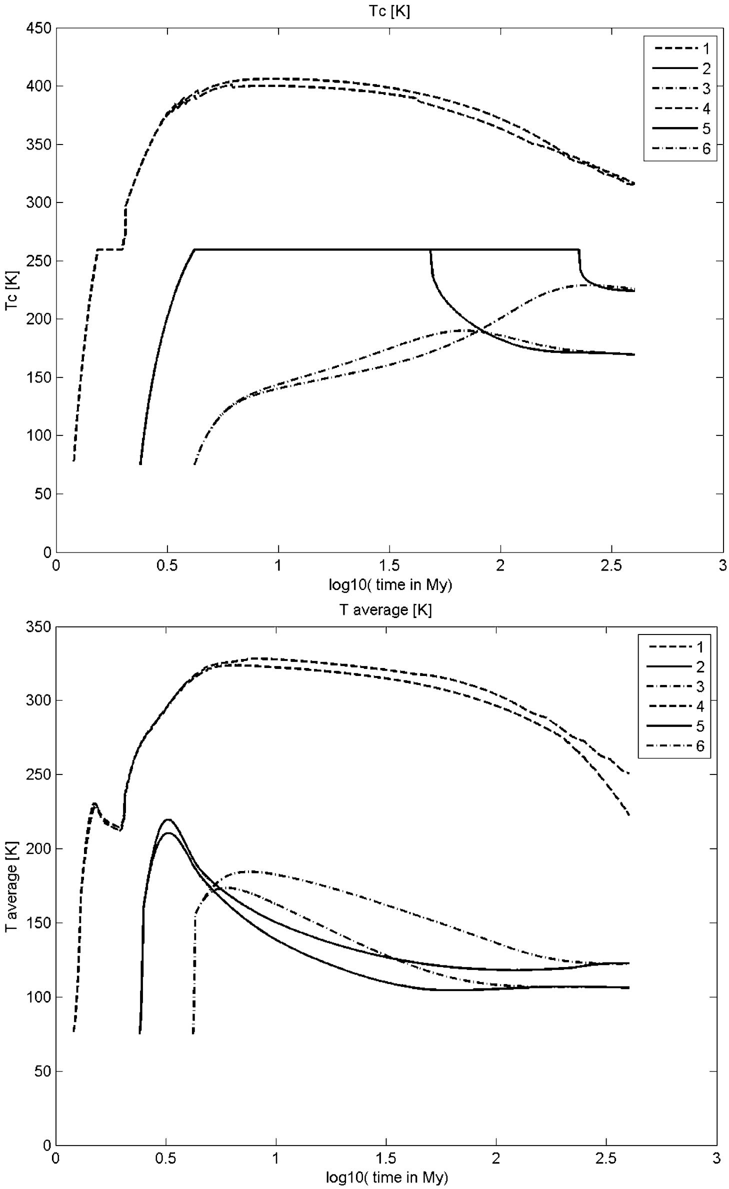


Fig. 6. The role of activation energy E for thermal evolution. Two values of activation energy are used: $E = 3\times10^4$ J mole^{-1} (thick lines) and $E = 5\times10^4$ J mole^{-1} (thin lines). The lower value of E increases the cooling effect of SSC in the solid part of the satellite. The rest of parameters have the "basic" values, *i.e.*, $t_{ac} = 0.1$ My, $\eta_0 = 10^{14}$ Pa s, $c_{serp} = 240\,000$ J kg^{-1}, $k_{sil} = 4.2$ W m^{-1} K^{-1}, $X_{NH3} = 0$. Lines from the left to the right correspond to $t_{ini} = 1.2$ My (dashed lines: 1 and 4), 2.4 My (solid lines: 2 and 5), 4.2 My (dashed-dotted lines: 3 and 6), respectively.

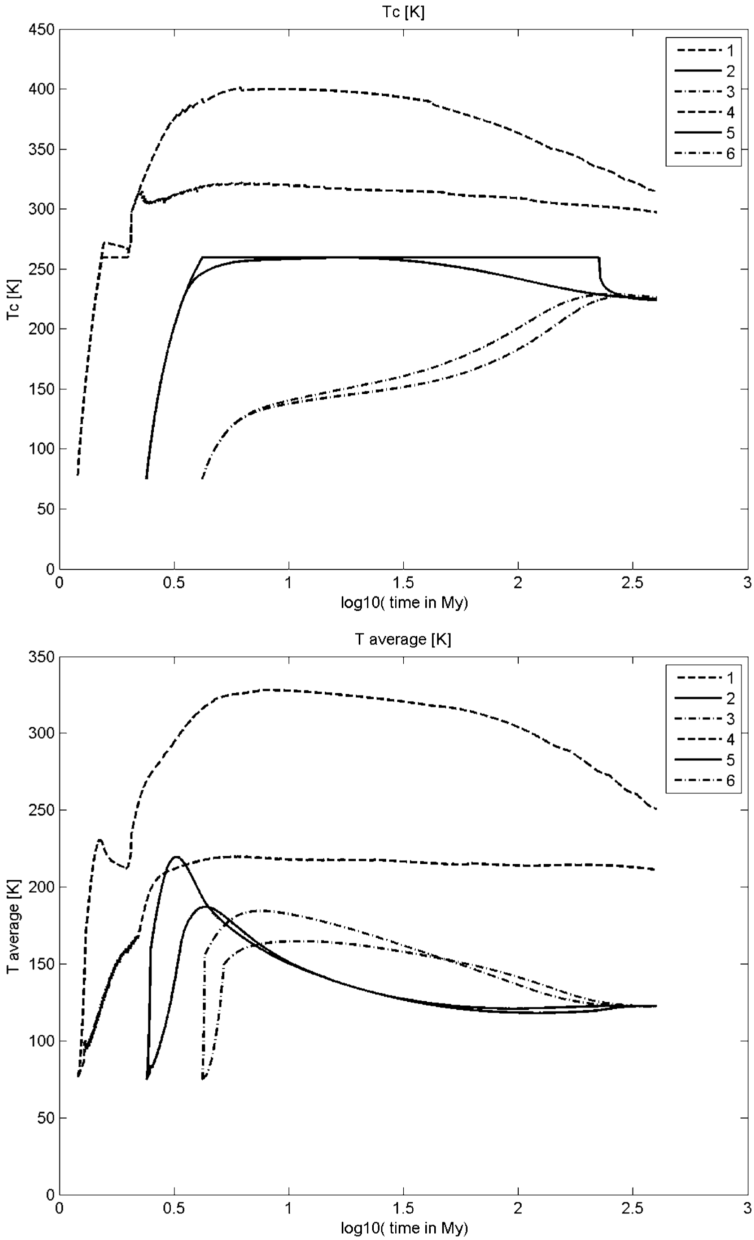


Fig. 7. The role of duration of accretion t_{ac} for thermal evolution. Two values of t_{ac} are used: $t_{ac} = 1$ My (thick lines) and $t_{ac} = 0.1$ My (thin lines). Lines from the left to the right correspond to $t_{ini} = 1.2$, 2.4, 4.2 My, respectively. The increased value of t_{ac} generally decreases temperature. The role of t_{ac} is similar to role of time of beginning of convection t_{ini}. The rest of parameters have their "basic" values.

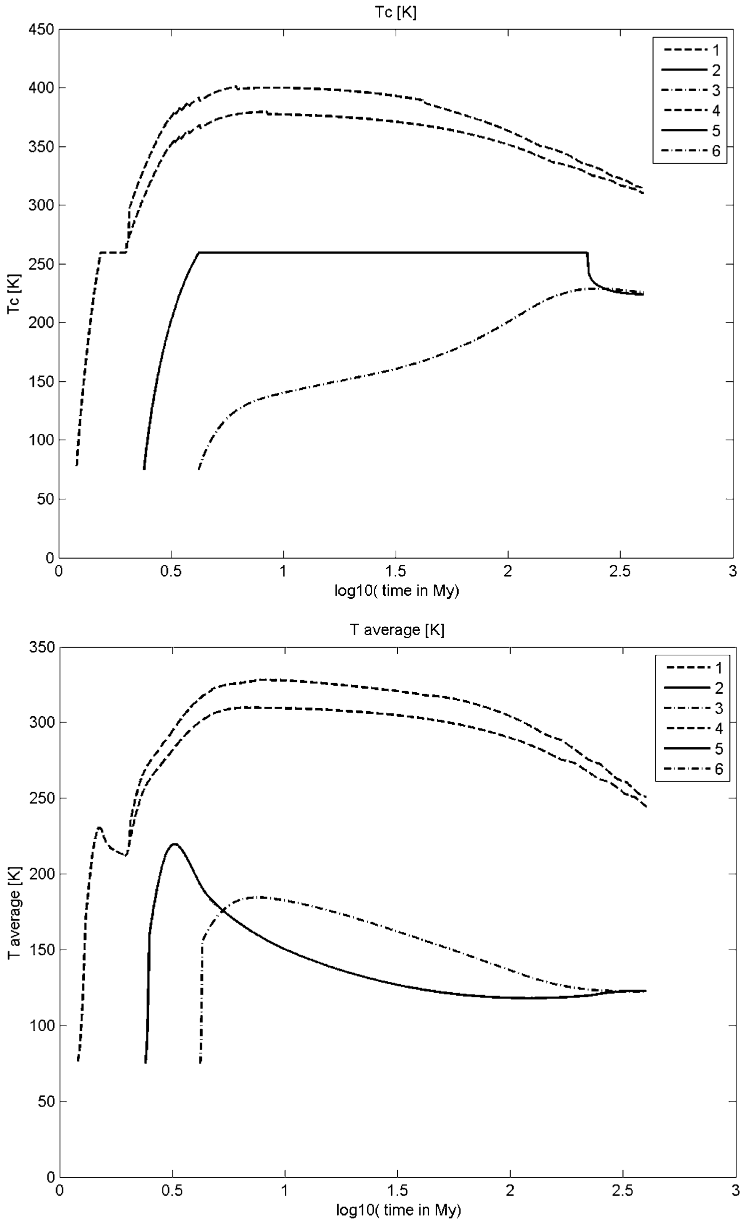


Fig. 8. The role of energy of serpentinization c_{serp} for thermal evolution. Two values of c_{serp} are used: $c_{serp} = 0$ J kg^{-1} (thick lines) and $c_{serp} = 240\,000$ J kg^{-1} (thin lines). Note that lines for $t_{ini} = 2.4$ and 4.2 My overlap indicating that melting and serpentinization do not occur. Other parameters have the "basic" values (see Fig. 4).

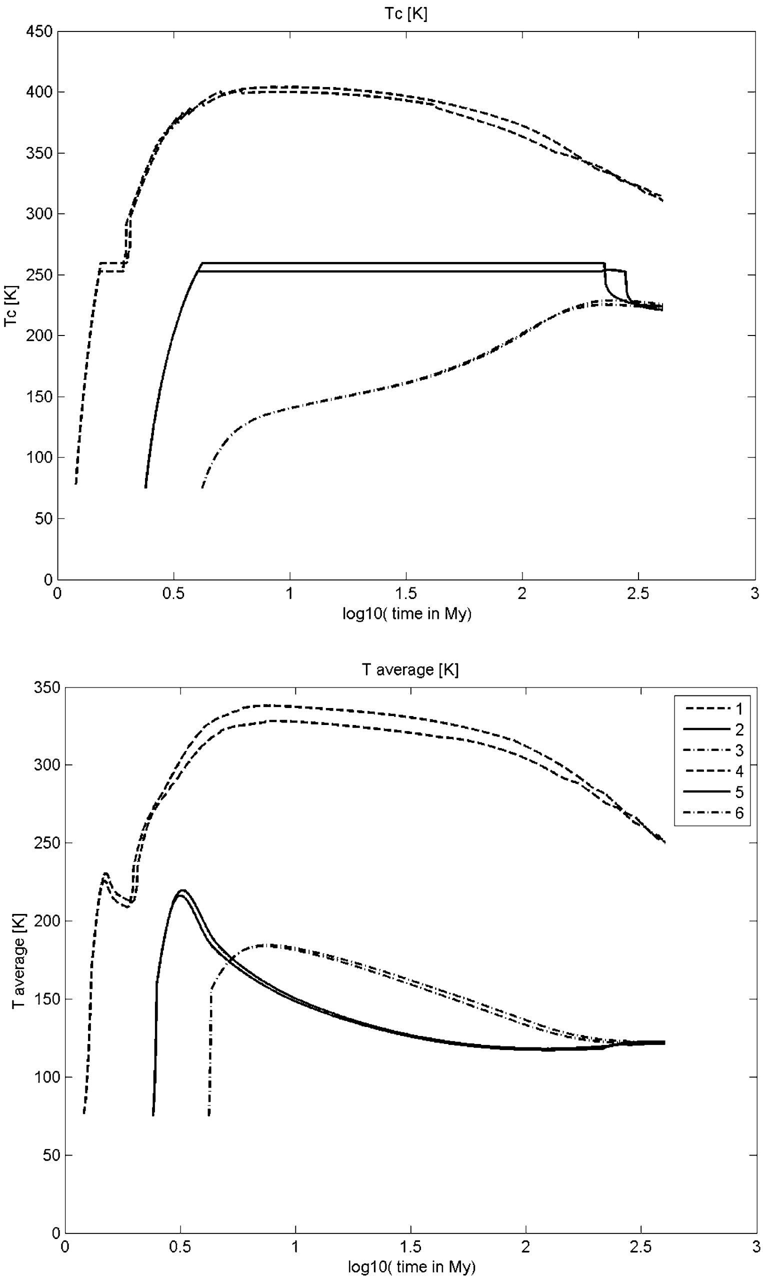


Fig. 9. The role of ammonia content X_{NH3} for thermal evolution. Two values of X_{NH3} are used: $X_{NH3} = 0.1$ (thick lines) and $X_{NH3} = 0$ (thin lines). Other parameters have the 'basic' values as in Fig. 4.

$t_{ini} = 1.2$ My and $X_{NH3} = 0.1$ for time 1.5-2 My, T_c and T_{ave} are below temperatures for the basic model. It is an effect of more intensive cooling by SSC. After 2 My, the melting starts reducing eventually the thickness of mantle and consequently reducing *Ra* for SSC. Lower *Ra* results in a lower cooling intensity and higher temperature. For $t_{ini} = 2.4$ My, when $T_c = T_m$, T_{ave} is slightly increased comparing to the basic model but still melting (and serpentinization) does not occur. For $t_{ini} = 4.2$ there is no visible effect of changed X_{NH3} at all.

5.3 Gravitational differentiation and LSC

Conclusions concerning gravitational differentiation and interaction with LSC are substantially changed in respect to Czechowski (2012) because he used only Stokes formula for the drag force. Here we use also more advanced formulae dependent on the Reynolds number. They are used for processes in Enceladus by Czechowski (2014). We use his approach to Rhea.

The melting allows motion of the silicate grains in the water. The grains' density is higher so negative buoyancy leads them to sink. However, we have also an opposite process, mixing by LSC. Let us consider interaction of these two processes.

The buoyancy force F_g acting on a spherical grain is:

$$F_g = (4/3)\,\pi\, r_g^{\,3}\,(\rho_g - \rho_w)\,g\,, \tag{18}$$

where r_g is the radius of grain, ρ_g is the density of grain, ρ_w is the density of water, and g is the gravity. The drag force F_D acts on the grain moving in fluid. F_D depends on the Reynolds number *Re*. *Re* is given by (*e.g.*, Landau and Lifshitz 1987, Malvern 1969, p. 465):

$$Re = \mathrm{v}\,\rho_w\,L\,/\,\eta_w\,, \tag{19}$$

where η_w is the viscosity of water (~10^{-3} Pa·s), L [m] is the largest dimension of the grain (we use here: $L = 2\,r_g$), v is the velocity of the grain [$\mathrm{m\cdot s^{-1}}$]. We considered two regimes of the flow: high *Re* flow when the drag force for the spherical grain is given by:

$$F_D = \tfrac{1}{2}\,C_D\,S_D\,\rho_w\,\mathrm{v}^2 = \tfrac{1}{2}\,0.45\,\pi\,r_g^{\,2}\,\rho_w\,\mathrm{v}^2 \tag{20}$$

and low *Re* flow when

$$F_D = 6\,\pi\,\eta\,r_g\,\mathrm{v}\,. \tag{21}$$

C_D is a drag coefficient (0.45 for a sphere), while S_D is the area of the cross section [$\mathrm{m^2}$]. Figure 10 presents results of the calculations. We do not know the size of typical grains. The structure of meteorite (*e.g.*, size of chondrules and CAI) suggests grains of the size of 10^{-3} m. Terminal velocity v_{term} of

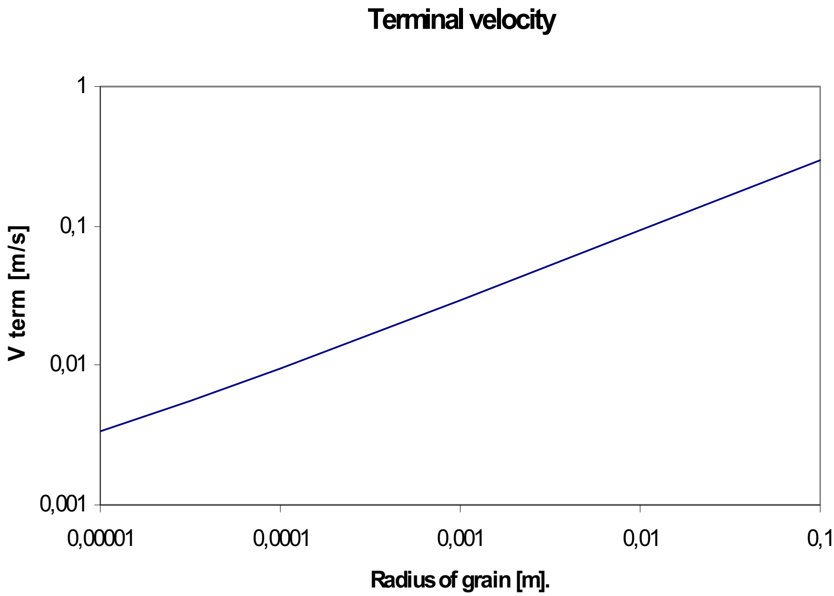


Fig. 10. Terminal velocity v_{term} for different values of radius of the grain r_g for: gravity $g = 0.1$ m s^{-2}, density difference 1500 kg m^{-3}, density of water 1000 kg m^{-3}, and the viscosity 10^{-3} Pa s. Nonlinear drag force given by Eq. 20 is used for high *Re* (>5), while the Stoke's formula Eq. 21 is used for low *Re* (<5). The values of terminal velocities presented here should be treated as the upper limits because spherical grains and large gravity are assumed.

such grains is a few cm s^{-1}, if gravity is 0.1 m s^{-2} (close to the center of satellite the gravity is lower than the surface value).

Consider now the interaction of the sinking grains and LSC. The average approximate velocity v_{LSC} of LSC in the molten interior of the satellite is given by the formula obtained from the theory of thermal boundary layer (*e.g.*, Turcotte and Schubert 2002, see also Czechowski 2014, for details):

$$v_{\mathrm{LSC}} = 0.354\,(\kappa/d)\,Ra^{1/2}\,, \tag{22}$$

where the Rayleigh number *Ra* is given by formula (8). The plot of v_{LSC} is shown in Fig. 11. Note that the velocity of convection v_{LSC} is higher than the terminal velocity v_{term} for considered range of $\boldsymbol{r_g}$ (compare Figs. 10 and 11). It means that differentiation is not possible because the mixing prevails. Czechowski (2014) indicates that even in smaller Enceladus, LSC convection is strong enough for mixing 1 mm grains and consequently not allowing core formation for more than 100 My, *i.e.*, even after decay of the short lived isotopes. It means that the time of core origin depends on the cooling rate of

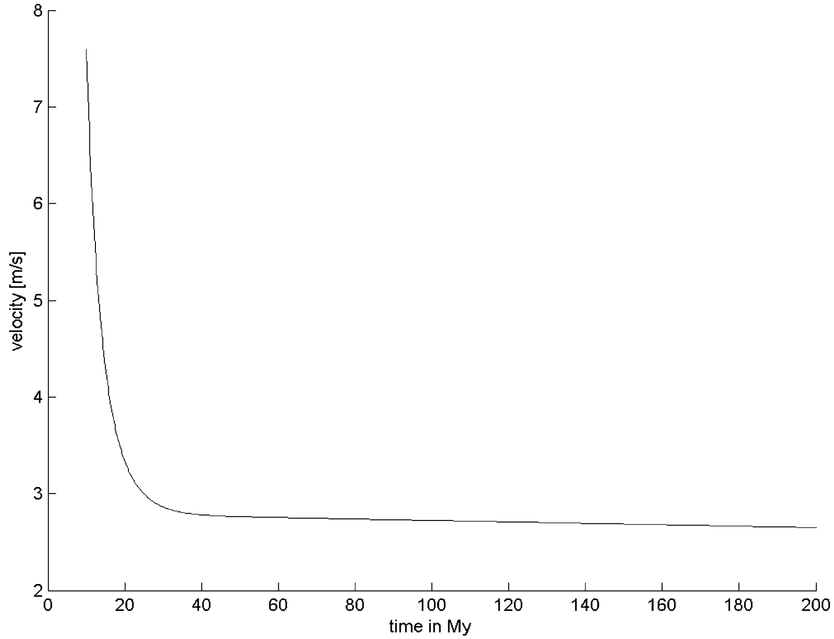


Fig. 11. Velocity of liquid state convection V_{LSC} according to formula 22 *versus* time for $g = 0.1$ m s^{-2}, and $d = 10^5$ m.

the satellite interior. For large Rhea, the bulk of the core is formed only ~200-500 My after CAI. Of course, even during intensive LSC, the largest grains (with high terminal velocity v_{term}) could sink forming a small "proto-core". In this way, LSC segregates the large silicate grains from the small ones.

Czechowski (2014) indicates also that the result of differentiation in Enceladus is a relatively cold core (temperature is close to the melting point of water) of loosely packed grains with water between them (see also Czechowski and Witek 2015). It is permeable for fluids because silicate rocks are permeable up to pressures of above 300 MPa (see discussion in Section 2.2 above). The same conclusion could be applied to Rhea. In fact, the potential temperature for Rhea is lower than for Enceladus (Fig. 3). The pressure in the centre of Rhea is still lower than 300 MPa (Fig. 1).

Temperature increase (above the boiling temperature) resulting from the tidal heating could be a mechanism that removes water from the core in Enceladus. In Rhea, probably there is/was no such mechanism. Eventually, water still could be present between grains and the density of core could be very low. This leads to the conclusion that the gravity data indicating low degree of differentiation could be interpreted also as a full differentiation but

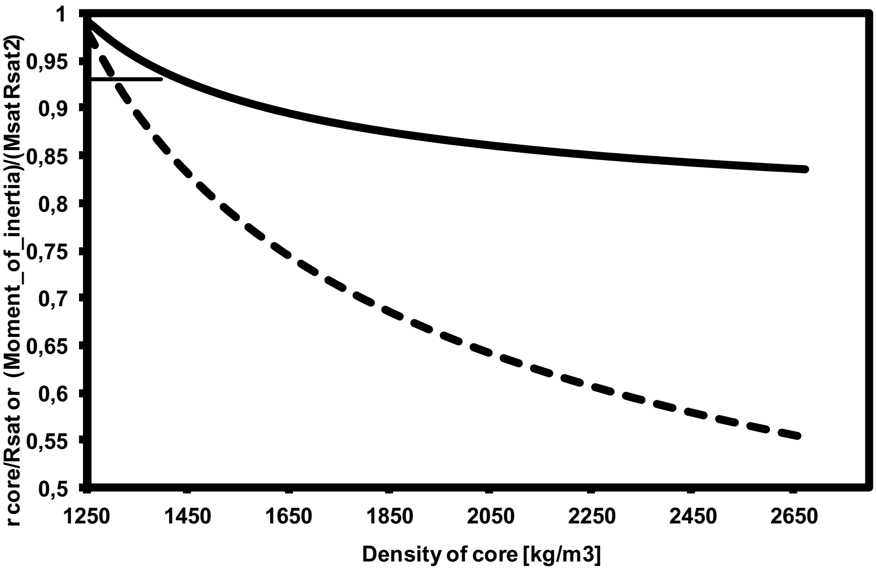


Fig. 12. Some parameters of Rhea as functions of the density of its core. The fully differentiated satellite is assumed for these calculations. Thick solid line – the ratio: (radius of core)/(R_{Sat}). Dashed line – the ratio: (moment of inertia)/($M_{Sat}R_{Sat}^2$). This ratio for an uniform satellite = 1. The short horizontal solid line shows the ratio (moment of inertia)/($M_{Sat}R_{Sat}^2$) corresponding to the value of this ratio obtained by Iess *et al.* (2007). It intersects the dashed line at ~1300 kg m^{-3}, so the full differentiation of Rhea is possible only if density of the core is ~1300 kg m^{-3}.

the density difference between the undifferentiated mantle and the differentiated porous core is very low (Fig. 12). Unfortunately, we do not know what content of inter-grain water is possible. It depends on the size and shape of the grains and properties of volatiles.

5.4 The role of Coriolis force

The role of Coriolis force for LSC in MIS was suggested already by Czechowski (2012). It could be an important factor determining direction of LSC. The Coriolis force is given by the formula:

$$\boldsymbol{F}_{\mathrm{Cor}} = 2\,\rho\,\boldsymbol{v}\times\boldsymbol{\omega}\,, \tag{23}$$

where $\boldsymbol{\omega}$ is the vector of angular velocity of rotation of Rhea. The ratio of Coriolis force $\boldsymbol{F}_{\mathrm{Cor}}$ to the viscous term in the Navier-Stokes equation is very large (>3×10^7) even for low spatial scale of the velocity changes, $s = 10^3$ m. The Coriolis force is latitude-dependent, so convection in the molten region and possible differentiation could also be latitude-dependent. Effects of Coriolis force give a possibility to explain some latitude-dependent features. Czechowski and Leliwa-Kopystynski (2013) use it to explain position of the Iapetus' equatorial ridge and bulge. An equatorial structure is found also on Rhea. According to Schenk *et al.* (2011): "equatorial configuration of Rhea's blue circle is surprisingly similar to that of Iapetus' equatorial ridge (Porco *et al.* 2005, Giese *et al.* 2008)". If the "blue circle" is of endogenic origin then the Coriolis force could be used for explanation of its position.

Note that the very large Prandtl number for SSC makes SSC independent of Coriolis force. However, SSC could be affected by some results of LSC (*e.g.*, by the latitude dependent differentiation). In this sense, SSC could be also latitude-dependent. Latitude-dependent tidal heating was probably never important in Rhea, so its role could be neglected.

5.5 Post differentiation scenarios

Total melting and total differentiation result in a stable situation: the icy mantle of low density overlies the silicate core. Note, however, that after partial differentiation there are 3 spherical regions: (i) the silicate core, (ii) the ice or water layer overlying it, and (iii) the primitive crust. This situation is gravitationally unstable because the density of ice or water is lower than the density of the undifferentiated crust above.

A few possible scenarios after differentiation were proposed – see Czechowski (2012) for more details. All of them lead to presently observed situations, *i.e.*, mostly undifferentiated satellite with the surface covered by depleted ice.

6. CONCLUSIONS AND FUTURE RESEARCH

- The properties of the CI and CM meteorites' matter can be used to estimate the properties of the silicate portion of icy satellites.
- We found that the time of beginning of accretion t_{ini} and the duration of accretion t_{ac} are crucial for the early evolution, especially for differentiation (see Figs. 4 and 7).
- Viscosity of ice close to the melting point η_0, activation energy in formula for viscosity E, and ammonia content X_{NH3} are very important for the evolution, but no dramatic differences are found if realistic values are considered (Figs. 5 and 6).
- The energy of serpentinization c_{serp} is important for the evolution, but its role is also not dominant.
- The LSC operating in the molten part could delay the differentiation and the core formation for a few hundred My.

- The gravity data could be interpreted as showing that Rhea is fully differentiated only if its core has high porosity and low density, ~1300 kg m^{-3}. The result of Anderson and Schubert (2007) allows for such a structure. In fact, there is no mechanism that could remove water from the molten core and the core of Rhea is probably porous. Note also that partial differentiation of a small body is possible only for very narrow time ranges. Moreover, the interpretation of Iess *et al.* (2007) should be reconsidered because of the results of Mackenzie *et al.* (2008) which indicates non-hydrostatic part of Rhea shape.
- LSC could lead to the generation of magnetic field by the planetary dynamo process (*e.g.*, Ivers and Phillips 2012).
- Our results suggest some extensions of the research. Similar models could be developed for other MIS if the gravity field data are provided.
- The role of Coriolis force for MIS should be investigated. It could be responsible for latitude-dependent features like the "blue circle" on the surface of Rhea.

A c k n o w l e d g m e n t s. This work was partially supported by the National Science Centre (grants 2011/01/B/ST10/06653 and 2013/08/S/ST10/00586). Computer resources of Interdisciplinary Centre for Mathematical and Computational Modeling of Warsaw University are also used in the research. We are very grateful to Dr. Maria Gritsevich and Dr. Daria Kuznetsova for their remarks and suggestions.

References

Abramov, O., and S.J. Mojzsis (2011), Abodes for life in carbonaceous asteroids?, *Icarus* **213**, 1, 273-279, DOI: 10.1016/j.icarus.2011.03.003.

Anderson, J.D., and G. Schubert (2007), Saturn's satellite Rhea is a homogeneous mix of rock and ice, *Geophys. Res. Lett.* **34**, 2, L02202, DOI: 10.1029/2006GL028100.

Barr, A.C., and R.M. Canup (2008), Constraints on gas giant satellite formation from the interior states of partially differentiated satellites, *Icarus* **198**, 1, 163-177, DOI:. 10.1016/j.icarus.2008.07.004.

Bland, P.A., M.D. Jackson, R.F. Coker, B.A. Cohen, B. Webber, M.R. Lee, C.M. Duffy, R.J. Chater, M.G. Ardakani, D.S. McPhail, D.W. McComb, and G.K. Benedix (2009), Why aqueous alteration in asteroids was isochemical: high porosity doesn't equal high permeability, *Earth Planet. Sci. Lett.* **287**, 3-4, 559-568, DOI: 10.1016/j.epsl.2009.09.004.

Brearley, A.J. (2004), Nebular versus parent-body processing. **In:** A.M. Davis, H.D. Holland, and K.K. Turekian (eds.), *Meteorites, Comets, and Planets. Treatise on Geochemistry*, Elsevier-Pergamon, Oxford, 247-268.

Brearley, A.J. (2006), The action of water. **In:** D.S. Lauretta, and H.Y. McSween Jr. (eds.), *Meteorites and the Early Solar System II*, University of Arizona Press, Tuscon, 587-624.

Brearley, A.J., and R.H. Jones (1998), Chondritic meteorites. **In:** J.J. Papike (ed.), *Planetary Materials*, Mineralogical Society of America, Reviews in Mineralogy, Vol. 36, 3-1–3-398.

Canup, R.M., and W.R. Ward (2009), Origin of Europa and the Galilean satellites. **In:** R.T. Pappalardo, W.B. McKinnon, and K. Khurana (eds.), *Europa*, University of Arizona Press in collaboration with Lunar and Planetary Institute, Tucson, 59-83.

Castillo-Rogez, J. (2006), Internal structure of Rhea, *J. Geophys Res*. **111**, E11, E11005, DOI: 10.1029/2004JE002379.

Castillo-Rogez, J., D. Matson, C. Sotin, T. Johnson, J. Lunine, and P. Thomas (2007), Iapetus' geophysics: Rotation rate, shape, and equatorial ridge, *Icarus* **190**, 1, 179-202, DOI: 10.1016/j.icarus.2007.02.018.

Christensen, U. (1984), Convection with pressure and temperature-dependent non-Newtonian rheology, *Geophys. J. Roy. Astr. Soc.* **77**, 2, 343-384, DOI: 10.1111/j.1365-246X.1984.tb01939.x.

Cogoni, M., B. D'Aguanno, L.N. Kuleshova, and D.W.M. Hofmann (2011), A powerful computational crystallography method to study ice polymorphism, *J. Chem. Phys*. **134**, 20, 204506, DOI: 10.1063/1.3593200.

Consolmagno, G.J., and D.T. Britt (1998), The density and porosity of meteorites from the Vatican collection, *Metorit. Planet. Sci.* **33**, 6, 1231-1241, DOI: 10.1111/j.1945-5100.1998.tb01308.x.

Coradini, A., G. Magni, and D. Turrini (2010), From gas to satellitesimals: Disk formation and evolution, *Space Sci. Rev*. **153**, 1, 411-429, DOI: 10.1007/s11214-009-9611-9.

Czechowski, L. (1993), Theoretical approach to mantle convection. **In:** R. Teisseyre L. Czechowski, and J. Leliwa-Kopystyński (eds.), *Dynamics of The Earth's Evolution*, Elsevier, Amsterdam, 161-271.

Czechowski, L. (2006), Parameterized model of convection driven by tidal and radiogenic heating, *Adv. Space Res*. **38**, 4, 788-793, DOI: 10.1016/j.asr.2005.12.013.

Czechowski, L. (2012), Thermal history and large scale differentiation of the Saturn's satellite Rhea, *Acta Geophys*. **60**, 4, 1192-1212, DOI: 10.2478/s11600-012-0041-9.

Czechowski, L. (2014), Some remarks on the early evolution of Enceladus, *Planet. Space Sci.* B. **104**, 185-199, DOI: 10.1016/j.pss.2014.09.010.

Czechowski, L., and K.J. Kossacki (2009), Thermal convection in the porous methane-soaked regolith of Titan: Investigation of stability, *Icarus* **202**, 2, 599-606, DOI: 10.1016/j.icarus.2009.02.032.

Czechowski, L., and K.J. Kossacki (2012), Thermal convection in the porous methane-soaked regolith in Titan: finite amplitude convection, *Icarus* **217**, 1, 130-143, DOI: 10.1016/j.icarus.2011.10.006.

Czechowski, L., and J. Leliwa-Kopystyński (2003), Tidal heating and convection in the medium size icy satellites, *Celest. Mech. Dyn. Astr.* **87**, 1, 157-169, DOI: 10.1023/A:1026136025400.

Czechowski, L., and J. Leliwa-Kopystynski (2013), Remarks on the Iapetus bulge and ridge, *Earth Planet Space* **65**, 8, 929-934, DOI: 10.5047/eps.2012.12.008.

Czechowski, L., and P.P. Witek (2015), Comparison of early evolutions of mimas and enceladus, *Acta Geophys*. **63**, 3, 900-921, DOI: 10.1515/acgeo-2015-0024.

Dalton, J.B., O. Prieto-Ballesteros, J. Kargel, C.S. Jamieson, J. Joliviet, and R. Quinn (2005), Spectral comparison of highly hydrated sulfate salts to disrupted terrains on Europa, *Icarus* **177**, 472-490.

Davaille, A., and C. Jaupart (1993), Transient high-Rayleigh-number thermal convection with large viscosity variations, *J. Fluid Mech*. **253**, 141-166, DOI: 10.1017/S0022112093001740.

Davies, G.F. (2007), Mantle regulation of core cooling: A geodynamo without core radioactivity? *Phys. Earth Planet. In.* **160**, 3-4, 215-229, DOI: 10.1016/j.pepi.2006.11.001.

De Pater, I., and J.J. Lissauer (2001), *Planetary Sciences*, Cambridge University Press, Cambridge, 528 pp.

Desch, S.J., J.C. Cook, T.C. Doggett, and S.B. Porter (2009), Thermal evolution of Kuiper belt objects, with implications for cryovolcanism, *Icarus* **202**, 2, 694-714, DOI: 10.1016/j.icarus.2009.03.009.

Dumoulin, C., M.P. Doin, and L. Fleitout (1999), Heat transport in stagnant lid convection with temperature- and pressure-dependent Newtonian or non-Newtonian rheology, *J. Geophys. Res.* **104**, B6, 12759-12777, DOI: 10.1029/1999JB900110.

Durham, W.B., S.H. Kirby, and L.A. Stern (1998), Rheology of planetary ices. **In:** B. Schmitt, C. de Bergh, and M. Festou (eds.), *Solar System Ices*, Kluwer Academic Publ., Dordrecht, 63-78.

Forni, O., A. Coradini, and C. Federico (1991), Convection and lithospheric strength in dione, an icy satellite of Saturn, *Icarus* **94**, 1, 232-245, DOI: 10.1016/0019-1035(91)90153-K.

Giese, B., T. Denk, G. Neukum, T. Roatsch, P. Helfenstein, P.C. Thomas, E.P. Turtle, A. McEwen, and C.C. Porco (2008), The topography of Iapetus' leading side, *Icarus* **193**, 2, 359-371, DOI: 10.1016/j.icarus.2007.06.005.

Goldsby, D.L., and D.L. Kohlstedt (1997). Grain boundary sliding in fine-grained Ice-I, *Scripta Mater.* **37**, 9, 1399-1405.

Gounelle, M., and M.E. Zolensky (2001), A terrestrial origin for sulphate veins in CI1 carbonaceous chondrites, *Meteorit. Planet. Sci.* **36**, 10, 1321-1329, DOI: 10.1111/j.1945-5100.2001.tb01827.x.

Grasset, O., and P.M. Parmentier (1998), Thermal convection in a volumetrically heated, infinite Prandtl number fluid with strongly temperature-dependent viscosity: Implications for planetary evolution, *J. Geophys. Res.* **103**, B8, 18171-18181, DOI: 10.1029/98JB01492.

Hobbs, P.V. (1974), *Ice Physics*, Oxford University Press, New York.

Hussmann, H., G. Choblet, D.L. Matson, C. Sotin, G. Tobie, and T. van Hoolst (2010), Implications of rotation, orbital states, energy sources, and heat transport for internal processes in icy satellites, *Space Sci. Rev.* **153**, 1-4, 317-348, DOI: 10.1007/s11214-010-9636-0.

Hutchison, R. (2004), *Meteorites: A Petrologic, Chemical and Isotopic Synthesis*, Cambridge University Press, Cambridge, 506 pp.

Iess, L., N.J. Rappaport, P. Tortora, J. Lunine, J.W. Armstrong, S.W. Asmar, L. Somenzi, and F. Zingoni (2007), Gravity field and interior of Rhea from Cassini data analysis, *Icarus* **190**, 2, 585-593, DOI: 10.1016/j.icarus.2007.03.027.

Ivers, D.J., and C.G. Phillips (2012), Anisotropic turbulent thermal diffusion and thermal convection in a rapidly rotating fluid sphere, *Physics Planet. In.* **190**, 1-9, DOI: 10.1016/j.pepi.2011.10.006.

Jaumann, R., R.N. Clark, F. Nimmo, A.R. Hendrix, B.J. Buratti, T. Denk, J.M. Moore, P.M. Schenk, S.J. Ostro, and R. Srama (2009), Icy satellites: Geological evolution and surface processes. **In:** M.K. Dougherty, L.W. Esposito, and S.M. Krimigis (eds.), *Saturn from Cassini-Huygens*, Springer, 637-681, DOI: 10.1007/978-1-4020-9217-6_20.

Jull, A.J.T., S. Cheng, J.L. Gooding, and M.A. Velbel (1988), Rapid growth of magnesium-carbonate weathering products in a stony meteorite from Antarctica, *Science* **242**,4877, 417-419, DOI: 10.1126/science.242.4877.417.

Kargel, J.S., and S. Pozio (1996), The volcanic and tectonic history of Enceladus, *Icarus* **119**, 2, 385-404, DOI: 10.1006/icar.1996.0026.

Kargel, J.S., J.Z. Kaye, J.W. Head, G.M. Marion, and R. Sassen, J.K. Crowley, O. Prieto-Ballesteros, S.A. Grant, and D.L. Hogenboom (2000), Europa's crust and ocean: origin, composition, and the prospects for life, *Icarus* **148**, 1, 226-265, DOI: 10.1006/icar.2000.6471.

Kuskov, O.L., and V.A. Kronrod (2005), Internal structure of Europa and Callisto, *Icarus* **177**, 2, 550-569, DOI: 10.1016/j.icarus.2005.04.014.

Landau, L.D., and E.M. Lifshitz (1987), *Fluid Mechanics*. *Vol. 6*, 2nd ed., Butterworth-Heinemann, 552 pp.

Leliwa-Kopystynski, J., M. Maruyama, and T. Nakajima (2002), The water-ammonia phase diagram up to 300 MPa: Application to icy satellites, *Icarus* **159**, 2, 518-528, DOI: 10.1006/icar.2002.6932.

Llana-Funez, S., K.H. Brodie, E.H. Rutter, and J.C. Arkwright (2007), Experimental dehydration kinetics of serpentinite using pore volumometry, *J. Metamorph. Geol.* **25**, 4, 423-428, DOI: 10.1111/j.1525-1314.2007.00703.x.

Losiak, A., and M.A. Velbel (2011), Evaporite formation during weathering of Antarctic meteorites – A weathering census analysis based on the ANSMET database, *Meteorit. Planet. Sci.* **46**, 3, 443-458, DOI: 10.1111/j.1945-5100.2010.01166.x.

Macke, R.J., G.J. Consolmagno, and D.T. Britt (2011), Density, porosity, and magnetic susceptibility of carbonaceous chondrites, *Meteorit. Planet. Sci.* **46**, 12, 1842-1862, DOI: 10.1111/j.1945-5100.2011.01298.x.

Mackenzie, R.A., L. Iess, P. Tortora, and N.J. Rappaport (2008), A non-hydrostatic Rhea, *Geophys. Res. Lett.* **35**, 5, L05204, DOI: 10.1029/2007GL032898.

Malamud, U., and D. Prialnik (2013), Modeling serpentinization: Applied to the early evolution of Enceladus and Mimas, *Icarus* **225**, 1, 763-774, DOI: 10.1016/j.icarus.2013.04.024.

Malvern, L.E. (1969), *Introduction to the Mechanics of a Continuous Medium*, Prentice-Hall Inc., Englewood Cliffs, 713 pp.

Mangold, N., P. Allemand, P. Duval, Y. Geraud, and P. Thomas (2002), Experimental and theoretical deformation of ice-rock mixtures: Implications on rheology and ice content of Martian permafrost, *Planet. Space Sci.* **50**, 4, 385-401, DOI: 10.1016/S0032-0633(02)00005-3.

Matson, D.L., J.C. Castillo-Rogez, G. Schubert, Ch. Sotin, and W.B. McKinnon (2009), The thermal evolution and internal structure of Saturn's mid-sized icy satellites. **In:** M.K. Dougherty, L.W. Esposito, and S.M. Krimigis (eds.), *Saturn from Cassini-Huygens*, Springer, 577-612, DOI 10.1007/978-1-4020-9217-6_18.

McKinnon, W.B. (1998), Geodynamics of icy satellites. **In:** B. Schmitt, C. de Bergh, and M. Festou (eds.), *Solar System Ices*, Kluwer Academic Publ., Dordrecht, 525-550.

McKinnon, W.B., and M.E. Zolensky (2003), Sulfate content of Europa's ocean and shell: Evolutionary considerations and some geological and astrobiological implications, *Astrobiology* **3**, 4, 879-897, DOI: 10.1089/153110703322736150.

Merk, E., D. Breuer, and T. Spohn (2002), Numerical modeling of ^{26}Al induced radioactive melting of asteroids concerning accretion, *Icarus* **159**, 1, 183-191, DOI: 10.1006/icar.2002.6872.

Mosqueira, I., P. Estrada, and D. Turrini (2010), Planetesimals and satellitesimals: Formation of the satellite systems, *Space Sci. Rev.* **153**, 1, 431-446, DOI: 10.1007/s11214-009-9614-6.

Multhaup, K., and T. Spohn (2007), Stagnant lid convection in the mid-sized icy satellite of Saturn, *Icarus* **186**, 2, 420-435, DOI: 10.1016/j.icarus.2006.09.001.

Opeil, C.P., G.J. Consolmagno, and D.T. Britt (2010), The thermal conductivity of meteorites: New measurements and analysis, *Icarus* **208**, 1, 449-454, DOI: 10.1016/j.icarus.2010.01.021.

Ostro, S.J., R.D. West, M.A. Janssen, R.D. Lorenz, H.A. Zebker, G.J. Black, J.I. Lunine, L.C. Wye, R.M. Lopes-Gautier, S.D. Wall, C. Elachi, L. Roth, S. Hensley, K. Kelleher, G.A. Hamilton, Y. Gim, Y.Z. Anderson, R.A. Boehmer, W.T.K. Johnson, and the Cassini RADAR Team (2006), Cassini RADAR observations of Enceladus, Thethys, Dione, Rhea, Iapetus, Hyperion, and Phoebe, *Icarus* **183**, 2, 479-490, DOI: 10.1016/j.icarus.2006.02.019.

Peacock, S.M. (2001), Are the lower planes of double seismic zones caused by serpentine dehydration in subducting oceanic mantle? *Geology* **29**, 4, 299-302, DOI: 10.1130/0091-7613(2001)029<0299:ATLPOD>2.0.CO;2.

Peltier, W.R., and G.T. Jarvis (1982), Whole mantle convection and the thermal evolution of the Earth, *Phys. Earth Planet. In*. **29**, 3-4, 281-304, DOI: 10.1016/0031-9201(82)90018-8.

Plescia, J.B. (1985), Geology of Rhea. **In:** *Lunar and Planetary Science Conference XVI*, 665-666.

Porco, C.C., E. Baker, J. Barbara, K. Beurle, A. Brahic, J.A. Burns, S. Charnoz, N. Cooper, D.D. Dawson, A.D. Del Genio, T. Denk, L. Dones, U. Dyudina, M.W. Evans, B. Giese, K. Grazier, P. Helfenstein, A.P. Ingersoll, R.A. Jacobson, T.V. Johnson, A. McEwen, C.D. Murray, G. Neukum, W.M. Owen, J. Perry, T. Roatsch, J. Spitale, S. Squyres, P.C. Thomas, M. Tiscareno, E. Turtle, A.R. Vasavada, J. Veverka, R. Wagner, and R. West (2005), Cassini imaging science: Initial results on Phoebe and Iapetus, *Science* **307**, 5713, 1237-1242, DOI: 10.1126/science.1107981.

Porco, C.C., P. Helfenstein, P.C. Thomas, A.P. Ingersoll, J. Wisdom , R. West, G. Neukum, T. Denk, R. Wagner, T. Roatsch, S. Kieffer, E. Turtle, A. McEwen, T.V. Johnson, J. Rathbun, J. Veverka, D. Wilson, J. Perry, J. Spitale, A. Brahic, J.A. Burns, A.D. Delgenio, L. Dones, C.D. Murray, and S. Squyres (2006), Cassini observes the active South Pole of Enceladus, *Science* **311**, 5766, 1393-1401, DOI: 10.1126/science.1123013.

Postberg, F., J. Schmidt, J.K. Hillier, S. Kempf, and R. Srama (2011), A salt-water reservoir as the source of a compositionally stratified plume on Enceladus, Nature **474**, 7353, 620-622, DOI: 10.1038/nature10175.

Prentice, A.J.R. (2006), Saturn's icy moon Rhea: A prediction for its bulk chemical composition and physical structure at the time of the Cassini spacecraft first flyby, *Publ. Astron. Soc. Aust*. **23**, 1, 1-11, DOI: 10.1071/AS05041.

Prialnik, D., and R. Merk (2008), Growth and evolution of small porous icy bodies with an adaptive-grid thermal evolution code. I. Application to Kuiper Belt objects and Enceladus, *Icarus* **197**, 1, 211-220, DOI: 10.1016/j.icarus.2008.03.024.

Prialnik, D., A. Bar-Nun, and M. Podolak (1987), Radiogenic heating of comets by Al^{26} and implications for their time of formation, *The Astrophys. J.* **319**, 993-1002.

Robuchon, G., G. Choblet, G. Tobie, O. Cadek, C. Sotin, and O. Grasset (2010), Coupling of thermal evolution and despinning of early Iapetus, *Icarus* **207**, 2, 959-971, DOI: 10.1016/j.icarus.2009.12.002.

Roscoe, R. (1952), The viscosity of suspensions of rigid spheres, *Brit. J. Appl. Phys.* **3**, 8, 267-269.

Rothery, D.A. (1992), *Satellites of the Outer Planets*, Clarendon Press, Oxford.

Rubin, A.E., J.M. Trigo-Rodrıguez, H. Huber, and J.T. Wasson (2007), Progressive alteration of CM carbonaceous chondrites, *Geochem. Cosmochim. Acta* **71**, 9, 2361-2382, DOI: 10.1016/j.gca.2007.02.008.

Rutter, E.H., S. Llana-Funez, and K.H. Brodie (2009), Dehydration and deformation of intact cylinders of serpentinite, *J. Struct. Geol.* **31**, 1, 29-43, DOI: 10.1016/j.jsg.2008.09.008.

Schenk, P., D.P. Hamilton, R.E. Johnson, W.B. McKinnon, C. Paranicas, J. Schmidt, and M.R. Showalter (2011), Plasma, plumes and rings: Saturn system dynamics as recorded in global color patterns on its midsize icy satellites, *Icarus* **211**, 1, 740-757, DOI: 10.1016/j.icarus.2010.08.016.

Schubert, G., T. Spohn, and R.T. Reynolds (1986), Thermal histories, compositions and internal structures of the moons of the solar system. **In:** J.A. Burns and M.S. Matthews (eds.), *Satellites*, The University of Arizona Press, Tucson, 224-292.

Schubert, G., D.L. Turcotte, and P. Olson (eds.) (2001), *Mantle Convection in the Earth and Planets*, Cambridge University Press, Cambridge, 940 pp., DOI: 10.1017/CBO9780511612879.

Schubert, G., J.D. Anderson, B.J. Travis, and J. Palguta (2007), Enceladus: Present internal structure and differentiation by early and long-term radiogenic heating, *Icarus* **188**, 2, 345-355, DOI: 10.1016/j.icarus.2006.12.012.

Schubert, G., H. Hussmann, V. Lainey, D.L. Matson, W.B. McKinnon, F. Sohl, C. Sotin, G. Tobie, D. Turrini, and T. Van Hoolst (2010), Evolution of icy satellites, *Space Sci. Rev.* **153**, 1, 447-484, DOI: 10.1007/s11214-010-9635-1.

Sharpe, H.N., and W.R. Peltier (1978), Parameterized mantle convection and the Earth's thermal history, *Geophys. Res. Lett.* **5**, 9, 737-740, DOI: 10.1029/GL005i009p00737.

Showman, A.P., and R. Malhotra (1999), The Galilean satellites, *Science* **286**, 5437, 77-84, DOI: 10.1126/science.286.5437.77.

Sinha, M., and R. Evans (2004), Mid-ocean ridges: Hydrothermal interactions between the lithosphere and oceans. **In:** C.R. German, J. Lin, L.M. Parson (eds.), *Mid-Ocean Ridges: Hydrothermal Interactions Between the Lithosphere and Oceans*, AGU Press, Washington, 19-62.

Sohl, F., T. Spohn, D. Breuer, and K. Nagel (2002), Implications from Galileo observations on the interior structure and chemistry of the Galilean satellites, *Icarus* **157**, 1, 104-119, DOI: 10.1006/icar.2002.6828.

Sohl, F., M. Choukroun, J. Kargel, J. Kimura, R. Pappalardo, S. Vance, and M. Zolotov (2010), Subsurface water oceans on icy satellites: Chemical composition and exchange processes, *Space Sci. Rev.* **153**, 1-4, 485-510, DOI: 10.1007/s11214-010-9646-y.

Solomatov, V.S. (1995), Scaling of temperature- and stress-dependent viscosity convection, *Phys. Fluids* **7**, 2, 266-274, DOI: 10.1063/1.868624.

Stein, C., S. Stein, and A. Pelayo (1995), Heat flow and hydrothermal circulation. **In:** S.E. Humphris, R.A. Zierenberg, L.S. Mullineaux, and R.E. Thomson (eds.), *Seafloor Hydrothermal Systems*, American Geophysical Union, Washington, D.C., 425-445, DOI: 10.1029/GM091p0425.

Tomeoka, K., and P.R. Buseck (1988), Matrix mineralogy of the Orgueil CI carbonaceous chondrite, *Geochim. Cosmochim. Ac.* **52**, 6, 1627-1640, DOI: 10.1016/0016-7037(88)90231-1.

Travis, B.J., J. Palguta, and G. Schhubert (2012), A whole-moon thermal history model of Europa: Impact of hydrothermal circulation and salt transport, *Icarus* **218**, 2, 1006-1019, DOI: 10.1016/j.icarus.2012.02.008.

Turcotte, D.L., and G. Schubert (eds.) (2002). *Geodynamics*, 2nd ed., Cambridge University Press, Cambridge, 450 pp.

Vance, S., J. Harnmeijer, J. Kimura, H. Hussmann, B. DeMartin, and J.M. Brown (2007), Hydrothermal systems in small ocean planets, *Astrobiology* **7**, 6, 987-1005, DOI: 10.1089/ast.2007.0075.

Velbel, M.A., D.T. Long, and J.L. Gooding (1991), Terrestrial weathering of Antarctic stone meteorites: Formation of Mg-carbonates on ordinary chondrites, *Geochim. Cosmochim. Ac.* **55**, 1, 67-76, DOI: 10.1016/0016-7037(91)90400-Y.

Velbel, M.A., E.K. Tonui, and M.E. Zolensky (2012), Replacement of olivine by serpentine in the carbonaceous chondrite Nogoya (CM2), *Geochim. Cosmochim. Ac.* **87**, 117-135, DOI: 10.1016/j.gca.2012.03.016.

Walsh, J.B., and W.E. Brice (1984), The effect of pressure on porosity and the transport properties of rock, *J. Geophys. Res.* **89**, NB11, 9425-9432, DOI: 10.1029/JB089iB11p09425.

Weisberg, M.K., T.J. McCoy, and A.N. Krot (2006), Systematics and evaluation of meteorite classification. **In:** D.S. Lauretta and H.Y. McSween Jr. (eds.), *Meteorites and the Early Solar System II*, University of Arizona Press, Tuscon, 19-52.

Young, E.D., K.K. Zhang, and G. Schubert (2003), Conditions for pore water convection within carbonaceous chondrite parent bodies — implications for planetesimal size and heat production, *Earth Planet. Sci. Lett*. **213**, 3-4, 249-259, DOI: 10.1016/S0012-821X(03)00345-5.

Zolotov, M.Y. (2007), An oceanic composition on early and today's Enceladus, *Geophys. Res. Lett.* **34**, 23, L23203, DOI: 10.1029/2007GL031234.

Zolotov, M.Y., and M.V. Mironenko (2007), Hydrogen chloride as a source of acid fluids in parent bodies of chondrites. **In:** *38th Lunar and Planetary Science Conference*, Abstract #2340.

Zolotov, M.Y., and E.L. Shock (2001), Composition and stability of salts on the surface of Europa and their oceanic origin, *J. Geophys. Res*. **106**, E12, 32815-32827, DOI: 10.1029/2000JE001413.

Received 2 July 2015
Received in revised form 17 November 2015
Accepted 22 January 2016

DE GRUYTER OPEN

Acta Geophysica

vol. 64, no. 6, Dec. 2016, pp. 2717-2733
DOI: 10.1515/acgeo-2016-0087

Dynamic and Thermal Processes in the Mid-Latitude Ionosphere over Kharkov, Ukraine (49.6° N, 36.3° E), During the 13-15 November 2012 Magnetic Storm: Calculation Results

Mykhaylo V. LYASHENKO

Institute of Ionosphere, Kharkov, Ukraine
e-mail: mlyashenko@ya.ru

Abstract

Calculation results of the variations of dynamic and thermal process parameters in geospace plasma during the 13-15 November 2012, magnetic storm (MS) over Kharkov are presented. The calculations were based on experimental data obtained on the Kharkov incoherent scatter radar, single in the European mid-latitudes. Calculations showed that during the MS there took place an increase, by modulus, of the values of vertical component of transfer velocity, due to ambipolar diffusion, up to a factor of 1.4-2.1. During the MS there took place a decrease of the values of energy input to the electron gas by about 20-35%. During the main phase of MS, the heat flux density transferred by electrons increased up to a factor of 2-2.5. Results of estimates of the zonal component electric field value E_y are presented. During the MS the value of E_y was –9.5 mV/m. The vertical component of plasma velocity due to electromagnetic drift v_{EB} has been calculated.

Key words: magnetic storm, dynamic and thermal processes, geospace.

Ownership: Institute of Geophysics, Polish Academy of Sciences;

1. INTRODUCTION

Dynamic and thermal processes play an important role in the formation of spatial structure of the ionospheric plasma at altitudes of the F2 region of the ionosphere. Parameters of the physical processes have characteristic variations depending on the season, time of the day, place of observation, as well as by space weather conditions. As shown by experimental studies, the effects of strong geospace disturbances (geospace storms) are well manifested in variations of the dynamic and thermal processes in the ionosphere. Strong magnetic storms lead to significant changes in the dynamic and thermal modes of the ionospheric plasma. Such effects have been recorded not only in the high latitudes, but in the middle and low latitudes too.

For investigation of the effects of the strong geospace storms, various tools of remote and local sounding of the near-Earth space have been used (ionosondes, worldwide network of incoherent scatter radars (ISRs), satellites, *etc*.).

Incoherent scatter radars are the most informative radiophysical tools for ionospheric plasma investigation. At present, about fifteen incoherent scatter radars operate over the word. The results obtained by a worldwide network of ISRs allow extending our knowledge about global processes in the ionosphere during strong geomagnetic disturbances in polar, middle and low latitudes (Buonsanto 1999, Buonsanto *et al*. 1999a, b; Goncharenko *et al*. 2005, Grigorenko *et al*. 2005a, b, 2007; Immel *et al*. 2015). Results of observations of the magnetic storm effects in ionospheric plasma variations obtained on ISRs have been described in numerous publications (Buonsanto 1999, Buonsanto *et al*. 1999a, b, Goncharenko *et al*. 2005, Grigorenko *et al*. 2005a, b, 2007; Immel *et al*. 2015).

In the middle latitudes of Europe, the incoherent scatter radar (ISR) in Kharkov, Ukraine (49.6° N, 36.3° E), is the only means for radio sounding of the ionosphere, providing almost entire set of basic parameters of the ionospheric plasma (Emelyanov and Zhivolup 2013). The experimental data obtained at the Kharkov incoherent scatter radar were used earlier for the study of unique events in the geospace − partial solar eclipses (Lyashenko 2013, Lyashenko and Chernogor 2013, Domnin *et al*. 2014a), geospace storms of different intensity (Grigorenko *et al*. 2007, Chernogor *et al*. 2007), as well as the wave processes in the ionospheric plasma (Burmaka and Chernogor 2012).

The aim of this study was the calculation of parameters of the dynamic and thermal process in the geospace plasma during the 13-15 November 2012 magnetic storm.

2. GENERAL INFORMATION ABOUT THE 13-15 NOVEMBER 2012 GEOMAGNETIC STORM

The magnetic storm began on 13 November at 15:00 UT. The main phase of the magnetic storm took place from 18:00 UT on 13 November to 06:00 UT on 14 November. The extreme values of the geomagnetic activity indices during the magnetic storm were: $AE_{\max} = 1009$ nT, $K_{p\,\max} = 6+$, $D_{st} = -108$ nT. The value of the IMF B_z-component was –(17-18) nT. The value of the Akasofu function was ~26-30 GJ/s. The solar radio flux index $F_{10.7}$ ranged from 141 to 146.

For comparison of the magnetic storm effects in the variations of dynamic and thermal processes, the reference period from 21 to 23 November was chosen, which was characterized by quiet heliogeophysical conditions. Geomagnetic and solar activity indexes for this period were $A_p = 2$-7, $K_p = 0$-3, $F_{10,7} = 126$-140.

3. THE OBSERVATION MEANS AND EXPERIMENTAL RESULTS

For modeling the parameters of dynamic and thermal processes, the Kharkov ISR (geographic coordinates: 49.6° N, 36.3° E; geomagnetic coordinates: 45.7°, 117.8°) data were used. At present, the Kharkov ISR is the only reliable and most informative data source of the geospace plasma state in the mid-latitudes of Central Europe. Radar allows measuring with high accuracy (usually, the error is of 1-10%) and acceptable altitude resolution (10-100 km) the following ionospheric parameters: electron density N, electron T_e and ion T_i temperatures, a vertical component of the plasma transfer velocity v_z, and ion composition. The investigated altitude range is 100-1500 km (depending on the solar activity level).

Domnin *et al.* (2014b) presents observation results of ionospheric plasma parameter variations during 13-15 November 2012 magnetic storm obtained on Kharkov incoherent scatter radar.

The magnetic storm was accompanied by ionospheric storm with two positive phases (δfoF2 > 0) and one negative phase (δfoF2 < 0).

The electron density in the F2-region maximum of the ionosphere, N_mF2, during the first positive phase of the ionospheric storm increased up to a factor of 3. Then there was a significant decrease in the N_mF2 up to a factor of 5 (negative phase of ionospheric storm). During the second positive phase of the ionospheric storm, the electron density N_mF2 increased up to a factor of 2.8. The greatest increase of F2-peak height h_mF2 (up to 400 km) took place at 03:30 UT, while during quiet conditions h_mF2 have not exceed 275 km.

During magnetic storm, the electron density N in the altitude range of 200-250 km decreased up to a factor 5. At altitudes of 300 and 350 km, the reduction of N was up to a factor of 3.5 and 3. The electron temperature in-

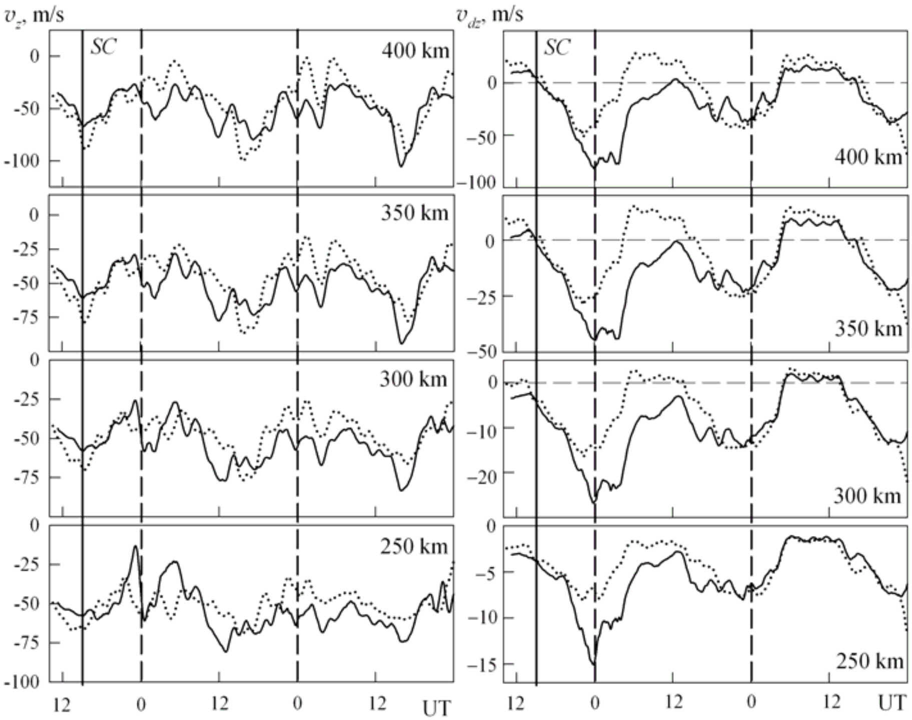


Fig. 1. Temporal variations of the vertical component of the plasma transfer velocity v_z (experimental results) and the vertical component of the plasma transfer velocity due to ambipolar diffusion v_{dz} (calculation) during 13-15 November 2012 magnetic storm (solid line) and quiet condition period on 21-23 November 2012 (dots).

creased up to 200-300 K in the altitude range of 200-700 km. The ion temperature for the considered altitudes increased up to 100-200 K.

The variations of the vertical component of plasma transfer velocity v_z during magnetic storm were small (see Fig. 1, left panel). In the main phase of magnetic storm (from 23:50 UT on 13 November to 16:30 UT on 14 November) the quasi-harmonic oscillations with a period from 3 h 30 min to 5 h 15 min and amplitude of 12-40 m/s were detected.

In greater detail, the results of the 13-15 November 2012, magnetic storm effects in variations of the main parameters of the ionospheric plasma have been described by Domnin *et al*. (2014b).

4. INITIAL THEORETICAL RELATIONS

To calculate the dynamic and thermal process parameters in geospace plasma, the following theoretical relations were used.

Plasma flux densities. Expressions for calculation of the plasma flux density in the vertical direction and the particle flux due to ambipolar diffusion are of the form:

$$\Pi_p = v_z N$$
$$\Pi_d = v_{dz} N ,$$

where v_z is the vertical component of the plasma velocity (Kharkov ISR data). Expression for calculating the plasma transfer velocity due to ambipolar diffusion v_{dz} has the form (Schunk and Nagy 2000)

$$v_{dz} = -D_a \sin^2 I \left(\frac{1}{H_p} + \frac{1}{N}\frac{\partial N}{\partial z} + \frac{1}{T_p}\frac{\partial T_p}{\partial z} \right),$$

where $D_a = (kT_p)/(m_i \nu_{in})$ is the longitudinal component of the ambipolar diffusion tensor, k the Boltzmann constant, $T_p = T_e + T_i$ the plasma temperature, m_i the O^+ ion mass, I the Earth's geomagnetic field inclination (for the Kharkov ISR coordinates $I = 66.85°$), $H_p = (kT_p)/(m_i g)$ the plasma scale height, g the gravity acceleration, and ν_{in} is the total collision frequency of ion with neutrals. Expression for ν_{in} has the form

$$\nu_{in} = \nu_{O^+,O} + \nu_{O^+,O_2} + \nu_{O^+,N_2} + \nu_{O^+,H} + \nu_{O^+,He} , \tag{1}$$

where $\nu_{O^+,O}$, ν_{O^+,O_2}, ν_{O^+,N_2}, $\nu_{O^+,H}$, $\nu_{O^+,He}$ are the collision frequencies of oxygen ions in the parent gas, with atoms and molecules of oxygen, nitrogen, hydrogen and helium, respectively. Each of the summands in Eq. 1 can be calculated from the following relations (Stubbe 1968):

$$\nu_{O^+,O} = 1.86 \cdot 10^{-9} N(O) \left(\frac{T_i + T_n}{2000} \right)^{0.37} ,$$

$$\nu_{O^+,O_2} = N(O_2) \cdot 10^{-9} ,$$

$$\nu_{O^+,N_2} = 1.08 \cdot 10^{-9} N(N_2) ,$$

$$\nu_{O^+,H} = 2.19 \cdot 10^{-9} N(H) ,$$

$$\nu_{O^+,He} = 0.88 \cdot 10^{-9} N(He) ,$$

where T_n, $N(O)$, $N(O_2)$, $N(N_2)$, $N(H)$, and $N(He)$ are the temperature of neutrals, concentrations of atomic oxygen, molecular oxygen and nitrogen, atomic hydrogen and helium calculated from NRLMSISE-00 model (Picone *et al*. 2002).

Zonal component of electric field. It is well known that in quiet conditions the contribution of magnetospheric sources in electric fields and currents in the middle and low latitudes is small. As follows from experimental studies and theoretical calculations (Blanc *et al*. 1977, Blanc and Amayenc 1979, Ogawa *et al*. 1975, Richmond *et al*. 1980), the magnitude of the electric field in the mid-latitude ionosphere without geomagnetic disturbances does not exceed several mV/m. At altitudes of the ionospheric F2-peak the plasma drift caused by these fields is small compared to the transport processes of charged particles due to ambipolar diffusion and neutral winds. During strong geomagnetic disturbances, the electric fields have penetration at the altitudes of the mid-latitude ionosphere, increasing the plasma velocity in the electric and magnetic fields they cross. It should be noted that the transfer of the plasma due to the electromagnetic drift during geomagnetic storms has a significant impact on the spatial distribution of parameters of the mid-latitude ionosphere.

As it is known, without perturbation the electric field effects in the mid-latitudes can be neglected. But during strong geomagnetic storms there is an amplification of electric fields due to the magnetospheric convection, which affects the dynamics of the mid-latitude ionosphere. During disturbed conditions in the mid-latitudes, and neglecting the effects of geomagnetic field declination, the main contribution to the vertical transport of plasma comes from the zonal electric field and neutral winds. Electric field directed to the east causes the plasma drift upwards and the field directed to the west causes the transfer of ionospheric plasma down. We estimate the value of the zonal component of the electric field, as well as the contribution of the vertical component of plasma motion due to the electromagnetic drift in the dynamic mode of the ionosphere during magnetic storm on 13-15 November 2012.

As shown in the calculations presented by Sergeenko (1982), there is a correlation between the *AE* index and the values of the zonal component of the electric field E_y. In this case, the electric fields of magnetospheric origin and geomagnetic plasma heating are the main sources of dynamic processes in the mid-latitude ionosphere during strong geomagnetic disturbances.

To calculate E_y we use the empirical relation between the magnitude of the electric field and auroral activity index, given by Sergeenko (1982):

$$E_y = \left(0.55 - 0.01AE\right)\cdot 10^{-3} ,$$

where *AE* is the auroral activity index (in nT).

Velocity of plasma transport due to electromagnetic drift. Expression to calculate the velocity of plasma transport due to electromagnetic drift has the form (Schunk and Nagy 2000)

$$v_{EB} = \left(E_x/B\right)\cos I \sin D + \left(E_y/B\right)\cos I \cos D ,$$

where B is the geomagnetic field modulus ($B \approx 4.4 \cdot 10^{-5}$ nT at $z = 300$ km (Finlay *et al*. 2010)), E_x and E_y are the meridional (positive direction to equator) and zonal (positive direction to east) components of electric field, and D is the declination of the magnetic field. Neglecting the effects of declination (for Kharkov city $D = 8.14°$), the expression for the calculation of the vertical component of the velocity of plasma transport by the electromagnetic drift has the form

$$v_{EB} = \left(E_y/B\right)\cos I .$$

Neutral wind in the ionosphere. In mid-latitudes, the vertical component of the velocity of ion wind drag is determined by the meridional component of the velocity of the neutral gas horizontal motion. Neutral wind directed toward the equator moves the plasma upward along the magnetic field lines, and poleward wind moves the plasma downward.

Expression for calculating the meridional component of the neutral wind velocity v_{nx}, upon neglecting the effects of the geomagnetic field declination, has the form (Schunk and Nagy 2000):

$$v_{nx} = \left(v_z - v_{dz} - v_{EB}\right)/\left(\sin I \cos I\right) .$$

Energy influx to the electron gas. In the ionospheric F region, the collision frequency of electrons with neutrals is smaller than electrons with ions. In this case, the main mechanisms of electron gas cooling are the heat loss in collisions of electrons and ions, excitation of the fine structure of the oxygen atoms and electron gas heat conductivity (Schunk and Nagy 2000). There is also the photoelectron transfer and the non-local heating of the electron gas, related to this transfer.

At altitudes $z \leq 350$ km, the heat conductivity of the electron gas can be neglected and the electron energy balance equation in the stationary case in the SI system has the form (Schunk and Nagy 2000)

$$Q = L_{ei} + L_e ,$$

$$L_{ei} = 8 \cdot 10^{-32} N^2 (T_e - T_i) T_e^{-3/2} ,$$

$$L_e = 6.4 \cdot 10^{-37} N N(\mathrm{O}) \left(T_e - T_i\right) T_n^{-1} ,$$

where Q is the energy transferred to the thermal electrons at the Coulomb collisions with photoelectrons, L_{ei} – the energy lost in electron-ion collisions, L_e – the energy needed to excite the fine structure of the oxygen atoms, N – the electron density (Kharkov ISR data), N(O) – the density of oxygen atoms, T_e and T_i – the electron and ion temperatures (Kharkov ISR data). The neutral temperature T_n and N(O) density has been calculated by the NRLMSISE-00 model (Picone *et al*. 2002).

Heat flux density. The heat balance of the electron gas depends on the heat flux carried by electrons from the plasmasphere into the ionosphere. The heat in the plasmasphere accumulates by suprathermal electrons escaping from the place of their formation in the topside ionosphere. Some of the electrons lose their energy in Coulomb collisions with thermal electrons and ions. Another part of the electrons fall into the magnetic flux tube. In the magnetic tube, trapped electrons become thermalized in multiple reflections from the ends of the tube. Thus, accumulated heat in the plasmasphere fed back into the ionosphere by means of the heat conductivity of the electron gas (Schunk and Nagy 2000).

Heat flux can be defined by the kinetic equation with the transport of suprathermal electrons. The expression to calculate the heat flux density from the plasmasphere in the vertical direction has the form (Schunk and Nagy 2000):

$$\Pi_T = -\kappa_e \sin^2 I \frac{\partial T_e}{\partial z} , \tag{2}$$

where $\kappa_e = (2.08 \cdot k^2 N T_e)/(m \nu_{ei})$ is the longitudinal component of the heat conductivity tensor of the electron gas, and m is the mass of the electron.

The collision frequency between electrons and O^+ ions to calculate the longitudinal component of the conductivity tensor in Eq. 2 can be found by using the expression (Lyashenko 2013, Lyashenko and Chernogor 2013, Domnin *et al*. 2014a):

$$\nu_{ei} \approx 5.5 \cdot 10^{-6} N T_e^{-3/2} \ln\left(2.2 \cdot 10^4 T_e N^{-1/3}\right) .$$

5. CALCULATION RESULTS

Particle transport velocity due to ambipolar diffusion. Temporal variations of the vertical component of the plasma transport velocity due to ambipolar diffusion v_{dz} at altitudes of 250-400 km during the 13-15 November 2012 magnetic storm is presented in Fig. 1 (right panel).

The calculations show that in quiet conditions (22 November 2012) at night there was a downward plasma transport ($v_{dz} < 0$). The velocity v_{dz} was about –7, –15, –25, and –50 m/s at altitudes of 250, 300, 350, and 400 km, respectively. In the daytime, the transfer plasma velocity due to ambipolar diffusion was negligible at altitudes of 250 and 300 km (no more than 2-3 m/s) and at altitudes of 350 and 400 km v_{dz} it was 15-25 m/s.

After the beginning of the magnetic storm (around 18:00 UT on 13 November), the behaviour of v_{dz} changed. On the night of 14 November at altitudes of 250, 300, 350, and 400 km, v_{dz} reached values of –15, –25, –40, and –70 m/s, respectively. During the daytime on 14 November, v_{dz} also differed from the variations in quiet conditions.

Plasma fluxes. Figures 2 and 3 show temporal variations of the density of the full plasma flux Π_p and the density of the plasma flux due to ambipolar diffusion Π_d in the altitude range of 250-400 km during 13-15 November 2012, magnetic storm and quiet condition period on 21-23 November 2012.

The full plasma flux Π_p is the total flux of charged particles caused the transfer of plasma due to ambipolar diffusion Π_d, neutral wind v_{nx} and plasma drift v_{EB} in crossed electric and magnetic fields. It should be noted that during the strong magnetic storms the contribution of neutral winds and electromagnetic drift in the plasma transfer is significant and comparable to the contribution of the diffusion flux.

The variations of full plasma flux density Π_p during the magnetic storm were as follows. After the magnetic storm beginning (November 13), there was a slight increase of Π_p by modulus compared to the reference day. On 14 November near midday (during the recovery phase of magnetic storm) there was a decrease in plasma flux density $|\Pi_p|$ by 55 and 25% at altitudes of 250 and 300 km.

On 15 November, diurnal variation of the plasma flux density began to recover.

The variations of the charged particle flux density due to ambipolar diffusion Π_d (Fig. 3) during the 13-15 November magnetic storm were complicated.

Temporal variations of the Π_d flux on 22 November kept their basic features of diurnal behaviour: upward flux after sunrise and during daytime and downward flux after sunset during nighttime. It can be seen that during the magnetic storm the Π_d flux values significantly differed from the values under quiet conditions. As can be seen from Fig. 3, the magnetic storm has led to the fact that on 13 and 14 November in the altitude range of 300-400 km there was a downward plasma flux due to ambipolar diffusion, both at night and daytime hours. On 15 November variations of Π_d were similar to the variations in plasma flux density due to ambipolar diffusion in quiet conditions.

The calculations show that in the nighttime (13 November) during the magnetic storm main phase, the Π_d value increased compared with the reference day about 1.25, 3 and 5.9 times at 300, 350 and 400 km. At $z = 250$ km, the magnetic storm effects in variations of Π_d were invisible.

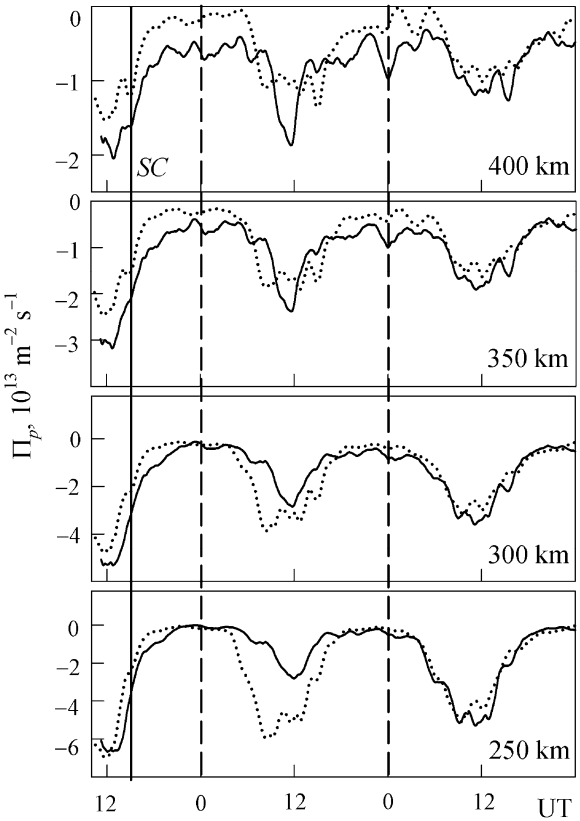


Fig. 2. Temporal variations of the full plasma flux density during 13-15 November 2012 magnetic storm (solid line) and quiet condition period on 21-23 November 2012 (dots).

Figure 4 shows the variation of the zonal electric field component E_y, the vertical component of the velocity of plasma transport due to the electromagnetic drift v_{EB}, and neutral wind velocity v_{nx} during the 13-15 November 2012 magnetic storm and the reference days on 21-23 November 2012.

Zonal component of the electric field. Figure 4 (top panel) shows the calculation of temporal variations of the electric field zonal component. Calculations showed that the value of the electric field zonal component was –9.5 mV/m during the 13-15 November 2012 magnetic storm. In quiet conditions, the value of E_y does not exceed a few mV/m.

Plasma drift velocity. Figure 4 (middle panel) shows the calculation results of the plasma drift velocity vertical component during magnetic storm and quiet conditions. It was found that during the main phase of the storm

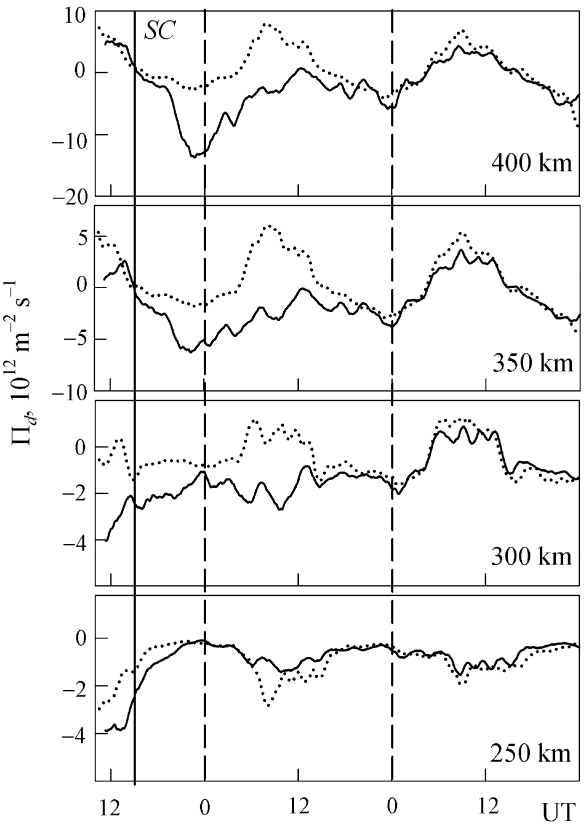


Fig. 3. Temporal variations of the plasma flux density due to ambipolar diffusion during 13-15 November 2012 magnetic storm (solid line) and quiet condition period on 21-23 November 2012 (dots).

the values of v_{EB} reached –85 m/s, whereas in quiet conditions the plasma transport due to the electromagnetic drift is absent.

Neutral wind. Figure 4 (bottom panel) shows the temporal variation of the meridional neutral wind velocity v_{nx} during the magnetic storm and quiet days at an altitude of 300 km. We see that in quiet conditions the wind velocity has poleward direction and ranges from 0 to –150 m/s. After the beginning of the magnetic storm and during the main phase, there was a change of the neutral wind direction to the equator, and calculations showed that the highest rate of v_{nx} was 150 m/s.

Variations of the meridional component of the neutral wind during 13-15 November 2012 magnetic storm are like the variations of v_{nx} during the 5-6 August 2011 magnetic storm. In both cases there is a change in the

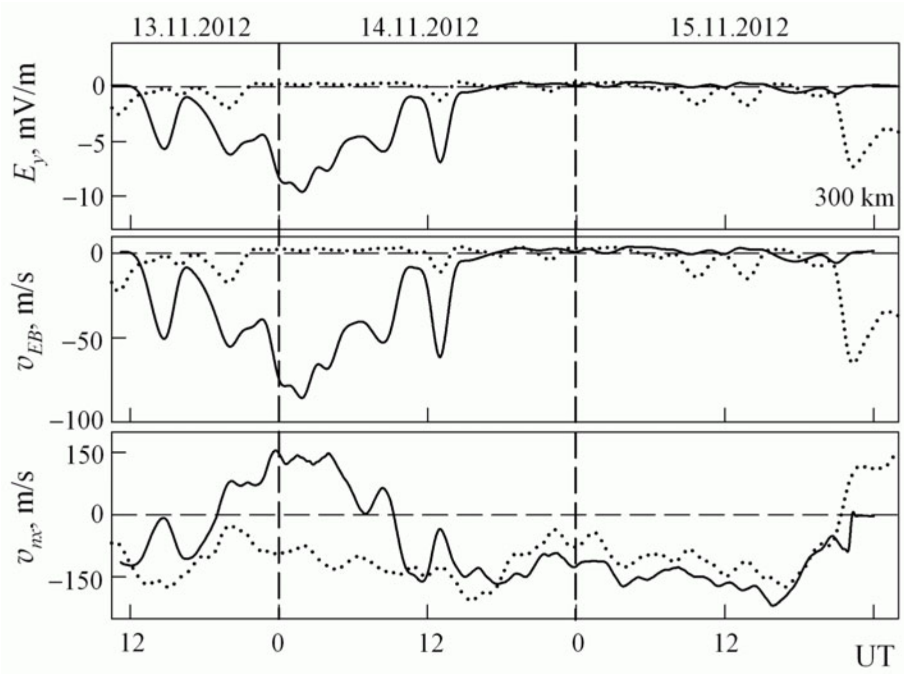


Fig. 4. Temporal variations of the electric field zonal component values, vertical component of the plasma drift velocity and neutral wind velocity during the 13-15 November 2012 magnetic storm (solid line) and quiet condition period on 21-23 November 2012 (dots).

direction of the neutral gas transport in the main phase of the magnetic storm (Domnin *et al*. 2014c).

Energy influx to the electrons. Figure 5 shows the calculation results of the energy *Q/N*, supplied to the electrons during the 13-15 November 2012 magnetic storm and quiet period of 21-23 November 2012.

In the disturbed day on 14 November 2012, during the negative phase of ionospheric storm and depressions of the electron concentration (F2-layer critical frequency foF2 ≈ 4 MHz at 07:30 UT) there was observed a dip in the *Q/N* variations to values of $2.5{\cdot}10^{-21}$, $1.2{\cdot}10^{-21}$, $0.7{\cdot}10^{-21}$, and $0.4{\cdot}10^{-21}$ J/s at altitudes of 200, 250, 300, and 350 km, respectively. On 22 November at the same time the values *Q/N* at the same altitudes were as follows: $3.5{\cdot}10^{-21}$, $1.8{\cdot}10^{-21}$, $0.9{\cdot}10^{-21}$, and $0.5{\cdot}10^{-21}$ J/s. At noon on 14 November the values of *Q/N* were $2.95{\cdot}10^{-21}$, $1.35{\cdot}10^{-21}$, $0.7{\cdot}10^{-21}$, and $0.5{\cdot}10^{-21}$ J/s. During the recovery phase of magnetic storm (15 November 2012) the value of *Q/N* increased and around noon the values of *Q/N* reached $4.2{\cdot}10^{-21}$, $1.9{\cdot}10^{-21}$, 10^{-21} and $0.5{\cdot}10^{-21}$ J/s at the corresponding altitudes.

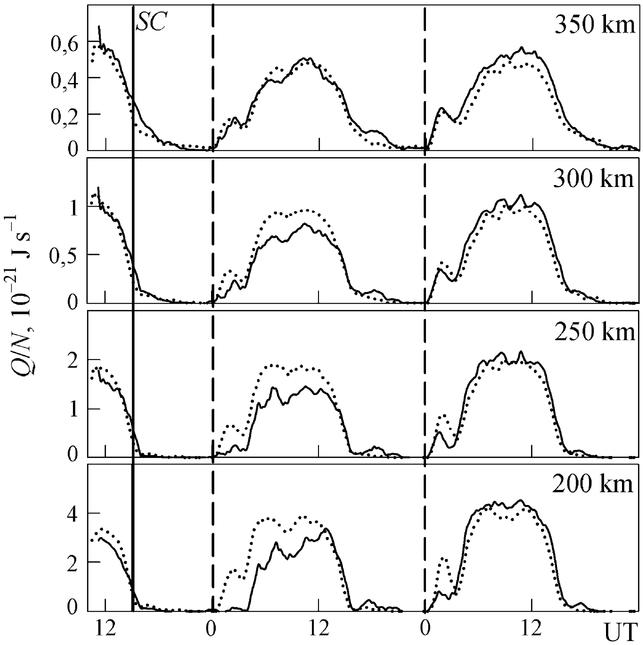


Fig. 5. Temporal variations of the energy supplied to the electron gas, *Q*/*N,* at the fixed altitudes during 13-15 November 2012 magnetic storm (solid line) and quiet condition period on 21-23 November 2012 (dots).

In general, the reduction of *Q/N* during the main phase of the magnetic storm compared to the reference days amounted to approximately 35, 25, and 20% at altitudes of 200, 250, and 300 km, respectively. At an altitude of 350 km the magnetic storm effects in the variations of the *Q/N* were not discernable.

Heat flux. The calculation results of the heat flux density Π_T transferred by electrons from the plasmasphere into the ionosphere during the 13-15 November 2012 magnetic storm and the quiet period of 21-23 November 2012 are shown in Fig. 6.

Calculations show that during the main phase of magnetic storm (14 November 2012), the absolute values of the heat flux density $|\Pi_T|$ increased relative to the values of Π_T in quiet conditions. This phenomenon is due to the unusual heating of the plasma in the night, when the temperatures of T_e and T_i almost reached the daily values. Plasma heating took place against the backdrop of a deep depression of the electron density in the F-region of the ionosphere (*N* decreased more than 5 times). During the greatest decrease in the electron density, the heat flux density Π_T was $-(1.3\text{-}1.2)\cdot10^{-5}$ W/m^2 in the altitude range of 200-350 km. The least value of Π_T on 14 November was

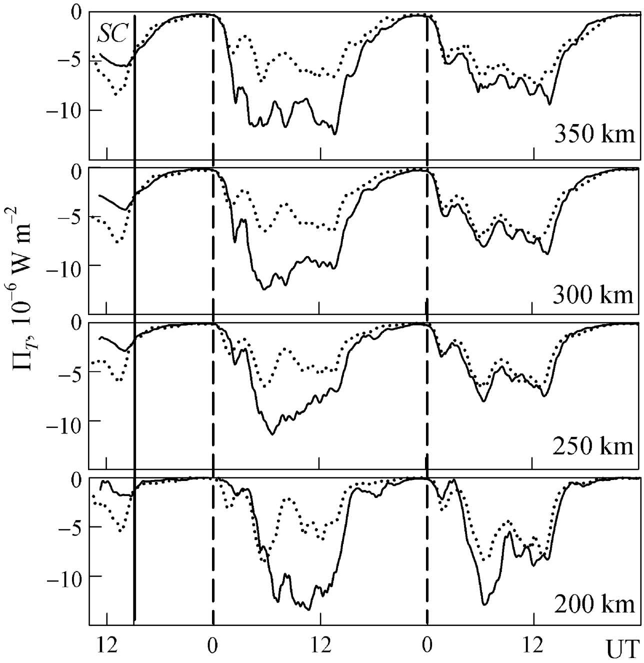


Fig. 6. The temporal variations of the heat flux density Π_T during 13-15 November 2012 magnetic storm (solid line) and quiet condition period on 21-23 November 2012 (dots).

observed at $z = 200$ km around 11:00 UT and was of about $-1.3 \cdot 10^{-5}$ W/m^2, whereas on 22 November the value was $\Pi_T \approx -0.5 \cdot 10^{-5}$ W/m^2.

On 15 November the absolute values of $|\Pi_T|$ were greater than on the reference day of 23 November. At $z = 200$ km, the least value of Π_T was about $-1.3 \cdot 10^{-5}$ W/m^2, and in the reference day at the same time $\Pi_T \approx -0.8 \cdot 10^{-5}$ W/m^2. In the afternoon, there was a recovery of values of heat fluxes.

6. CONCLUSIONS

The effects of the strong geospace storm on 13-14 November 2012 were well marked in variations of the electron density, electron and ion temperatures, and the vertical component of the plasma transfer velocity in the ionosphere over Kharkov.

Calculations show that during the magnetic storm on 13-14 November 2012 there has been a significant change in the dynamic and thermal modes of the ionospheric plasma in a wide altitude range.

On the background of a significant decrease in the electron density during the main phase of the magnetic storm on 14 November 2012, the rate of heating of the electron gas, Q/N, was approximately 2 times less than in the reference day on November 15. This led to an increase in the heat flux density caused by the electron temperature increases.

Due to the decrease in electron density, the full plasma flux values on 14 November were generally less than for undisturbed conditions. At altitudes of 350 and 400 km in the noon hours there was an upsurge in the Π_p variations, which is higher than the value observed on 22 November.

Strong geomagnetic storm on 13-14 November 2012, has caused a whole range of processes accompanying the plasma disturbances, electric and magnetic fields in various areas of the near-Earth space. The storm effects in the variations of dynamic and thermal processes in the ionosphere are evident. Temporal variations in the absolute values of the vertical component of velocity due to ambipolar diffusion and plasma transfer due to the electromagnetic drift during ionospheric disturbance have increased. There was an increase in the v_{dz} velocity by two times at $z = 250$ km and v_{EB} from –85 to 0 m/s at an altitude of 300 km. In turn, the meridional component of the neutral wind velocity changed in the opposite direction and became directed to the equator. The highest value of v_{nx} was about 150 m/s.

The results of calculations of dynamic and thermal process parameters in mid-latitude ionosphere during 13-14 November 2012 geospace storm supplement the global picture of the storm and its effects in the geospace.

Acknowledgements. The author thanks Prof. Igor F. Domnin and Prof. Leonid F. Chernogor for their interest in this study, as well as Dr. Leonid Ya. Emelyanov, Anatoly F. Kononenko and Yakov N. Chepurnyy for conducting experiments on Kharkov incoherent scatter radar.

References

Blanc, M., and P. Amayenc (1979), Seasonal variations of the ionospheric E×B drift above Saint-Santin on quite days, *J. Geophys. Res.* **84**, A6, 2691-2704, DOI: 10.1029/JA084iA06p02691.

Blanc, M., P. Amayenc, P. Bauer, and C. Taieb (1977), Electric field induced from the French incoherent scatter facilities, *J. Geophys. Res.* **82**, 1, 87-97, DOI: 10.1029/JA082i001p00087.

Buonsanto, M.J. (1999), Ionospheric storms: A review, *Space Sci. Rev.* **88**, 3-4, 563, DOI: 10.1023/A:1005107532631.

Buonsanto, M.J., S.A. Gonzalez, G. Lu, B.W. Reinisch, and J.P. Thayer (1999a), Coordinated incoherent scatter radar study of the January 1997 storm, *J. Geophys. Res.* **104**, A11, 24625-24637, DOI: 10.1029/1999JA900358.

Buonsanto, M.J., S.A. Gonzalez, X. Pi, J.M. Ruohoniemi, M.P. Sulzer, W.E. Swartz, J.P. Thayer, and D.N. Yuan (1999b), Radar chain study of the May, 1995 storm, *J. Atmos. Sol.-Terr. Phys*. **61**, 3-4, 233-248, DOI: 10.1016/S1364-6826(98)00134-5.

Burmaka, V.P., and L.F. Chernogor (2012), Wave disturbances in the ionosphere during a lasting solar activity minimum, *Geomagn. Aeron.* **52**, 2, 183-196, DOI: 10.1134/S001679321202003X.

Chernogor, L.F., Ye.I. Grigorenko, V.N. Lysenko, and V.I. Taran (2007), Dynamic processes in the ionosphere during magnetic storms from the Kharkov incoherent scatter radar observations, *Int. J. Geomagn. Aeron.* **7**, GI3001, DOI: 10.1029/2005GI000125.

Domnin, I.F., L.Ya. Emelyanov, M.V. Lyashenko, and L.F. Chernogor (2014a), Partial solar eclipse of January 4, 2011 above Kharkiv: observation and simulation results, *Geomagn. Aeron.* **54**, 5, 583-592, DOI: 10.1134/S0016793214040112.

Domnin, I.F., L.Ya. Emelyanov, S.V. Katsko, and L.F. Chernogor (2014b), Ionospheric effects of geospace storm of November 13-14, 2012, *Radio Phys. Radio Astron.* **19**, 2, 170-180 (in Russian).

Domnin, I.F., C. La Hoz, and M.V. Lyashenko (2014c), Variation of the electric field zonal component, the vertical component of the plasma drift and neutral wind velocities in ionosphere over Kharkov (Ukraine) during August 5-6, 2011 and November 13-15, 2012 magnetic storms, *Bull. Nation. Tech. Univ. "Kharkiv Polytechnic Institute". Series: Radio Physics and Ionosphere* **47**, 15-21.

Emelyanov, L.Ya., and T.G. Zhivolup (2013), History of the development of IS radars and founding of the Institute of Ionosphere in Ukraine, *Hist. Geo Space Sci*. **4**, 1, 7-17, DOI: 10.5194/hgss-4-7-2013.

Finlay, C.C., S. Maus, C.D. Beggan, T.N. Bondar, A. Chambodut, T.A. Chernova, A. Chulliat, V.P. Golovkov, B. Hamilton, M. Hamoudi, R. Holme, G. Hulot, W. Kuang, B. Langlais, V. Lesur, F.J. Lowes, H. Lühr, S. MacMillan, M. Mandea, S. McLean, C. Manoj, M. Menvielle, I. Michaelis, N. Olsen, J. Rauberg, M. Rother, T.J. Sabaka, A. Tangborn, L. Tøffner-Clausen, E. Thébault, A.W.P. Thomson, I. Wardinski, Z. Wei, and T.I. Zvereva (2010), International geomagnetic reference field: the eleventh generation, *Geophys. J. Int.* **183**, 3, 1216-1230, DOI: 10.1111/j.1365-246X.2010.04804.x.

Goncharenko, L.P., J.E. Salah, A. van Eyken, V. Howells, J.P. Thayer, V.I. Taran, B. Shpynev, Q. Zhou, and J. Chau (2005), Observations of the April 2002 geomagnetic storm by the global network of incoherent scatter radars, *Ann. Geophys.* **23**, 1, 163-181.

Grigorenko, E.I., V.N. Lysenko, V.I. Taran, and L.F. Chernogor (2005a), Specific features of the ionospheric storm of March 20-23, 2003, *Geomagn. Aeron.* **45**, 6, 745-757.

Grigorenko, E.I., S.A. Pazyura, V.I. Taran, L.F. Chernogor, and S.V. Chernyaev (2005b), Dynamic processes in the ionosphere during the severe magnetic storm of May 30-31, 2003, *Geomagn. Aeron.* **45**, 6, 758-777.

Grigorenko, E.I., V.N. Lysenko, S.A. Pazyura, V.I. Taran, and L.F. Chernogor (2007), Ionospheric disturbances during the severe magnetic storm of November 7-10, 2004, *Geomagn. Aeron.* **47**, 6, 720-738, DOI: 10.1134/S0016793207060059.

Immel, T.J., G. Liu, S.L. England, L.P. Goncharenko, P.J. Erickson, M.V. Lyashenko, M. Milla, J. Chau, H.U. Frey, S.B. Mende, Q. Zhou, A. Stromme, and L.J. Paxton (2015), The August 2011 URSI World Day campaign: Initial results, *J. Atmos. Sol.-Terr. Phys*. **134**, 47-55, DOI: 10.1016/j.jastp.2015.09.005.

Lyashenko, M.V. (2013), The effects of the partial solar eclipse on January 4, 2011 in the variety of thermal process parameters in ionosphere, *Sun Geosph.* **8**, 1, 15-18.

Lyashenko, M.V., and L.F. Chernogor (2013), Solar eclipse of August 1, 2008, above Kharkov: 3. Calculation results and discussion, *Geomagn. Aeron.* **53**, 3, 367-376, DOI: 10.1137/S0016793213020096.

Ogawa, T., Y. Tanaka, A. Huzita, and M. Yasuhara (1975), Horizontal electric fields in the middle latitude, *Planet. Space Sci*. **23**, 5, 825-830, DOI: 10.1016/0032-0633(75)90019-7.

Picone, J.M., A.E. Hedin, D.P. Drob, and A.C. Aikin (2002), NRLMSISE-00 empirical model of the atmosphere: statistical comparisons and scientific issues, *J. Geophys. Res*. **107**, A12, 1468-1483, DOI: 10.1029/2002JA009430.

Richmond, A.D., M. Blanc, B.A. Emery, R.H. Wand, B.G. Fejer, R.F. Woodman, S. Ganguly, P. Amayenc, R.A. Behnke, C. Calderon, and J.V. Evans (1980), An empirical model of quite-day ionospheric electric fields at middle and low latitudes, *J. Geophys. Res*. **85**, A9, 4658-4664, DOI: 10.1029/JA085iA09p04658.

Schunk, R.W., and A.F. Nagy (2000), *Ionospheres: Physics, Plasma Physics, and Chemistry*, Cambridge Atmospheric and Space Science Series, Cambridge University Press, New York, 555 pp.

Sergeenko, N.P. (1982), Estimates of electric fields during ionospheric disturbances. **In:** R.A. Zevakina, and N.P. Sergeenko (eds.), *Ionospheric Forecasting*, Nauka, Moscow, 91-96 (in Russian).

Stubbe, P. (1968), Frictional forces and collision frequencies between moving ion and neutral gases, *J. Atmos. Terr. Phys.* **30**, 12, 1965-1985, DOI: 10.1016/0021-9169(68)90004-4.

Received 27 February 2015
Received in revised form 10 November 2015
Accepted 18 February 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2734-2747
DOI: 10.1515/acgeo-2016-0102

Seasonal Variations of Mid-Latitude Ionospheric Trough Structure Observed with DEMETER and COSMIC

Barbara MATYJASIAK, Dorota PRZEPIÓRKA, and Hanna ROTHKAEHL

Space Research Centre of Polish Academy of Sciences, Warsaw, Poland,
e-mail: bmatyjasiak@cbk.waw.pl

Abstract

The mid-latitude ionospheric trough is a depleted region of ionospheric plasma observed in the topside ionosphere. Its behavior can provide useful information about the magnetospheric dynamics, since its existence is sensitive to magnetospherically induced motions. Mid-latitude trough is mainly a night-time phenomenon. Both, its general features and detailed characteristics strongly depend on the level of geomagnetic disturbances, time of the day, season, and the solar cycle, among others. Although many studies provide basic information about general characteristics of the main ionospheric trough structure, an accurate prediction of the trough behavior in specific events is still understood poorly. The paper presents the mid-latitude trough characteristics with regard to the geomagnetic longitude and season during a solar activity minimum, as based on the DEMETER in situ satellite measurements and the data retrieved from FORMOSAT-3/COSMIC radio occultation measurements.

Key words: main ionospheric trough, ionosphere, electron density, DEMETER, FORMOSAT-3/COSMIC.

Ownership: Institute of Geophysics, Polish Academy of Sciences

1. INTRODUCTION

The main ionospheric trough (MIT) is a typical feature observed in the sub-auroral ionospheric region which has been studied for many years (Muldrew 1965, Spiro 1987, Karpatchev 2003, He *et al.* 2011, Zaalova *et al.* 2013). The MIT strongly affects the propagation of different natural and artificial signals, thus itis important for radio communication as well as communication and space industry in general (Clilverd *et al.* 1995).

The trough structure is most developed for altitudes of the F2 layer maximum where significant depletion in electron concentration occurs (during a whole year). The structure seems to exist in the areas where the stagnation of ionospheric plasma takes place, between a region of eastward (corotating) plasma drift and a region of westward drift in the auroral zone (Mallis and Essex 1993, Whalen 1989, Aladjev *et al.* 2001). Several convection patterns which do not interact with a production region for several hours are present in this area, which allows the plasma density decay to a low value.

The trough shape is narrow in latitudes but extended in longitudes. It is typical that poleward edge is a steep wall with sudden increase of electron density, and that the amplitude even increases during the higher geomagnetic activity. The equatorial edge is flatter and it reaches the normal night-time F layer electron density level at lower latitudes in a less steep manner (Muldrew 1965, Werner and Prölss 1997). Position of MIT poleward edge is considered to be connected with the magnetospheric plasmapause. Yizengaw *et al.* (2005) have shown that the two features lie on the same field lines during geomagnetically quiet time period, which became an important step in modeling the plasmapause position at ionospheric heights. However, the correlation disappears during strong geomagnetic disturbances and plasmapause occurs inside the trough wall (Rothkaehl *et al.* 1997). MIT is also connected with an auroral oval, because the research of this structure indirectly gives us information about the polar region physics as well as the dynamics.

Many studies show that MIT mostly occurs during the night-time and its intensity is changing together with the distance from local midnight (Moffett and Quegan 1983). The mid-latitude trough is a very dynamic structure which is highly-reactive to changes in geomagnetic conditions. Studies have shown that the structure moves to the lower latitudes and becomes shallower with increasing level of geomagnetic activity (Moffett and Quegan 1983). Some changes in MIT location for different longitudes were also observed. Even though the dependence of the main ionospheric structures has been detected, its direct source is still not well known.

In this paper we study the general characteristics of the trough structure during the last long solar minimum. This solar cycle decline, and in particular minimum phase, seems to be unusual and that appears to be related to weak solar polar magnetic fields. Thus, studies of the behavior of the main ionospheric trough during this period are important for good understanding of ionospheric response to specific solar conditions. The analysis was made in respect to the season and level of the geomagnetic activity. The data from two different missions were compared.

2. DATA COLLECTION

In order to analyze the characteristic behavior and features of the main ionospheric trough, data received from two missions were used, namely DEMETER satellite and FORMOSAT-3/COSMIC constellation. DEMETER was a micro-satellite operating from June 2004 to 2010, located at circular orbit with inclination of 98.3°. Originally orbiting at an altitude of 710 km, it lowered to 660 km in 2005. One of the instruments present onboard DEMETER was Langmuir probe which provides measurements of electron density and temperature (Lebreton *et al.* 2006). Due to limited time of the instrument operation per orbit, the data from mission reach only ±65° of geomagnetic latitude. Because of the night-time character of MIT, data from night orbits were chosen and divided into two groups, depending on K_p index value. The first group, with index between 0 and 2, was considered as a low geomagnetic activity, while the second one, with values in the range of 4-6, as a high geomagnetic activity (*e.g.*, magnetic storm).

FORMOSAT-3/COSMIC is a joint U.S.-Taiwanese project consisting of six micro-satellites using radio occultation method to obtain information about meteorological conditions in the Earth's atmosphere. The mission was successfully launched into a circular low-Earth orbit on 15 April 2006. The initial altitude of micro-satellites was 512 km and it was gradually rising to the final altitude of about 800 km, an orbit plane inclination angle of 72° and a separation angle between neighboring orbit planes of 24° in longitude. This RO method uses signal of operating GPS satellites to determine, *e.g.*, atmospheric profiles, vertical water vapor content over the whole Earth. Specific character of this method provides measurements in multiple points. From one day of observation, even 2500 atmospheric profiles can be calculated. It allows us to get 3-dimensional information about atmospheric conditions.

The data from FORMOSAT-3/COSMIC's were selected at an altitude of 650 km to make them comparable with those from DEMETER satellite. Also time range was chosen so that it covers local night-time. Both data sets are presented in geomagnetic dipole coordinate system. To distinguish from different geomagnetic conditions that cause varied ionospheric behavior, data also were divided into two groups, basing on value of Kp index (Kp = (0,2) and Kp = (4,6)). For both datasets seasonal maps were created in order to observe and analyze typical behavior of trough structure depending on solar flux condition and geomagnetic activity.

3. SEASONAL CHANGES OF MAIN IONOSPHERIC TROUGH POSITION

The main ionospheric trough is a phenomenon strongly correlated with solar flux and geomagnetic activity conditions. Studies show that its location and

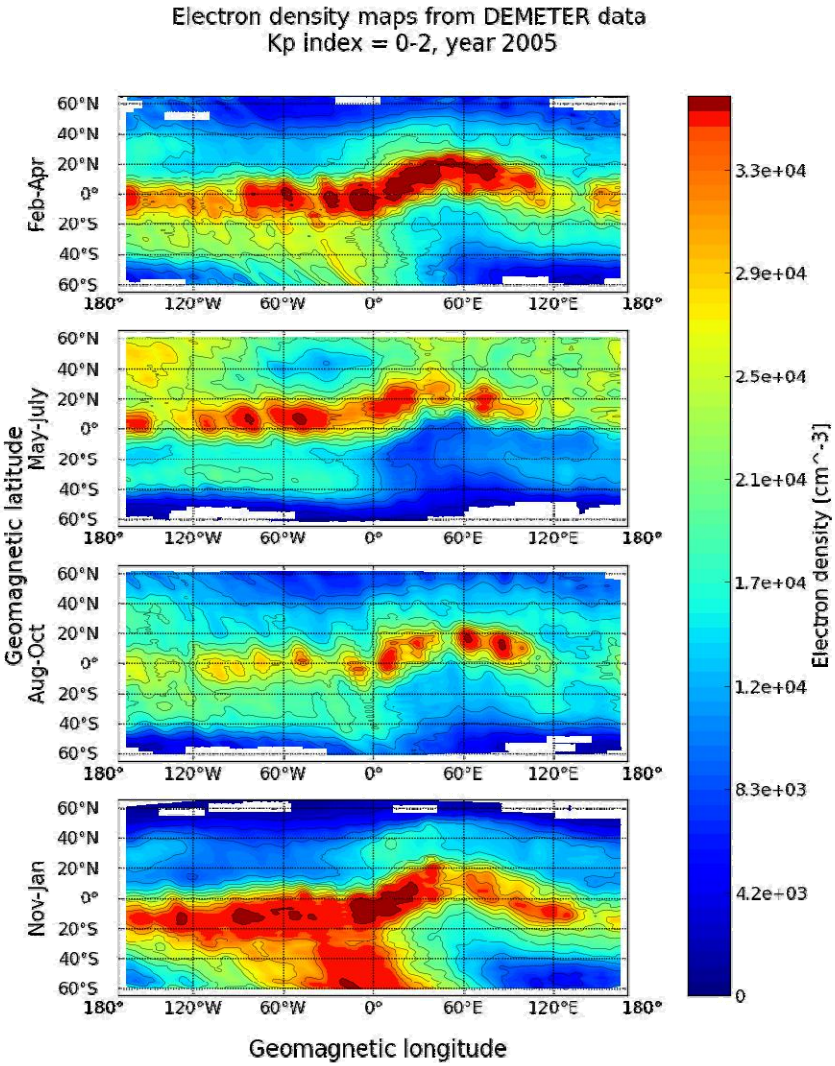


Fig. 1. Maps of seasonal changes in electron density derived from DEMETER data for K_p index in the range of 0-2 in 2005.

shape change during different seasons. The data from DEMETER mission were selected to find typical insolation-correlated behavior of MIT on both hemispheres. The measurements for night-time orbits from Langmuir probe instrument were picked and divided into four time periods (representing the four seasons), *i.e.*, February-April, May-July, August-October, and November-January. Data points were gridded over almost whole range of geomagnetic longitude and for geomagnetic latitudes from -65° to 65°. Similar

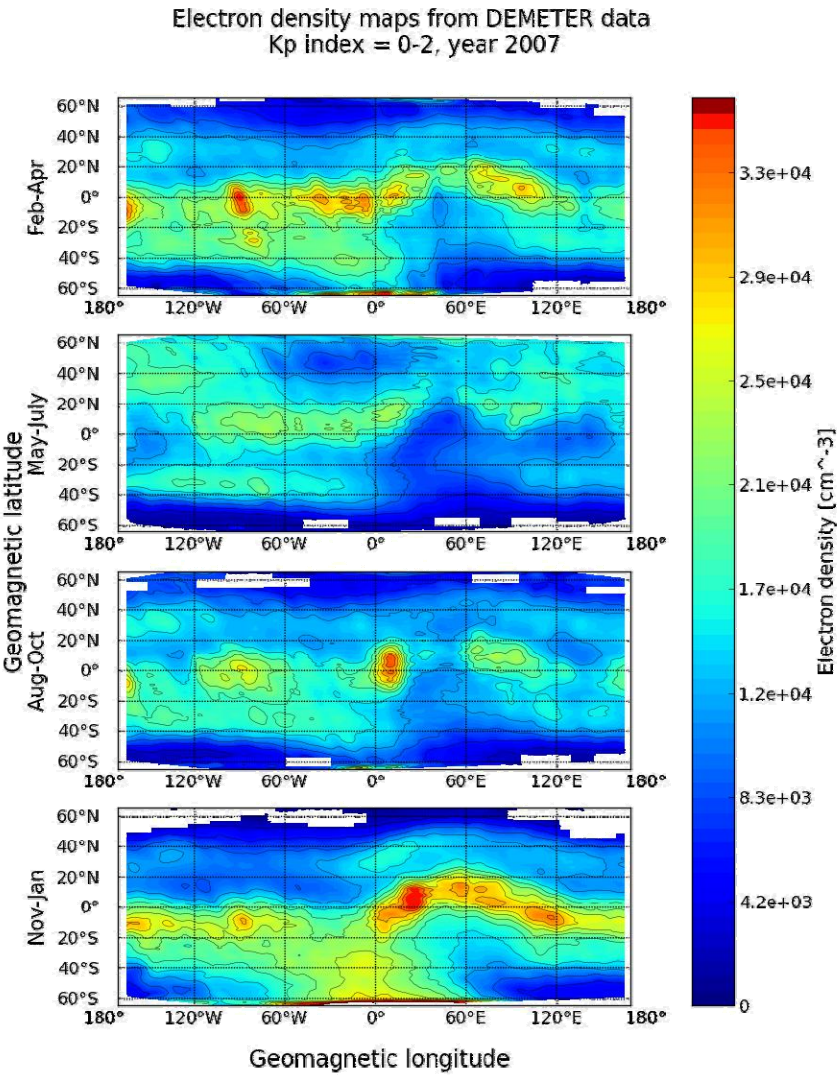


Fig. 2. Maps of seasonal changes in electron density derived from DEMETER data for K_p index in the range of 0-2 in 2007.

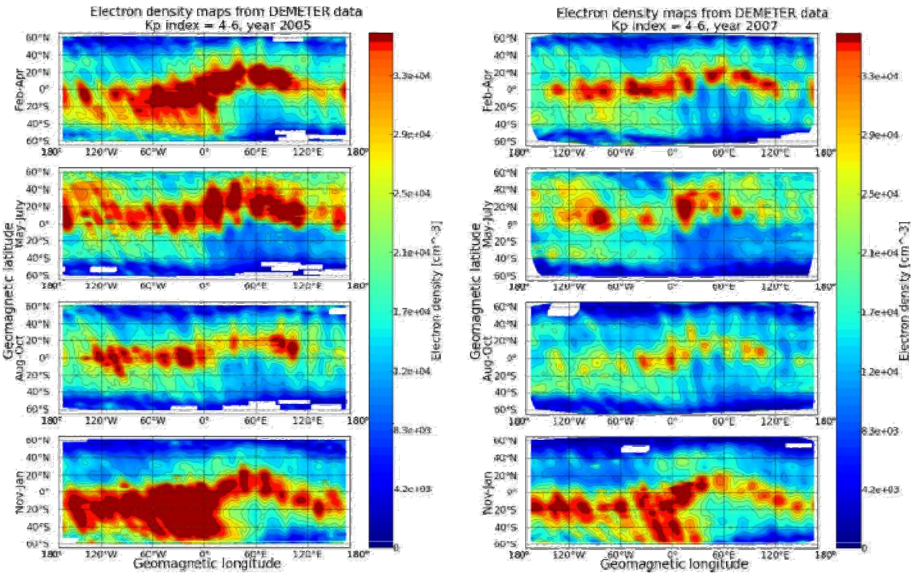


Fig. 3. Maps of seasonal changes in electron density derived from DEMETER data for K_p index in the ranges of 4-6 in 2005 (left figures) and 2007 (right figures).

processing was applied for FORMOSAT-3/COSMIC data but the range available for geomagnetic latitude was higher. This allowed better analysis of the trough and its adjacent region. Also, to compare with DEMETER night-time data (local time 10:20 PM) only profiles in time range from 9:00 PM to 3:00 AM were chosen.

We can observe that the trough structure is most visible during the local winter on both hemispheres. It happens because of the shorter time when the ionosphere is sunlit. As mentioned before, the limits of DEMETER data result in the fact that only the equatorial edge of the trough and part of inner structure can be observed (blue edges of the map Figs. 1 and 2). The data show that there are regions where MIT exists almost all year long and where it is most visible under favorable conditions. For the northern hemisphere the trough can be easily observed during winter solstice months, *i.e.*, from November till January, both on maps from DEMETER (bottom panel Figs. 1 and 2) and FORMOSAT-3 data (bottom left panel Fig. 4). It covers almost whole range of geomagnetic longitudes and the equatorial edge is located at about 50º-55º of geomagnetic latitude. On the contrary, during the same period of time, the structure on the southern hemisphere is hardly visible, and for some longitudes only. We can find weak trough for geomagnetic longitudes from 80º E to 160º W at typical latitudinal location, *i.e.* about -62º.

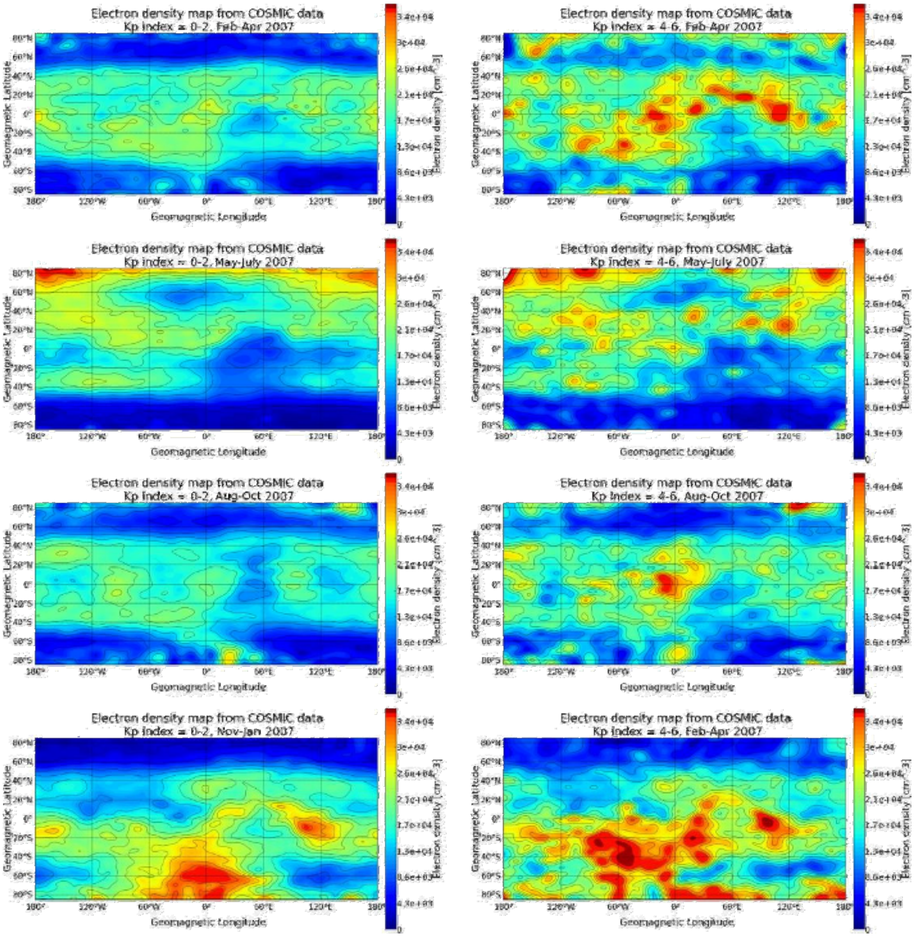


Fig. 4. Maps of seasonal changes in electron density derived from FORMOSAT-3/COSMIC data for K_p index in the ranges of 0-2 (left figures) and 4-6 (right figures) in 2007. Data selected for 650 km altitude.

When comparing troughs on the southern and the northern hemisphere we find out that during equinoxes and even local summer the trough on the southern hemisphere appears to be deeper and more visible locally than northern trough in similar solar flux conditions (Fig. 2 and left panel of Fig. 4). It could underpin the conjecture that mechanisms and physics responsible for MIT formation and behavior on the southern and northern hemispheres are quite different. In addition, the presence of Weddel Sea Anomaly on the southern hemisphere may have a significant impact on the appearance of the MIT. The night-time trough is more developed over the region of Eastern Indian Ocean because of the higher ionization available

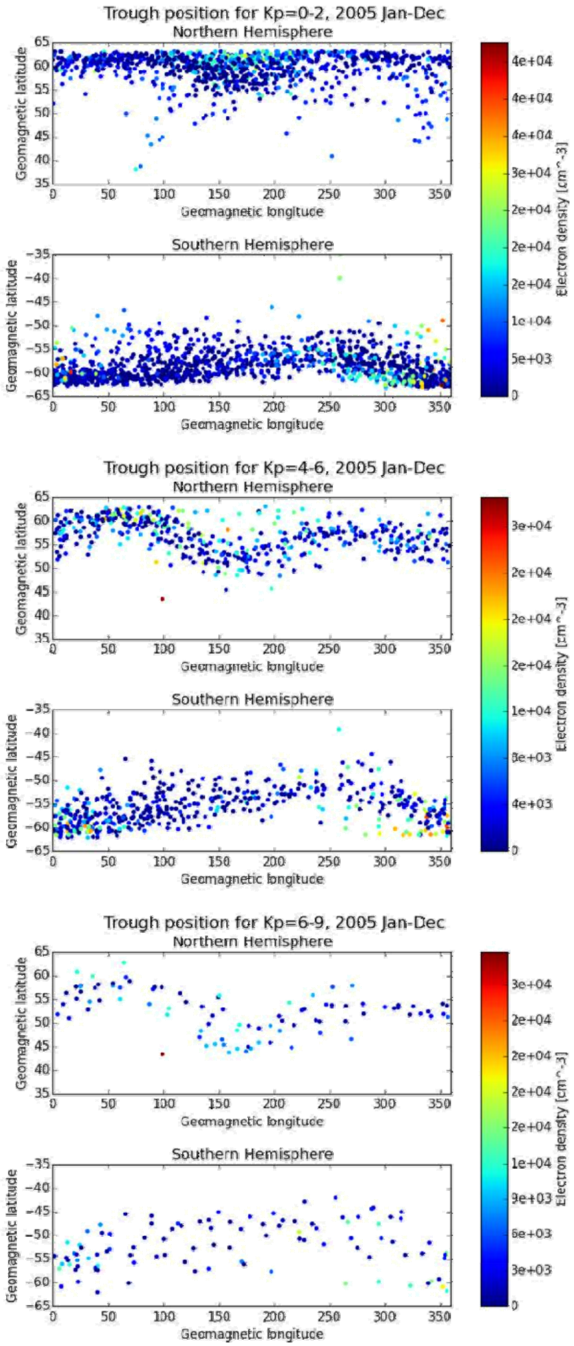


Fig. 5. Trough minimum position derived from DEMETER data for K_p index in the ranges of 0-2 (top figure), 4-6 (middle figure), and 6-9 (bottom) in 2005.

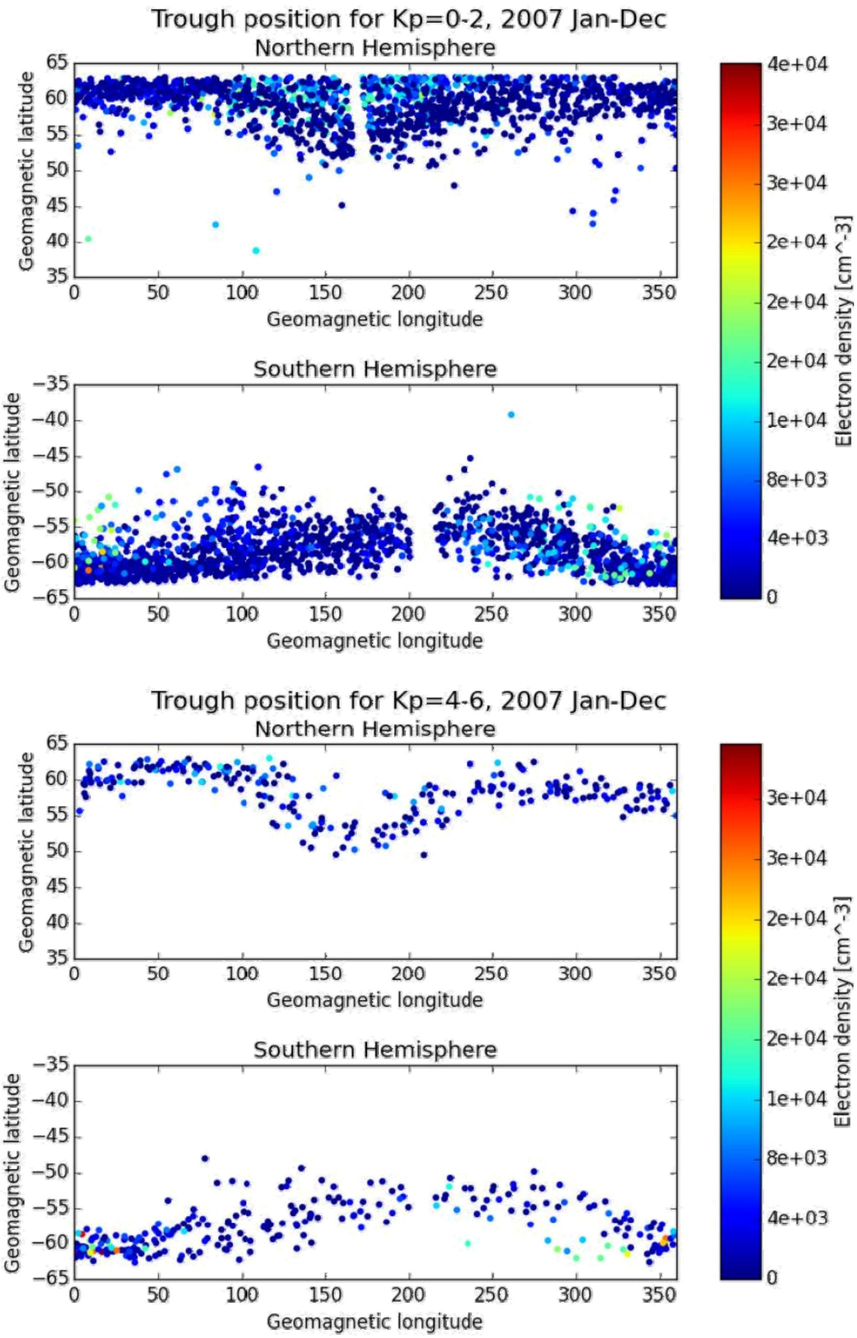


Fig. 6. Trough minimum position derived from DEMETER data for K_p index in the ranges of 0-2 (top figure) and 4-6 (bottom figure) in 2007.

(due to Weddell Sea Anomaly) than over the Mid Indian Ocean where the electron density level is lower (Horvath and Essex 2011).

In terms of geomagnetic conditions, one of the features observed is that the higher electron densities come with the higher K_p index. This can be easily explained by higher particle precipitation and the trough bottom filled by electrons, through expansion of auroral oval region. We can also find greater amplitude between regions with high and low electron density (Fig. 3 and right panels of Fig. 4). Another important thing is that general annual geomagnetic conditions affect MIT characteristics as can be observed when we compare the maps created from 2005 and 2007 year. For DEMETER data from 2007, which is considered as relatively quiet period, global electron densities are lower than for data from 2005 and for the same K_p index range (compare Figs. 1, 2, and 3). During 2005 we have experienced high solar activity with numerous strong solar flares and coronal mass ejection events, thus the Earth magnetic field was often disturbed. It resulted in general higher geomagnetic activity and enhanced energetic particles precipitation and accumulation during the disturbed periods.

The comparison between data from DEMETER and FORMOSAT-3 shows good agreement of global structures. As a result, we can consider the behavior of mid-latitude trough presented here as typical for seasonal changes.

4. TROUGH MINIMUM POSITION

The typical location for MIT structure during the solar activity maximum is estimated to be about 60º-65º of geomagnetic latitude. Simple analysis was made in order to find the main trough location for time period corresponding to solar activity minimum. From DEMETER data, night-time orbits have been selected and ranges corresponding to the main ionospheric trough location have been picked (*i.e.*, 35º to 65º geomagnetic latitude). The region where density minimum was clearly visible and which was correlated with electron temperature maximum was considered as trough minimum. Orbits were divided into three groups, for low, moderate and high geomagnetic activity. Kp index in the ranges: 0-2, 4-6, and 6-9 has been chosen as an indicator of geomagnetic activity. In Fig. 5 plots for the trough minimum position in 2005 for different geomagnetic conditions are shown. From the quiet period we can find that the trough minimum locates at about 60º-63º, which is consistent with literature. Geomagnetic disturbances cause shift of the minimum trough location as expected (Muldrew 1965) and for moderate conditions it is placed at about 55º-60º (middle panel). There were several extremely strong geomagnetic events in 2005; thus, it was possible to make some statistics in case when K_p index exceeded the value of 6. Bottom panel

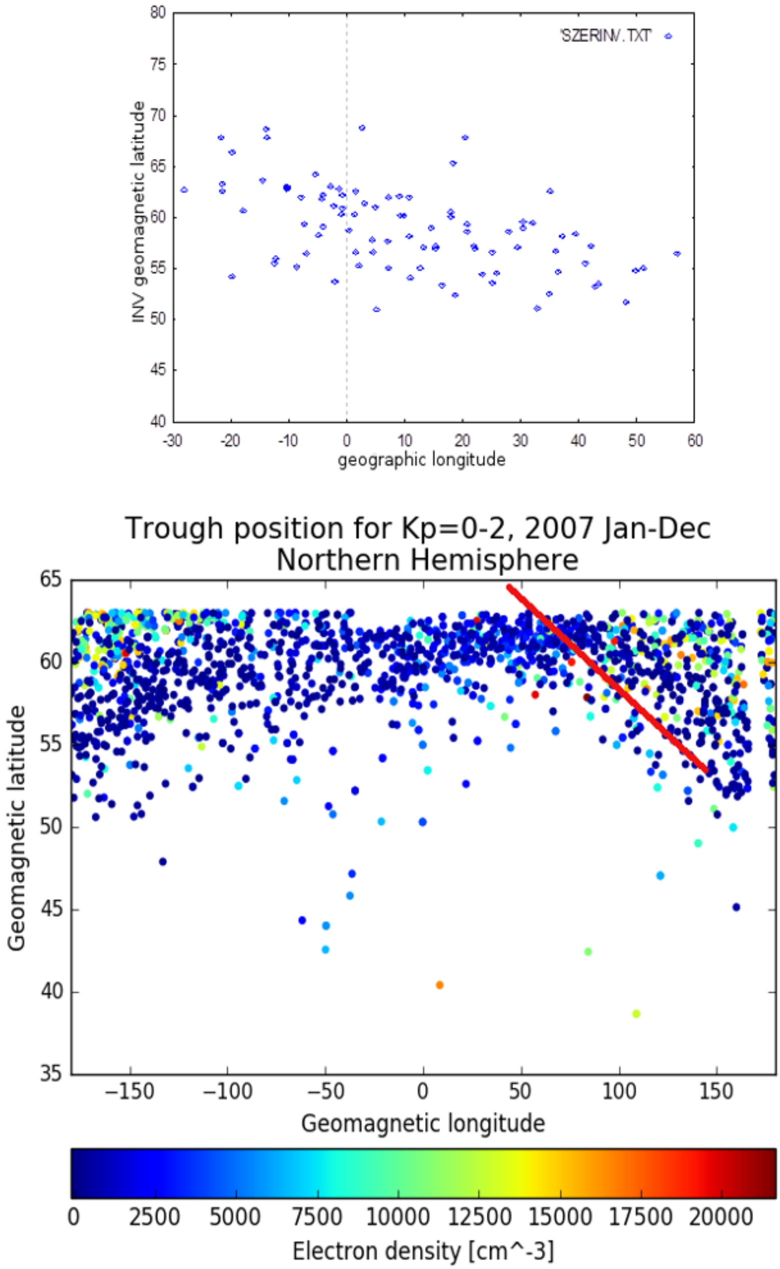


Fig. 7. Trough minimum position derived from radio measurements in the topside ionosphere from Magion-3 spacecraft (top) and from DEMETER data for K_p index in the range of 0-3 in 2005 (bottom). Red line corresponds to the position of Magion-3 measurements.

shows that for high geomagnetic activity the MIT minimum moves equatorward and can reach even 47º-50º of geomagnetic latitude. Figure 6 represents the estimated trough minimum positions for year 2007. Only two plots were made for low and moderate activity, as in 2007 extremely strong events were rarer. We find similar position for low K_p as in 2005. For the range 4-6, the trough minimum moves few degrees towards equator.

Analyzing the MIT minimum we can observe that its location differs with longitudes. There are regions where the structure estimated minimum is shifted to the lower latitudes. This situation takes place for longitudes in sector from 100º to 160º for the northern hemisphere. Similar behavior can be found for the southern hemisphere trough. Position of the minimum moves to the lower latitudes between 170º and 300º of geomagnetic longitude. Analyzing Fig. 7 we can find out that the trough position shift for this longitude sector is not an accidental feature, but rather a characteristic attribute since it can be found in measurements from different missions. For example, the trough minimum location displacement has been observed by Magion-3, operating between 18 December 1991 and December 1992 during the maximum solar activity. In contrary, the measurements conducted by DEMETER fall into the minimum solar activity. This leads to the hypothesis that the MIT longitudinal dependence may be associated with the Earth's local conditions such as neutral winds or particle flow rather than only with the geomagnetic conditions imposed by the Sun. However, further studies are needed.

5. CONCLUSIONS

The data from two space missions, DEMETER and FORMOSAT-3/COSMIC, were used to analyze the behavior of the mid-latitude trough structure during the solar minimum period. Annual variations of the shape, location and intensity of the phenomenon were examined by means of electron density at specified altitude, latitude and longitude. It was shown that longitudinal variations of position of mid-latitude trough are strongly correlated with seasonal changes and insolation. Local winter is for both hemispheres the period when the structure appears to be the deepest and well developed, while during local summer it becomes hardly visible and much weaker. However, the southern hemisphere trough structure appears as deeper and better visible even during the local summer.

What plays an important role in MIT characteristic changes are also the Earth's magnetic field disturbances and particle precipitation. It has been shown that the typical latitudinal location of mid-latitude trough minimum is about ±62º, but during the higher geomagnetic activity it moves equatorward even up to 5º. We can assume this to be a characteristic behavior since the data analysis from 2005 and 2007 year have shown some consistent results. Observations show also that the MIT tends to shift equatorward for some longitudes in both hemispheres. The mid-latitude trough is a dynamic structure sensitive to magnetic and particle disturbances. To better understand the complexity of its behavior, several further studies are needed.

Acknowledgements This work was conducted in the frame of the Polish-Taiwanese Joint Research Project for years 2012-2014 under the agreement between the Polish Academy of Sciences in Warsaw, Poland, and the National Science Council in Taipei, Taiwan. The work was partially funded by RELEC grant no. NCN 2012/07//04414. We acknowledge the University Corporation for Atmospheric Research (UCAR) for providing the COSMIC Data. We would like to thank the DEMETER Team for preparing the data. Additionally, we would like to thank M. Parrot, CNRS/LPCE Laboratory in Orleans, France and CDPP team for making the data available.

References

Aladjev, G.A., O.V. Evstaev, V.S. Mingalev, G.I. Mingaleva, E.D. Tereshchenko, and B.Z. Khudukon (2001), Interpretation of ionospheric F-region structures in the vicinity of ionization trough observed by satellite radio tomography, *Ann. Geophys.* **19**, 1, 25-36.

Clilverd, M.A., N.R. Thomson, and A.J. Smith (1995), The effect of the mid-latitude ionospheric trough on whistler mode ducting during magnetic storm, *J. Atmos. Terr. Phys.* **57**, 8, 945-954, DOI: 10.1016/0021-9169(94)00083-Z.

He, M., L. Liu, W. Wan, and B. Zhao (2011), A study on the night-time midlatitude ionospheric trough, *J. Geophys. Res.* **116**, A5, A05315, DOI: 10.1029/2010JA016252.

Horvath, I., and E.A. Essex (2003), The Weddell Sea Anomaly observed with the Topex satellite data, *J. Atmos. Sol. Terr. Phys.* **65**, 6, 693-706, DOI: 10.1016/S1364-6826(03)00083-X.

Karpachev, A.T. (2003), The dependence of the main ionospheric trough shape on longitude, altitude, season, local time, and solar and magnetic activity, *Geomagn. Aeron.* **43**, 2, 239-251.

Lebreton, J.-P., S. Stverak, P. Travnicek, M. Maksimovic, D. Klinge, S. Merikallio, D. Lagoutte, B. Poirier, P.-L. Blelly, Z. Kozacek, and M. Salaquarda (2006), The ISL Langmuir probe experiment processing onboard DEMETER: Scientific objectives, description and first results, *Planet. Space Sci.* **54**, 5, 472-486, DOI: 10.1016/j.pss.2005.10.017.

Mallis, M., and E. Essex (1993), Diurnal and seasonal variability of the southern-hemisphere main ionospheric trough from differentialphase measurements, *J. Atmos. Terr. Phys*. **55**, 7, 1021-1037, DOI: 10.1016/0021-9169(93) 90095-G.

Moffett, R.J., and S. Quegan (1983), The midlatitude trough in the electron concentration of the ionospheric Flayer: A review of observations and modeling, *J. Atmos. Terr. Phys*. **45**, 5, 315-343, DOI: 10.1016/S0021-9169(83)80038-5.

Muldrew, D.B. (1965), Flayer ionization trough deduced from Alouette data, *J. Geophys. Res.* **70**, 11, 2635-2650, DOI: 10.1029/JZ070i011p02635.

Rothkaehl, H., F. Jiricek, J. Smilauer, and M. Förster (1997), Dynamic changes in the outer ionosphere in the region of the ionospheric trough during and intense magnetic storm, *Adv. Space Res*. **20**, 3, 409-414, DOI: 10.1016/ S0273-1177(97)00701-1.

Spiro, R.W. (1978), Ion convection and the formation of the midlatitude F region ionization trough, *J. Geophys. Res.* **83**, A9, 4255-4264, DOI: 10.1029/ JA083iA09p04255.

Werner, S., and G.W. Prölss (1997), The position of the ionospheric trough as a function of local time and magnetic activity, *Adv. Space Res*. **20**, 9, 1717-1722, DOI: 10.1016/S0273-1177(97)00578-4.

Whalen, J.A. (1989), The daytime F layer trough and its relation to ionospheric-magnetospheric convection, *J. Geophys. Res*. **94**, A12, 17169-17184, DOI: 10.1029/JA094iA12p17169.

Yizengaw, E., H. Wei, M.B. Moldwin, D. Galvan, L. Mandrake, A. Mannucci, and X. Pi (2005), The correlation between mid-latitude trough and the plasmapause, *Geophys. Res. Lett*. **32**, 10, L10102, DOI: 10.1029/2005GL022954.

Zaalov, N.Y., H. Rothkaehl, A.J. Stocker, and E.M. Warrington (2013), Comparison between HF propagation and DEMETER satellite measurements within the mid-latitude trough, *Adv. Space Res*. **52**, 5, 781-790, DOI: 10.1016/j.asr. 2013.05.023.

Received 15 September 2015
Received in revised form 29 March 2016
Accepted 26 July 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2748-2760
DOI: 10.1515/acgeo-2016-0089

A Tiny Fabry–Perot Interferometer with Postpositional Filter for Measurement of the Thermospheric Wind

Houmao WANG[1,2], Yongmei WANG[1], and Jianguo FU[1]

[1]National Space Science Center, Chinese Academy of Sciences, Beijing, China
e-mails: hmwang@nssc.ac.cn (corresponding author), wym@nssc.ac.cn, fujianguo@nssc.ac.cn

[2]University of Chinese Academy of Sciences, Beijing, China

Abstract

A tiny and low-cost ground-based Fabry–Perot interferometer (FPI) was designed using a filter behind etalon and Galilean telescope system for the thermospheric wind observation with OI 630.0 nm nightglow emissions (~250 km). Based on the instrument, experiments were carried out at Langfang (39.40° N, 116.65° E) site for a rough comparison and Kelan (38.71° N, 111.58° E) site for a detailed validation. Wind results of Langfang experiment are well consistent with measurements of two other FPIs deployed at Xinglong (40.40° N, 117.59° E) and Kelan which are retrieved by the American National Center for Atmospheric Research (A-NCAR). In Kelan experiment, the averaged wind deviation between our FPI and A-NCAR FPI is 11.8 m/s. The averaged deviation of wind measurement error between them is 2.9 m/s. The comparisons suggest good agreement. Then, the analysis of influencing factors was made. The center determination offset has an exponential relation with wind deviation, while the radius calculation offset is linear with wind deviation.

Key words: Fabry–Perot interferometer (FPI), postpositional filter, thermospheric wind, 630.0 nm nightglow.

Ownership: Institute of Geophysics, Polish Academy of Sciences;

1. INTRODUCTION

The observations of thermosphere neutral winds are used for studying the dynamics and behavior of the thermospheric atmosphere and for developing forecasting capabilities for the space environment. The importance of measuring thermospheric winds is widely recognized in many research organizations' strategy, such as the latest NASA Heliophysics Roadmap (NASA 2009). However, the only technique that can probe thermospheric winds (~250 km) from the ground is the passive optical technique, among which the Fabry–Perot interferometer is currently the leading instrument available for making wind measurements in this region.

To date, wind measurements using Fabry–Perot interferometers from the ground have been made for several decades by many scientists. Various investigations and improvements have been reported for FPIs, for example, studies featuring long-term measurements (Hernandez and Roble 1995, Biondi *et al.* 1999), use of cooled CCD detectors (Biondi *et al.* 1995, Shiokawa *et al.* 2001, 2003), two-dimensional imaging capability (Rees *et al.* 1984, Niciejewski *et al.* 1994, Nakajima *et al.* 1995, Ishii *et al.* 1997, Conde *et al.* 2001, Sakanoi *et al.* 2009, Kosch *et al.* 2010), and daytime measurement (Gerrard and Meriwether 2011, Wu *et al.* 2012). Most of these instruments are designed using large-aperture etalons and filters. Some FPIs have been reported using small-aperture etalons to lower the costs (Makela *et al.* 2009, 2011; Shiokawa *et al.* 2012).

In this paper, a new tiny and low-cost FPI instrument with a smaller post-filter is proposed for thermospheric wind retrieval using 630.0 nm airglow emission. Based on the instrument, wind measurements were made for several days in Langfang and Kelan. Therefore, the detailed information of the instrument is given in Section 2. The methodology for wind retrieval is presented in Section 3. In Section 4, the observed data and retrieval results are shown. Subsequently, sensitivity studies for wind retrieval of our FPI are presented in Section 5.

2. INSTRUMENTATION

Figure 1 shows the schematic diagram of prime optical system (a) and configuration (b) of our FPI. The sky scanner located above the optical system consists of two motors and two 45° parallel mirrors for pointing the FPI optical axis toward any point of sky. The incident light passes through an etalon (fixed-gap spacing: 15 mm, diameter: 130 mm, reflectivity: 76%), a Galilean telescope system, a band-pass filter (diameter: 70 mm), a focus lens, and arrives at a 1024×1024 CCD detector (size: 13 mm).

The instrument is mainly composed of four parts: (1) The front optics sky-scanner constructed with two mirrors has a field-of-view of 2.25°, in

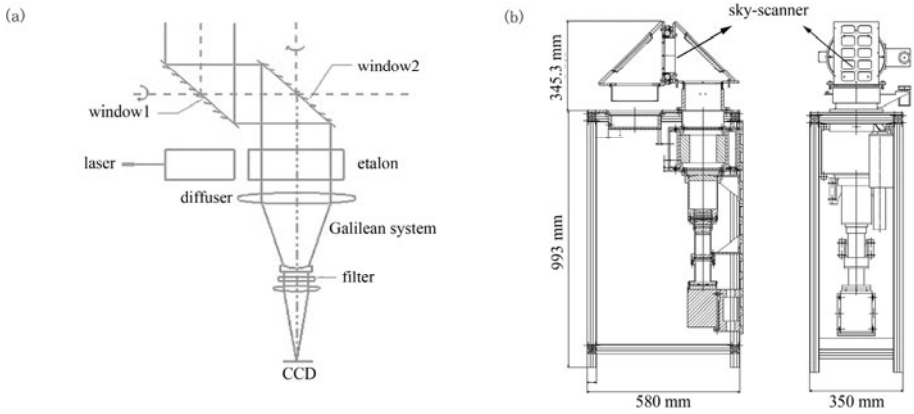


Fig. 1. The prime optical system with the filter behind etalon (a) and the corresponding configuration (b) of the Fabry–Perot (etalon fixed-gap spacing: 15 mm, etalon diameter: 130 mm, etalon reflectivity: 76%, filter diameter: 70 mm, CCD detector size: 13 mm, FPI size: $1.34 \times 0.58 \times 0.35$ m^3).

which airglow emission can be assumed uniform. The diameter of the two mirrors is 10 cm. One of the mirrors rotates to change the zenith, and the other is to change the azimuth. A cylinder hood and several stops are used in the first lens tube to prevent emissions out of field-of-view from entering instrument; (2) The etalon has a diameter of 130 mm with clear aperture of 100 mm. It is mounted in a thermally isolated enclosure that is stabilized within 0.1 °C at 30 °C for minimizing thermal drift; (3) The band-pass filter is fixed behind the etalon and Galilean telescope system to reduce the diameter to 70 mm. This design is for saving the fabricating cost; (4) Finally, approximately seven full interference fringes are imaged using a focus lens (f = 33.4 cm) onto the CCD chip. The CCD is with low readout (~3 e/pixel) and dark current (~0.0004 e/sec/pixel at a temperature of −70 °C). That allows for a long integration (~5 min) for analysis of the 630.0 nm emission. SParameters of other typical FPIs similar to our instrument are listed in Table 1.

All of these parts except for two sky-scanning mirrors are installed assembly in a box of $0.58 \times 0.35 \times 0.993$ m^3. Therefore, the tiny and compact instrument can be transported conveniently from one place to another using a car for wind observations.

The ground-based FPI measures the Doppler shift of the airglow emissions in a zenith direction and four azimuthal directions of eastward (E), westward (W), southward (S), and northward (N) with a zenith angle of 45° (Fig. 2). The FPI sky-scanner is sequentially steered to the five directions (Zenith, N, S, E, and W) to observe the OI 630.0 nm airglow emission with

Table 1

The parameters of optical components of typical FPI

FPI Institute	Our instruments National Space Science Center, China	A-NCAR National Center for Atmospheric Research, USA	J-NU Nagoya University, Japan
Filter diameter/mm	70	–	120/118.5
Etalon diameter/mm	130	132	116
Etalon fix-gap/mm	15	15/20	15
Etalon reflectivity	76%	80%	85%
CCD size/μm	13.312	13.312	13.312
CCD pixel number	1024 × 1024	1024 × 1024	1024 × 1024

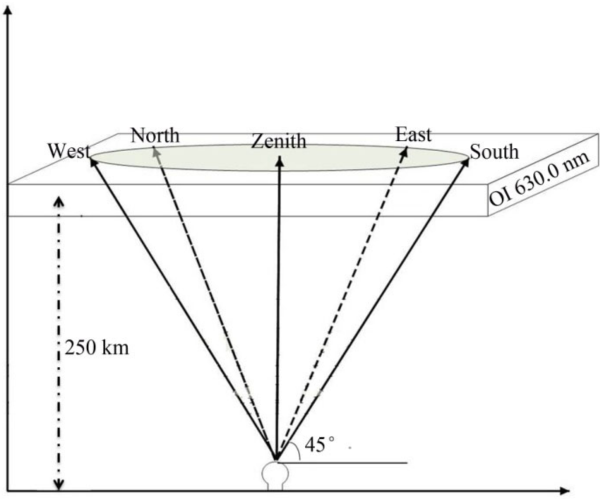


Fig. 2. The diagram of the ground-based Fabry–Perot observation with OI 630.0 nm airglow at ~250 km.

an exposure time of five minutes in each direction and ~ 30 min for a cycle. At the beginning of night observation, the calibration is utilized with the sky-scanner pointed towards a diffuser box uniformly illuminated by a HeNe frequency-stabilized laser. For acquirement of stable laser, calibrations are made five times sequentially.

3. METHODOLOGY

As the equal inclination interference formed by FPI, the interference order can be derived from the radius of fringe peak after series of transformations (Shiokawa *et al.* 2001):

$$m_i = \frac{2\mu t}{\lambda_0\left(1-\frac{v}{c}\sin\theta\right)}\left(1-\frac{r_i^2}{2f^2}\right), \tag{1}$$

where m is the interference order, subscript i presents the number of fringe from the center to the edge, μ is the refractive index, t is the etalon spacing, λ_0 is the central wavelength of the airglow without the Doppler shift, r is the fringe peak radius in one observation direction, f is the focal length, c is the light speed, $\theta = 45°$ is the view zenith angle, and v is horizontal wind velocity. Assuming that the vertical wind velocity was negligible and that the wind velocity was uniform between the two opposite directions (distance: ~500 km) holding in all areas except for high-latitude auroral zones (Smith 1998), zonal and meridional wind can be derived from the radiuses of eastward/westward and northward/southward fringe peaks, respectively. The equations are as follows:

$$v_N = \frac{c}{\sin\theta}\frac{r_S^2 - r_N^2}{4f^2-\left(r_S^2+r_N^2\right)}, \tag{2}$$

$$v_E = \frac{c}{\sin\theta}\frac{r_W^2 - r_E^2}{4f^2-\left(r_E^2+r_W^2\right)}, \tag{3}$$

where r_N and r_S are the radiuses of the northward and southward fringe peaks, respectively. r_E and r_W are the radiuses of the eastward and westward fringe peaks, respectively.

Based on Eqs. 2 and 3, wind velocity can be derived from the fringe radius. The details of procedure of data processing are provided in the literature (Wang and Wang 2015) and are briefly outlined here. First, we reduce noise using several filters to smooth out noise counts and neutralize background trend. Then, the fringe center is determined from the laser calibration image by fitting of a Gaussian function in horizontal and vertical cross sections (Kubota 1996). With the center determination, the fringe peak radius is determined using a Gaussian fitting based on annular-summed fringes. Finally, wind velocities are calculated from peak radius using Eqs. 2 and 3.

4. OBSERVATION AND RESULTS

Three observation sites in Xinglong, Langfang, and Kelan are used here for the validation of our FPI. The Langfang station is located to the east of Kelan with a linear distance of ~440 km and to the south of Xinglong with a linear distance of ~140 km, so it is applicable to make a rough comparison of wind retrieval between them. Besides, measurements of two A-NCAR FPIs deployed at Xinglong and Kelan sites are considered as reference data.

Therefore, two experiments were carried out for three days in Langfang in the end of September and for ten days in Kelan in the end of October in 2014.

Hampered by weather conditions, wind measurements of a clear night (24 September 2014) at Langfang site are compared with wind velocities of A-NCAR FPIs. The comparisons are shown in Fig. 3 that shows the wind velocity *versus* universal coordinated time (UTC) at Langfang, Xinglong, and Kelan sites. It suggests that our measurements at Langfang site are well consistent with wind variations at Xinglong and Kelan sites. However, due to weaker airglow emission at midnight, the retrieval errors become much larger with the largest values of 14 m/s (Langfang), 34 m/s (Kelan), and 18 m/s (Xinglong) at ~16:30 UTC (Fig. 3, error bar), and wind deviations between the three stations become larger during a period of 15:00~ 18:30 UTC. In Figure 3, it should be noted that the wind retrieval error of FPI at Langfang site is calculated from both the Gaussian fitting error of each fringe and the standard error of wind velocities (Wang and Wang 2015). It should be also noted that wind velocities are retrieved using 1st-7th fringes.

For further validation, our FPI was subsequently deployed in Kelan on 16-26 October 2014 for a robust comparison with A-NCAR FPI. Due to the cold weather and the tiny shabby house, the temperature of our FPI is a little difficult to be controlled. Besides, there were some cloudy or rainy days during our experiment. Therefore, wind measurements of five nights are obtained.

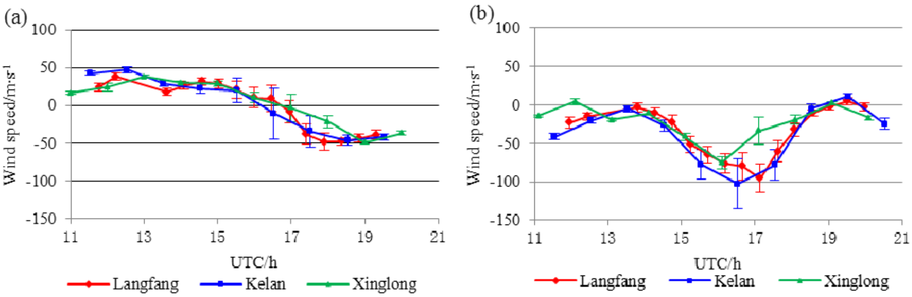


Fig. 3. Zonal (a) and meridional (b) wind retrieval of three FPI instruments on 24 September 2014, as a function of universal time (UTC). The diamond represents wind from our FPI at Langfang site. The square denotes wind results of A-NCAR FPI instrument located at Kelan site. The triangle denotes wind results of A-NCAR FPI instrument located at Xinglong site. The error bar represents twice of the standard error (1σ), which demonstrates wind retrieval uncertainty and signal-to-noise variation.

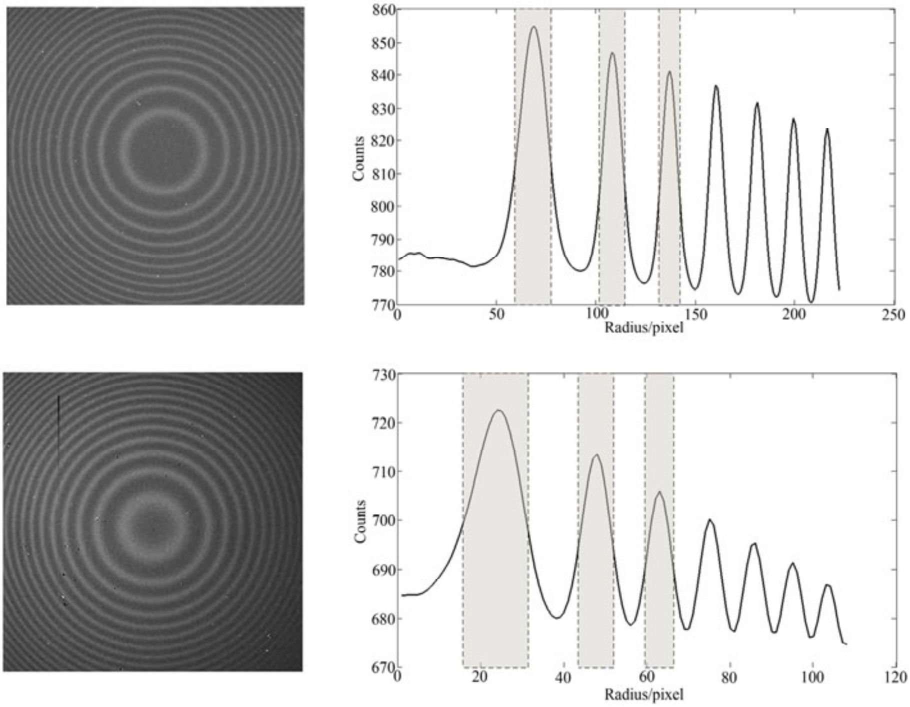


Fig. 4. The interference fringe (left) and annular-summation (right) of the two instruments in the same direction. The upper panel is from our FPI, and the lower panel is from A-NCAR FPI. The curves in gray dashed frame in the right panels indicate the Gaussian functions fitted to the fringes. Seven fringes were all used for the retrieval, but only three fringes used for Gaussian fitting are shown in the gray dashed frames here.

The interference fringes of the same direction of our FPI and A-NCAR FPI are compared and shown in Fig. 4 (left column). A corresponding cross-section of interference fringes is also shown in Fig. 4 (right column) that is a one-dimensional representation after annular-summation. It should be noted that 2×2 binning of the pixels is done for all the observed data to advance signal-to-noise, while the A-NCAR instrument uses 4×4 binning. Background count of our FPI (~780, Fig. 4, upper panel) is a little larger than that of the A-NCAR instrument (~680, Fig. 4, lower panel). One reason for the phenomenon is that more stray light is reflected in our tiny house. Another is that a little more background emissions from the sky out of the field-of-view enter into our FPI because of city light surrounded the Kelan station. Wind results of five days are compared in Fig. 5. The averaged deviation of wind velocities between the two FPIs is 11.8 m/s. The averaged measurement er-

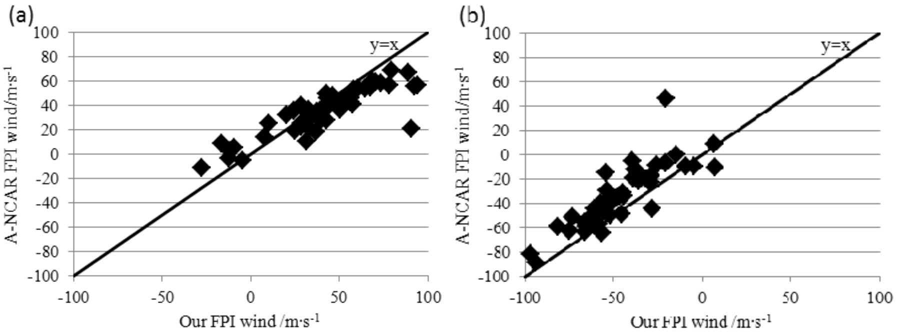


Fig. 5. Zonal (a) and meridional (b) wind of two FPI instruments in Kelan experiment. The black line (y = x) indicates values one-to-one.

ror of our instrument is 16.4 m/s, and the averaged measurement error of A-NCAR FPI is 13.5 m/s. It means that the averaged deviation of wind measurement error is 2.9 m/s between the two instruments. The comparisons show good agreement except for several large-deviation points which is due to the disturbance of cloud and stray lights (*e.g.*, city-light or lamp-light) in field-of-view.

Figure 6 shows a further comparison of measurements of one night (25 October 2014). It suggests good agreement between them. The average deviation of time-coincident retrievals at this night is 8.60 m/s. Additionally, the retrieval method used in this paper, different from the A-NCAR retrieval method, is also used for wind retrieval in combination with A-NCAR FPI data. These results called "A-NCAR retrieval" are shown with triangle in Fig. 6.

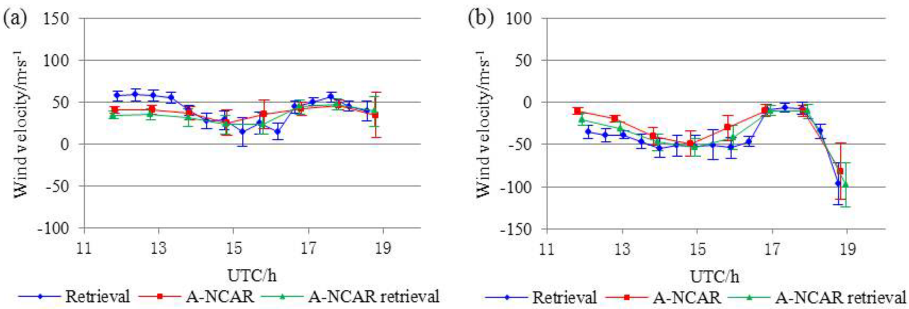


Fig. 6. Zonal (a) and meridional (b) wind retrieval of two FPI instruments as a function of universal time (UTC) on 25 October 2014. The diamond represents wind velocities from FPI. The square denotes wind products of A-NCAR FPI instrument. The triangle represents wind results of A-NCAR FPI data using our retrieval method.

5. ANALYSIS AND DISCUSSION

The precision of wind retrieval mainly depends on the airglow intensity, noise (dark noise and background), and radius fitting, *etc*. (Wang and Wang 2015). Due to the low intensity of 630.0 nm airglow, the common retrieval error of a scanning FPI is 4-30 m/s for ground-based measurements. The averaged wind measurement error of a FPI fabricated by Boston University is 15 m/s (Martinisa *et al.* 2001). The measurement error of a FPI fabricated by Japanese Nagoya University is 4-27 m/s (Shiokawa *et al.* 2012), and the measurement error of A-NCAR FPI is 6-10 m/s (Wu *et al.* 2004, Yu *et al.* 2014). The retrieval error of our FPI for 630.0 nm airglow is 4-20 m/s. Besides, the lower intensity of 630.0 nm airglow during 14:00-17:00 UTC (with large error bars) leads to larger uncertainties of wind velocities, which will enlarge the wind deviation between them. Therefore, the deviation of 11.8 m/s between our FPI and A-NCAR FPI demonstrates that the wind results of the two instruments are well consistent with each other.

The main factors (airglow intensity, instrument noise, and background emissions, *etc.*) work on precision of wind measurements according to two main parameters which are calculated during wind retrieval. One is the fringe center, and another is the fringe radius.

5.1 Center determination

When the wind velocities retrieved from two FPI instruments in Kelan (our FPI and A-NCAR FPI) are nearly equal with each other, the averaged wind is considered as the real wind. The offset of the center corresponding to real wind is set to zero. Then, center offsets in 8 azimuthal directions (45°-360°, step = 45°) are carried out, and corresponding wind deviation is calculated simultaneously for each center offset. Based on the method, we obtained several groups of wind deviation *versus* center offset. The averaged result is shown in Fig. 7a. The relationship between them is an exponential curve which shows that 2 pixels center offset can lead to 4.3 m/s wind deviation and that 6 pixels offset causes a much larger wind deviation: 114 m/s. When the center offset is 7 pixels, wind deviation is up to several thousand meters per second. Therefore, for our FPI instrument and retrieval method, the offset of center determination should be within 2 pixels which can ensure enough accuracy for wind retrieval of thermospheric atmosphere.

5.2 Radius calculation

With the same analysis method as the influence of center offset on wind deviation, the influence of radius offset on the wind retrieval is made (Fig. 7b). The relationship between radius offset and wind deviation is linear, with a

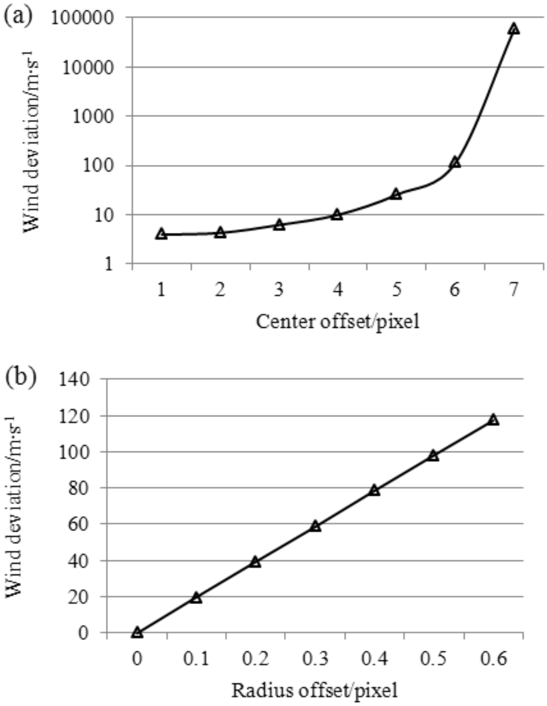


Fig. 7. The wind deviation caused by offsets of center determination (a) and radius calculation (b). The step of center offset is 1 pixel, and the step of radius offset is 0.1 pixels.

slope 19.7, which suggests that radius offsets of 0.1 pixels will cause a wind deviation of 19.7 m/s.

6. CONCLUSION

We successfully performed measurements of airglow with a tiny ground-based FPI with a band-pass filter behind etalon and Galilean telescope system. The main purpose of these experiments was to provide an independent ground-based validation of our FPI including FPI operation, detection capability, data processing method, and the retrieval precision. The observations are performed at Langfang site for a rough validation and subsequently at Kelan site for a detailed comparison with A-NCAR FPI measurements. At Langfang site, the variable trends of wind measurement of our FPI are consistent with that of A-NCAR FPI. At Kelan site, the averaged deviation of wind measurements between our FPI and A-NCAR FPI is 11.8 m/s, and the averaged deviation of wind measurement error between the two instruments is 2.9 m/s. The comparisons of the inferred zonal and meridional wind velocities show good agreement with each other.

Based on experiments, influencing factors of wind retrieval of our FPI are analyzed. One is center determination, and another is radius calculation.

The offset of the former one has an exponential relation with wind deviation. The center offset of 2 pixels can lead to 4.3 m/s wind deviation, while the center offset of 7 pixels will cause a much larger wind deviation which is up to several thousand meters per second. The offset of the latter one is linear with deviation of wind retrieval. The radius offsets of 0.1 pixels can cause a wind deviation of 19.7 m/s. Therefore, the precision of wind retrieval is much more sensitive to radius calculation than center determination.

Acknowledgments. Work about the wind observation experiment was supported by Langfang and Kelan observation stations. Thank to them for supplying places and observation conditions. Work about the comparison of FPI wind retrieval was supported by National Satellite Meteorological Center and State Key Laboratory of Space Weather (China). The authors would like to thank all of them for their outstanding support.

References

Biondi, M.A., D.P. Sipler, M.E. Zipf, and J.L. Baumgardner (1995), All-sky Doppler interferometer for thermospheric dynamics studies, *Appl. Opt.* **34**, 10, 1646-1654, DOI: 10.1364/AO.34.001646.

Biondi, M.A., S.Y. Sazykin, B.G. Fejer, J.W. Meriwether, and C.G. Fesen (1999), Equatorial and low latitude thermospheric winds: Measured quiet time variations with season and solar flux from 1980 to 1990, *J. Geophys. Res.* **104**, A8, 17091-17106, DOI: 10.1029/1999JA900174.

Conde, M., J.D. Craven, T. Immel, E. Hoch, H. Stenbaek-Nielsen, T. Hallinan, R.W. Smith, J. Olson, Wei Sun, L.A. Frank, and J. Sigwarth (2001), Assimilated observations of thermospheric winds, the aurora, and ionospheric currents over Alaska, *J. Geophys. Res.* **106**, A6, 10493-10508, DOI: 10.1029/2000JA000135.

Gerrard, A.J., and J.W. Meriwether (2011), Initial daytime and nighttime SOFDI observations of thermospheric winds from Fabry–Perot Doppler shift measurements of the 630-nm OI line-shape profile, *Ann. Geophys.* **29**, 9, 1529-1536, DOI: 10.5194/angeo-29-1529-2011.

Hernandez, G., and R.G. Roble (1995), Thermospheric nighttime neutral temperature and winds over Fritz Peak Observatory: Observed and calculated solar cycle variation, *J. Geophys. Res.* **100**, A8, 14647-14659, DOI: 10.1029/95JA00565.

Ishii, M., S. Okano, E. Sagawa, S. Watari, H. Mori, I. Iwamoto, and Y. Murayama (1997), Development of Fabry–Perot interferometers for airglow observations, *Proc. NIPR Symp. Upper Atmos. Phys.* **10**, 97-108.

Kosch, M.J., C. Anderson, R.A. Makarevich, B.A. Carter, R.A.D. Fiori, M. Conde, P.L. Dyson, and T. Davies (2010), First E region observations of mesoscale neutral wind interaction with auroral arcs, *J. Geophys. Res.* **115**, A2, A02303, DOI: 10.1029/2009JA014697.

Kubota, M. (1996), A study on middle-scale variations of thermospheric neutral winds associated with auroral activity over Syowa Station, Antarcica, Tohoku University, Japan.

Makela, J.J., J.W. Meriwether, J.P. Lima, E.S. Miller, and S.J. Armstrong (2009), The remote equatorial nighttime observatory of ionospheric regions project and the International Heliospherical Year, *Earth Moon Planet* **104**, 1, 211-226, DOI: 10.1007/s11038-008-9289-0.

Makela, J.J., J.W. Meriwether, Y. Huang, and P.J. Sherwood (2011), Simulation and analysis of a multi-order imaging Fabry–Perot interferometer for the study of thermospheric winds and temperatures, *Appl. Opt.* **50**, 22, 4403-4416, DOI: 10.1364/AO.50.004403.

Martinisa, C., J. Meriwether, R. Niciejewski, M. Biondi, C. Fesen, and M. Mendillo (2001), Zonal neutral winds at equatorial and low latitudes, *J. Atmos. Solar-Terr. Phys.* **63**, 14, 1559-1569, DOI: 10.1016/S1364-6826(01)00022-0.

Meriwether, J.W., J.J. Makela, Y. Huang, D.J. Fisher, R.A. Buriti, A.F. Medeiros, and H. Takahashi (2011), Climatology of the nighttime equatorial thermospheric winds and temperatures over Brazil near solar minimum, *J. Geophys. Res.* **116**, A4, A04322, DOI: 10.1029/2011JA016477.

Nakajima, H., S. Okano, H. Fukunishi, and T. Ono (1995), Observations of thermospheric wind velocities and temperatures by the use of a Fabry–Perot Doppler imaging system at Syowa Station, Antarctica, *Appl. Opt.* **34**, 36, 8382-8395, DOI: 10.1364/AO.34.008382.

NASA (2009), Heliophysics: the solar and space physics of a new era, Heliophysics Roadmap Team, NASA Advisory Council.

Niciejewski, R., T.L. Killeen, and M. Turnbull (1994), Ground-based Fabry–Perot interferometry of the terrestrial nightglow with a bare charge-coupled device: remote field site deployment, *Opt. Eng.* **33**, 2, 457-465, DOI: 10.1117/12.155931.

Rees, D., A.H. Greenaway, R. Gordon, I. McWhirter, P.J. Charleton, and A. Steen (1984), The Doppler imaging system: Initial observations of the auroral thermosphere, *Planet. Space Sci.* **32**, 3, 273-285, DOI: 10.1016/0032-0633(84)90163-6.

Sakanoi, T., H. Fukunishi, K. Igarashi, S. Okano, and N. Nishitani (2009), Neutral-ion interaction in the auroral E region obtained from coordinated Fabry–Perot imager and VHF radar observations, *J. Geophys. Res.* **114**, A9, A09305, DOI: 10.1029/2008JA013956.

Shiokawa, K., T. Kadota, M. K. Ejiri, Y. Otsuka, Y. Katoh, M. Satoh, and T. Ogawa (2001), Three-channel imaging Fabry–Perot interferometer for measurement of mid-latitude airglow, *Appl. Opt.* **40**, 24, 4286-4296, DOI: 10.1364/AO.40.004286.

Shiokawa, K., T. Kadota, Y Otsuka, T. Ogawa, T. Nakamura, and S. Fukao (2003), A two-channel Fabry–Perot interferometer with thermoelectric-cooled CCD detectors for neutral wind measurement in the upper atmosphere, *Earth Planets Space* **55**, 5, 271-275, DOI: 10.1186/BF03351759.

Shiokawa, K., Y. Otsuka, S. Oyama, S. Nozawa, M. Satoh, Y. Katoh, Y. Hamaguchi, and Y. Yamamoto (2012), Development of low-cost sky-scanning Fabry–Perot interferometer for airglow and auroral stydies, *Earth Planet. Space* **64**, 11, 1033-1046, DOI: 10.5047/eps.2012.05.004.

Smith, R.W. (1998), Vertical winds: A tutorial, *J. Atmos. Terr. Phys*. **60**, 14, 1425-1434, DOI: 10.1016/S1364-6826(98)00058-3.

Wang, H., and Y. Wang (2015), Error calculation and analysis for an improved wind retrieval method based on the ground-based Fabry–Perot interferometer measurements, *Adv. Space Res*. **56**, 9, 1815-1821, DOI: 10.1016/j.asr.2015.03.010.

Wu, Q., R.D. Gablehouse, S.C. Solomon, T.L. Killeen, and C.-Y. She (2004), A new Fabry–Perot interferometer for upper atmosphere research, *Proc. SPIE* **5660**, 218-227, DOI: 10.1117/12.573084.

Wu, Q., W. Wang, and R.G. Roble, I. Häggström, and A. Strømme (2012), First daytime thermospheric wind observation from a balloon-borne Fabry–Perot interferometer over Kiruna (68N), *Geophys. Res. Lett.* **39**, 14, L14104, DOI: 10.1029/2012GL052533.

Yu, T., C. Huang, G. Zhao, T. Mao, Y. Wang, Z. Zeng, J. Wang, and C. Xia (2014), A preliminary study of thermosphere and mesosphere wind observed by Fabry–Perot over Center China, *J. Geophys. Res.* **119**, 6, 4981-4997, DOI: 10.1002/2013JA019492.

Received 29 March 2015
Received in revised form 23 October 2015
Accepted 21 March 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2761-2780
DOI: 10.1515/acgeo-2016-0115

Comparison of Selected Geopotential Models in Terms of the GOCE Orbit Determination Using Simulated GPS Observations

Andrzej BOBOJĆ

University of Warmia and Mazury in Olsztyn, Institute of Geodesy, Olsztyn, Poland; e-mail: altair@uwm.edu.pl

Abstract

This work contains a comparative study of the performance of six geopotential models in an orbit estimation process of the satellite of the Gravity Field and Steady-State Ocean Circulation Explorer (GOCE) mission. For testing, such models as ULUX_CHAMP2013S, ITG-GRACE 2010S, EIGEN-51C, EIGEN5S, EGM2008, EGM96, were adopted. Different sets of pseudo-range simulations along reference GOCE satellite orbital arcs were obtained using real orbits of the Global Positioning System satellites. These sets were the basic observation data used in the adjustment. The centimeter-accuracy Precise Science Orbit (PSO) for the GOCE satellite provided by the European Space Agency (ESA) was adopted as the GOCE reference orbit. Comparing various variants of the orbital solutions, the relative accuracy of geopotential models in an orbital aspect is determined. Full geopotential models were used in the adjustment process. The solutions were also determined taking into account truncated geopotential models. In such case, an accuracy of the solutions was slightly enhanced. Different arc lengths were taken for the computation.

Key words: geopotential models, GOCE satellite orbit.

1. INTRODUCTION

Over the past few decades, numerous models describing the gravity field have been determined. Some of them were based on both terrestrial and satellite data, as, for example, EGM96 (Lemoine *et al.* 1998) or EGM2008 (Pavlis *et al.* 2012). Others have been obtained using solely satellite data, for example GOGRA02S (Yi 2012) or JYY_GOCE02S (Yi *et al.* 2013). The high quality of gravity field models is essential for modeling, for example, satellite orbits and geoid. The accuracy of these models can be described by internal and external quality parameters. Many models contain quality information expressed by variance-covariance matrices connected with a least squares solution or a Monte-Carlo approach (Gruber *et al.* 2011). However, this internal error characteristic requires verification by obtaining external quality parameters to estimate the performance of these models in several aspects. Generally, an evaluation of gravity field models may refer to two types of tests. The first one concerns satellite orbit determination, where RMS of observation residuals and orbit predictions are used for gravity model evaluations (Sośnica 2014). The second test type is connected with comparisons of geoid and derivative quantities (Gruber *et al.* 2011). Both test types are complementary because the first one investigates the quality of long wavelength part of gravity field models (in the orbit determination procedure), whereas the second one allows a model performance to be assessed in the spatial domain (in geoid comparisons) (Gruber *et al.* 2011).

Many different works refer to considering the issue of a gravity model evaluation. For example, Lejba *et al.* (2007) and Sośnica *et al.* (2012) compared the impact of selected gravity field models on the estimation of Laser Geodynamics Satellites (LAGEOS) orbits, taking into account the RMS values of satellite laser ranging (SLR) residuals. Sośnica (2014) also presents the results of different gravity model validation, where the estimated LAGEOS satellite orbits were compared with the predicted orbits. The obtained orbits estimated using the tested gravity models were also compared directly with each other. In this work, the strong dependence of orbital solutions obtained on the quality of the C_{20} coefficient for the tested gravity models is emphasized. This can also be observed in the estimated sine term of the once-per-revolution cross track acceleration (Sośnica 2014). In turn, Gruber *et al.* (2011) evaluated three geopotential models derived from the Gravity Field and Steady State Ocean Circulation Explorer (GOCE) mission taking into account orbital residuals for the very precise reduced-dynamic orbit of a satellite of the Challenging Mini-satellite Payload mission (CHAMP) (Reigber *et al.* 2005) and of the satellites of the Gravity Recovery and Climate Experiment (GRACE) mission (Tapley *et al.* 2004). These orbits were converted into the inertial Earth-centered reference frame. The

Cartesian coordinates X, Y, Z for these orbits were then treated as observations in a fully dynamic orbit determination (Gruber *et al.* 2011). On the other hand, a comparison of geoid heights computed from validated models with corresponding values obtained from GPS-leveling points was performed in an area of medium-to-higher spatial scales of the gravity field of the Earth (Gruber *et al.* 2011). The different truncation of gravity models allowed to recognize to which degree the models give significant results. This can be described equivalently as the degree from which the tested models start to lose signal. It is caused by the attenuation of a gravity signal with satellite height (Gruber *et al.* 2011).

In a similar way, Förste *et al.* (2014) tested the combined model EIGEN-6C4p. They investigated the performance of this model by fitting estimated orbital arcs of the GOCE satellite into the positions originating from the precise science orbit (PSO orbit) (Bock *et al.* 2011). These positions were treated as observations (Förste *et al.* 2014). Papanikolaou and Tsoulis (2014) validated GOCE gravity models using an adapted dynamic orbit determination algorithm. They compared selected models of the gravity field by the band-limited performance in the orbit determination and by generated orbit perturbations. Cheng and Ries (2015) compared the performance of selected GOCE gravity field models for satellite orbit determination based on SLR observations. Comparisons showed a similar performance of all recent GOCE and GRACE based models in terms of the RMS fit of the SLR observations. They also showed that the estimate of C_{20} is a dominant factor in the long-wavelength error of gravity field models (Cheng and Ries 2015).

GOCE-derived geopotential models were also locally tested in certain areas, *e.g.*, in Norway (Šprlák *et al.* 2015), Germany (Voigt and Denker 2015), in the Mediterranean area (Carrion *et al.* 2015) and in South America (de Matos *et al.* 2015). Usually, such quantities as gravity anomalies and height anomalies computed based on tested gravity field models are compared with corresponding quantities obtained independently using local terrestrial data. Hirt *et al.* (2015) compared gravity from the GOCE mission and from RET2014 topography. The degree of similarity between both signals was taken as an indicator of quality for the GOCE gravity field models. It was found that the 5th-generation GOCE models describe parts of the gravity field down to about 70 km spatial scales (Hirt *et al.* 2015).

The GOCE satellite was a key component of the aforementioned the Gravity Field and Steady State Ocean Circulation Explorer mission, which was the first Earth explorer core mission of the European Space Agency (ESA). This satellite was launched on 17 March 2009 into a sun-synchronous dusk-dawn orbit with a very low initial altitude of about 280 km (Bock *et al.* 2011). The core onboard three-axis gradiometer performed measurements of gravity gradients whereas the 12-channel dual-frequency Global

Positioning System (GPS) receiver delivered phase-code observations of GPS satellites. The extremely low altitude of this satellite was necessary to ensure a proper level of gravity signal. Such altitude was realized by a drag-free flight, which was maintained using the drag-free and attitude control system (DFACS). This means that mainly the atmospheric drag acting on the satellite in flight direction was compensated (Bock *et al.* 2011). However, direct and indirect solar radiation pressure also occurs. The collected time series of gravity gradient observations and GPS code-phase measurements were used for the realization of the main objective of the mission which was the estimation of the new generation of the Earth's static gravity field models (Pail *et al.* 2011). Code-phase GPS measurements were also used to estimate the GOCE satellite precise orbit as a reduced-dynamic orbit and a kinematic orbit. The precise orbit was needed for the geolocation of observations derived from the mission (Bock *et al.* 2011). The satellite positions taken from the kinematic orbit were used for the recovery of both the static and time-variable gravity fields (Jäggi *et al.* 2015). This also illustrates the sensitivity of the very low orbit of GOCE satellite to temporal variations of the gravity field. The GOCE mission ended with the re-entry of the satellite in dense layers of the atmosphere 13 November 2013 (ESA 2014).

This work contains various tests, in which the performance of six selected gravity field models in the process of determining the GOCE satellite orbit was compared. Since an average operational altitude of GOCE satellite was equal to about 255 km (Rummel *et al.* 2009), a comparison of the performance of selected gravity models therefore refers to the extremely low Earth orbit.

The aim of this work was to compare the quality of long-wavelength parts of selected gravity field models in the aspect of GOCE satellite orbit determination with an indication of the preferred models. The obtained results may be useful, for example, for a fully dynamic approach (Casotto *et al.* 2013) for the GOCE orbit determination by using the preferred gravity field models.

The tested models were based on the data coming from terrestrial sources and such space missions as CHAMP and GRACE. GOCE-derived gravity models were not included in this study. This is caused by the fact that the same data sets were partially used for the estimation of these models and the GOCE orbit.

2. RESEARCH

The basic tool used in this work is a software package called the Orbital Computation System (OCS), which is an extension of the Toruń Orbit Processor (TOP) package (Drożyner 1995). An important task realized by the OCS package is to determine a satellite orbit in the field of gravitational and non-gravitational perturbing forces. Taking into account the Cowell 8th order method, the equation motion of a satellite is numerically integrated in order to obtain a time series of position and velocity vectors. The right-hand side of the equation of motion contains the vector presenting the satellite's Keplerian motion in the central gravity field and the vector describing the effect of perturbing forces (Eshagh and Najafi-Alamdari 2007). In the framework of the OCS software, a mathematical model of the forces governing the satellite motion is created at the given epoch during the computation (Drożyner 1995). This model, for the GOCE satellite, includes the gravitational accelerations generated by: the geopotential (a given tested gravity field model), ocean tides and Earth tides, the third body effect and the relativity effects. The ocean and Earth tides were described by the MERIT (Monitoring Earth Rotation and Intercomparison of Techniques) standards model (Melbourne *et al.* 1983), whereas the acceleration due to the third body effect was computed using the planetary ephemerides DE200/LE200 (Standish *et al.* 1992). The relativistic acceleration was computed by means of the Painleve formulation, taking into account spherical symmetrical space-time with the Schwarzschild metric. This formulation was implemented in the Toruń Orbit Processor software (Drożyner 1995). Mathematical formulas describing the MERIT standards model and the relativistic acceleration are presented in detail by Bobojć and Drożyner (2011). The satellite orbit computation is a part of the orbit estimation process, which is the main task realized by the OCS package.

The simulated observations of pseudo-ranges between the GOCE satellite and maximum twelve Global Positioning System (GPS) satellites at a given epoch are used in the orbit determination. These simulations were obtained taking into account the reduced-dynamic PSO orbit of GOCE satellite and the orbits of GPS satellites provided by the ESA as an L2 product of the GOCE mission (ESA 2010, Bock *et al.* 2011). The computation of sets of the pseudo-ranges was based on the Cartesian coordinates of GOCE and of GPS satellites with respect to the inertial reference frame (IRF) of standard epoch J2000.0 (ESA 2010). Additionally, the time of GPS signal travel between the GPS satellites and the GOCE satellite was taken into account in this computation. In order to express of the GOCE and GPS satellite coordinates with respect to IRF, the elements of orientation of ITRF2005 and IRF (ESA 2010) were used. These elements in terms of quaternions were obtained through the ESA and they were also used in the orbit determination.

The orbit estimation process is based on the following observation equation:

$$D_{jk}^{o}+v_{jk}=D_{jk}^{c}+\frac{\partial D_{jk}^{c}}{\partial \boldsymbol{r}}\frac{\partial \boldsymbol{r}}{\partial(\boldsymbol{r}_o,\dot{\boldsymbol{r}}_o)}\left[\Delta\boldsymbol{r}_o,\Delta\dot{\boldsymbol{r}}_o\right]^T. \tag{1}$$

In the above formula:

D_{jk}^{o}, D_{jk}^{c} – observed and computed pseudo-range between the GOCE satellite and the j-th GPS satellite at epoch k, respectively,

v_{jk} – correction to the observed pseudo-range D_{jk}^{o},

$\partial D_{jk}^{c}/\partial \boldsymbol{r}$ – partial derivative of the computed pseudo-range D_{jk}^{c} with respect to the position vector $\boldsymbol{r}=[x,y,z]^T$ at epoch k,

$\boldsymbol{r}_o,\dot{\boldsymbol{r}}_o$ – initial position vector and initial velocity vector, respectively, at initial epoch t_o, $\boldsymbol{r}_o=[x_o,y_o,z_o]^T$ and $\dot{\boldsymbol{r}}_o=\left[\dot{x}_o,\dot{y}_o,\dot{z}_o\right]^T$,

$\partial\boldsymbol{r}/\partial(\boldsymbol{r}_o,\dot{\boldsymbol{r}}_o)$ – partial derivative of the position vector with respect to the initial state vector $\boldsymbol{p}_o=\left[\boldsymbol{r}_o,\dot{\boldsymbol{r}}_o\right]^T$,

$\Delta\boldsymbol{r}_o,\Delta\dot{\boldsymbol{r}}_o$ – unknown correction vectors to the initial position and velocity vectors; $\Delta\boldsymbol{r}_o=\left[\Delta x_o,\Delta y_o,\Delta z_o\right]^T$, and $\Delta\dot{\boldsymbol{r}}_o=\left[\Delta\dot{x}_o,\Delta\dot{y}_o,\Delta\dot{z}_o\right]^T$,

In the orbit estimation process, the unknown corrections to the initial state vector are estimated in successive iterations using the classical least squares method until a convergence. Finally, the initial state vector, corrected in the last iteration, enables determining the satellite orbit.

In order to estimate the quality of obtained solutions, the root mean square (RMS) of the difference between the determined orbit and the reference one is used. This parameter expresses the accuracy of the given solution. It is determined by means of the following formula:

$$\mathrm{RMS}=\sqrt{\sum_{i=1}^{3}\left(\mathrm{RMS}_i\right)^2}. \tag{2}$$

The quantities of $\mathrm{RMS}_i\,(i=1,2,3)$ are computed using the expression:

$$\mathrm{RMS}_i=\sqrt{\frac{\sum_{j=1}^{n}\left[(x_i)_j-(x_i)_{j\mathrm{REF}}\right]^2}{n}}, \tag{3}$$

where $(x_i)_j$, $(x_i)_{j\mathrm{REF}}$ (i = 1, 2, 3; $x_1=x$, $x_2=y$, $x_3=z$) are the satellite's Cartesian coordinates at epoch j w.r.t. IRF, in the estimated orbit and in the reference orbit, respectively, and n is the total number of epochs – the same for both orbits. The RMS parameter can be explained as the mean distance between corresponding points of both orbits (for the same epoch) or as a measure of the fit of the orbit determined to the reference orbit. In the case of GOCE orbit computed directly without adjustment, the RMS value determines the threshold of the orbital solution effectiveness, *i.e.*, solutions with RMS values less than the corresponding threshold values can be treated as effectively estimated orbital variants.

The reference orbit – the reduced-dynamic PSO GOCE orbit, is acquired through ESA as a Level 2 GOCE mission product and was estimated relying on the GPS observations. The gravity field model and the remaining dynamical models were also used in the estimation process. The generation of this orbit was a multi-step process. In the first step, an approximate orbit using pseudo-range measurements was determined. In the next step, this orbit was improved in an iterative procedure using zero-difference phase observations. After an appropriate number of iterations, this orbit was determined by six initial osculating elements, three constant empirical accelerations and a set of pseudo-stochastic piecewise constant accelerations in 6-minute intervals (Bock *et al.* 2007). These pseudo-stochastic accelerations absorb the errors of dynamic models used and the accelerations induced by non-gravitational forces such as the direct solar radiation pressure and the Earth albedo. The effect of the dynamic models on this orbit is limited by estimated pseudo-stochastic parameters. Hence, this orbit is called a reduced-dynamic orbit. An approach applied to the estimation of reduced-dynamic orbit has also been successfully used to determine, for example, the satellite orbit of the CHAllenging Minisatellite Payload (CHAMP) mission with an accuracy of about 3 cm (Jäggi *et al.* 2006). In turn, the accuracy of the PSO orbit of GOCE satellite is at a level of 2 cm, which is based on the SLR validation (Bock *et al.* 2011). The precise orbit of GOCE satellite was also determined as a kinematic solution. This kinematic orbit was estimated solely based on the GPS phase measurements, *i.e.*, no dynamic models were taken into account. Thus, this geometrical solution is more sensitive to changes in the quality of GPS measurements. The accuracy of the kinematic orbit is at a similar centimeter level as in the case of a reduced-dynamic orbit (Bock *et al.* 2011).

The estimated GOCE orbits were integrated and compared with the reference orbit (the reduced-dynamic PSO orbit) with respect to IRF, whose origin is located at the center of mass of the Earth. The reference orbit was previously transformed from ITRF2005 to IRF using the instantaneous rotation matrices generated on the basis of a given set of ESA-delivered quaternions.

Ten orbital arcs were selected for the orbit estimation process in which the corrections to the corresponding initial state vectors were estimated. These initial state vectors were taken from the reduced-dynamic PSO orbit of the GOCE satellite (Bock *et al.* 2011) at the following epochs [UTC]:

6 November 2009,	23 h 59 m 45.00 s,
19 November 2009,	23 h 59 m 45.00 s,
2 December 2009,	23 h 59 m 45.00 s,
18 December 2009,	23 h 59 m 45.00 s,
29 December 2009,	23 h 59 m 45.00 s,
6 January 2010,	23 h 59 m 45.00 s,
16 January 2010,	23 h 59 m 45.00 s,
26 January 2010,	23 h 59 m 45.00 s,
5 February 2010,	23 h 59 m 45.00 s,
10 February 2010,	23 h 59 m 45.00 s.

In order to obtain the computed (approximated) orbital arcs, the same set of initial state vectors was used.

Taking into account the mentioned variants of orbital arcs and six selected geopotential models, the different solutions of orbit estimation process were determined. The geopotential models are expressed in terms of spherical harmonic coefficients according to the following formula (Heiskanen and Moritz 1967):

$$V(r,\theta,\lambda) = \frac{\mathrm{GM}}{r} \sum_{n=0}^{N_{\max}} \left(\frac{a}{r}\right)^n \sum_{m=0}^{n} \left(\bar{C}_{nm} \cos m\lambda + \bar{S}_{nm} \sin m\lambda\right) \bar{P}_{nm}(\cos\theta). \tag{4}$$

In this equation: V is the potential of the gravity field; r, θ, λ – geocentric coordinates of a given point: r – distance from Earth's center, $\theta = 90° - \varphi$ – colatitude, and φ – geocentric latitude, λ – geocentric longitude; a – equatorial radius of the Earth ellipsoid; $\bar{C}_{nm}, \bar{S}_{nm}$ – spherical harmonic coefficients (Stokes' coefficients) of degree n ($n = 0, 1, ..., N_{\max}$, $N_{\max}$ – the maximum degree of the spherical harmonic expansion) and order m ($m = 0, 1, ..., n$), and $\bar{P}_{nm}(\cos\theta)$ – normalized associated Legendre function of degree n and order m.

All geopotential models used in this work are listed in Table 1. They were obtained through International Center for Global Earth Models (ICGEM) at Deutsches GeoForschungsZentrum Potsdam (Drewes 2012). The ICGEM is one of six centers of the International Gravity Field Service of the International Association of Geodesy. The library of available models is constantly updated by the ICGEM.

All of the parameters constituting the tested models were adopted for computation without any changes. The exception was the ITG-GRACE 2010S model, where the non-zero coefficients of first degree were replaced by zero values. It was done to remove the effect of geocenter shift.

Table 1

List of the gravity field models used in this work

Gravity field model	Reference
EGM2008	Pavlis *et al.* 2012
EIGEN-5S	Förste *et al.* 2008
EIGEN-51C	Bruinsma *et al.* 2010
ITG-GRACE2010S	Mayer-Gürr *et al.* 2011
ULux_CHAMP2013S	Weigelt *et al.* 2013
EGM96	Lemoine *et al.* 1998

3. NUMERICAL TEST RESULTS

Using selected six gravity field models and the corresponding sets of pseudo-ranges, the different variants of GOCE orbital arcs were determined using the OCS software package taking into account the remaining dynamical models. The accuracies of estimated orbital arcs were expressed by the RMS difference between these arcs and the corresponding reference ones (reduced-dynamic PSO orbital arcs). The variants of estimated orbital arcs were obtained taking into account the first three initial epochs and all ten initial epochs listed in the previous section. Thus, the three and the ten solutions of the orbit determination process were estimated for each geopotential model taking into account a given orbital arc length. Subsequently, the RMS parameters computed for obtained solution variants for a given gravity field model were averaged.

Table 2 shows the mean values (for three orbital arcs) of the RMS parameter for particular geopotential models for the 1-day orbital arcs. Nominally, the reference orbit (the GOCE reduced-dynamic PSO orbit) was just obtained in the form of the 1-day arcs (Bock *et al.* 2011).

There are two basic options in Table 2. Both assume, of course, the use of the given gravity field model in the orbit determination. However, the first one additionally implies the use of the additional dynamical models with or without the observations (G-DM and G-DM-O modes in Table 2), whereas the second one uses the geopotential model only with or without the observations or additionally the truncated geopotential model (G, G-O, and G-O-Gtr modes in Table 2). In both options, the addition of observations (simulated pseudo-ranges) is connected with the orbit improvement with respect to the corresponding approximated orbit (computed without using the observations). An effectiveness threshold of improvement is defined by the $\overline{\mathrm{RMS}}$ values, which are obtained without using the measurements.

Table 2

Mean RMS (for three arcs) of differences between the estimated 1-day orbital arcs and the corresponding reference arcs (reduced-dynamic PSO arcs), depending on the applied geopotential model

Gravity field model	$\overline{\text{RMS}}$ [m]				
	G-DM[1]	G-DM-O[2]	G[3]	G-O[4]	G-O-Gtr[5]
EIGEN-51C	15.97	2.52	30.90	8.01	7.72 / ~92
EGM2008_360×360	15.76	2.72	31.68	8.08	7.79 / ~104
EIGEN-5S	15.61	2.55	31.45	8.05	7.84 / ~92
ITG-GRACE2010S	10.03	2.47	29.99	8.13	7.84 / ~92
ULux_CHAMP2013S	9.21	2.56	30.69	8.14	7.88 / ~93
EGM96	43.46	4.48	47.90	8.42	8.36 / ~124

Explanations: Data set used in the orbit determination for the following modes:
1) geopotential model (G), dynamical models (DM) – ocean tides, solid Earth tides, third body effect, relativity;
2) geopotential model, dynamical models, observations (O),
3) geopotential model,
4) geopotential model and observations,
5) geopotential model, observations and the truncation of geopotential model (Gtr) for the improved initial state vector resulting from the G-O variant.
The results are given in the form: $\overline{\text{RMS}}$ / mean degree and order of the truncation.

It is clearly seen that all solutions in Table 2 are effective because they have the $\overline{\text{RMS}}$ values after using the observations several times less than the ones obtained for the approximated orbit. As can be seen, the older model, *i.e.*, EGM96 generates clearly worse solutions (higher $\overline{\text{RMS}}$ values) than the newer ones. In the framework of the first option (using the additional dynamical models in orbit determination), the best results are achieved for such models as ITG-GRACE2010S, EIGEN-51C, EIGEN-5S, and ULux_CHAMP2013S where the corresponding $\overline{\text{RMS}}$ values are equal to 2.47, 2.52, 2.55, and 2.56 m, respectively. These values were determined taking into account the G-DM-O mode.

In order to isolate and emphasize the impact of the gravity field models and to obtain an independence of the results of the orbit determination from the specified set of dynamic models used, the second aforementioned option is selected which comprises the use of only the geopotential models and observations in the orbit improvement process. The smallest $\overline{\text{RMS}}$ values in this option are given by the application of gravity field models truncated at the same degree and order (G-O-Gtr mode). This degree and order is obtained by an analysis of the RMS values of the difference between the orbit computed for the estimated state vector (taken from the estimation based on the full geopotential model and observations) and the reference orbit (reduced-dynamic PSO orbit). Due to the changing of geopotential model truncation, the RMS values change, reaching the minimum at the specific value of degree and order of truncation. The values presented in Table 2 (last column) are the averages taking into account the RMS values and the degree and order truncation values for the three orbital arcs, which were the subject of estimation. For the degree and order of truncation, it is the mean value rounded to the nearest integer. For all tested geopotential models, the truncation to the specified degree and order causes the increase of determined arc accuracy of a few dm. This indicates the well-known fact that a usable signal for the given gravity field model in an orbital aspect primarily covers the long wavelength part. On the other hand, the obtained degree and order of the gravity models truncation may be also connected with the different ways of regularizations, which were used in the process of the model estimation. Unlike in the previous option (the use of remaining dynamical models) the best result is obtained for the EIGEN-51C model – $\overline{\mathrm{RMS}}$ equals 7.72 m, but the successive places are occupied by the geopotential models, such as EGM2008, EIGEN-5S, ITG-GRACE2010S with the $\overline{\mathrm{RMS}}$ values from 7.79 to 7.84 m (G-O-Gtr mode). These models are involved with the GRACE mission data. Slightly worse results are obtained by the use of ULux_CHAMP2013S model ($\overline{\mathrm{RMS}}$ at a level of 7.88 m) which is based on the CHAMP mission data. Clearly inferior values of $\overline{\mathrm{RMS}}$ were obtained for the older EGM96 model (G-O and G-O-Gtr mode).

Table 3 contains similar results as in Table 2. In order to significantly reduce of errors due to the disabling of the remaining dynamic models in the G-O and G-O-Gtr modes, a much shorter orbital arc, which is equal to about 90 minutes, is selected. This is approximately the period of GOCE satellite revolution. Compared to Table 2, $\overline{\mathrm{RMS}}$ values for the best solutions decreased from meters to a level of decimeters. It is caused by the aforementioned decrease of errors, which are connected mainly with the use of only geopotential model with the observations (G-O and G-O-Gtr mode) and a simplified parameterization of the orbit determination process. This parameterization only includes the estimation of six corrections to the initial state vector of satellite. The orbital errors increase with increasing length of the determined arc. As for the results in Table 2, slightly better results were achieved using the truncated geopotential models in the orbit estimation. The decrease of $\overline{\mathrm{RMS}}$, obtained as a difference between the G-O mode and the

G-O-Gtr mode in Table 3, is included in the range of 0.7 cm for the ULux_CHAMP2013S model to 10.0 cm for the EGM96 model.

Table 3

Mean RMS of differences between the estimated 90-minute orbital arcs and the corresponding reference ones (reduced-dynamic PSO orbital arcs) depending on the applied geopotential model. Mean RMS values in the columns 2, 3, 4 are obtained for the three arcs. The results in the column 5 refer to the ten arcs

Gravity field model	$\overline{\text{RMS}}$ [cm]			
	G[1]	G-O[2]	G-O-Gtr[3]	G-DM-O[4]
EGM2008_360×360	218.6	56.3	55.0 / ~ 126	14.0
EIGEN-5S	220.4	56.6	55.1 / ~ 116	13.8
EIGEN-51C	217.7	56.4	55.2 / ~ 124	13.7
ITG-GRACE2010S	210.3	58.3	57.2 / ~ 124	13.9
ULux_CHAMP2013S	216.5	59.6	58.9 / ~ 110	15.3
EGM96	564.6	123.1	113.1 / ~ 93	120.4

Explanations: Data set used in the orbit determination for the following modes:
1) gravity field model only,
2) gravity field model and observations,
3) geopotential model, observations and the truncation of geopotential model (Gtr) for the corrected initial state vector resulting from the G-O variant,
4) geopotential, dynamical models (ocean tides, solid Earth tides, third body effect, relativity), observations.
The results are given in the form: $\overline{\text{RMS}}$ / mean degree and order of the truncation.

Unlike for the longer arcs, this time the best result in the G-O-Gtr mode was obtained for the EGM2008 gravity field model ($\overline{\text{RMS}}$ = 55.0 cm – the G-O-Gtr mode). The next best solutions were obtained using the EIGEN-5S and EIGEN-51C models with the $\overline{\text{RMS}}$ value equal to 55.1 and 55.2 cm, respectively (Table 3). The use of the ITG-GRACE2010S and ULux_CHAMP 2013S generates slightly worse results (57.2 and 58.9 cm, respectively). The result for the EGM96 model (113.1 cm) clearly differs from the other ones.

The values of $\overline{\text{RMS}}$ in the G-DM-O mode (Table 3) were determined taking into account the ten orbital arcs with the initial epochs listed in the previous section. In this mode, the obtained results are about four times better than in the G-O mode due to the addition of dynamic models. The result for the EGM96 model is worse than the results for the remaining models by almost one order of magnitude. For this model, a relatively small increase in the accuracy of about 1.02 times in the G-DM-O mode w.r.t. the G-O mode occurs. The best results are achieved for such models as EIGEN-51C ($\overline{RMS}$ = 13.7 cm) and EIGEN-5S ($\overline{RMS}$ = 13.8 cm). Slightly worse result can be seen for the ULux_CHAMP2013S model ($\overline{RMS}$ = 15.3 cm).

In turn, each of the values in Table 4 is computed using ten RMS values obtained for ten 45.0-minute orbital arcs with the initial epochs listen in the Section 2. Comparing the results in the G-O mode in Tables 2-4, a similar order of the gravity models can be noted, except the EGM2008 model, which changes position from third (Table 2, $\overline{RMS}$ = 8.08 m) through the first (Table 3, $\overline{RMS}$ = 56.3 cm) to fourth (Table 4, $\overline{RMS}$ = 23.43 cm). The relative order (G-O mode for increasing values) of the remaining models is the same in all three cases (Tables 2-4): EIGEN-51C, EIGEN-5S, ITG-GRACE2010S, ULux_CHAMP2013S, EGM96.

Just as in Table 3, the best results were achieved using the G-DM-O mode, taking into account the dynamic models (Table 4). As for the 90-minute arcs, the accuracy of solutions increases about four times w.r.t. the G-O mode with the exception of solutions for the older EGM96 model, where the increase in accuracy is characteristically low – about 1.02 times. The considered results confirm the dominance of the EIGEN-5S model ($\overline{RMS}$ = 4.9 cm) and the EIGEN-51C model ($\overline{RMS}$ = 5.1 cm). On the other hand, the slightly worse results for the ULux_CHAMP2013S model and the much worse results for the EGM96 model are also confirmed.

Table 4

Mean RMS (for ten arcs) of differences between the estimated 45.0-minute orbital arcs and the corresponding reference ones (reduced-dynamic PSO orbital arcs) depending on the applied geopotential model for the G-O mode and the G-DM-O mode

Gravity field model	$\overline{RMS}$ [cm]	
	G-O[1]	G-DM-O[2]
EGM2008_360x360	23.43	5.3
EIGEN-5S	23.29	4.9
EIGEN-51C	23.13	5.1
ITG-GRACE2010S	23.34	5.2
ULux_CHAMP2013S	23.52	5.8
EGM96	44.15	41.1

Explanations: Data set used in the orbit determination for the following modes:
1) gravity field model and observations,
2) geopotential, dynamical models (ocean tides, solid Earth tides, third body effect, relativity), observations.

Direct comparison between the performance of all gravity models is given in Table 5 (the G-O mode). This comparison is done for the 45-minute arcs and the 1-day arcs. The presented results are the absolute values of differences of $\overline{\text{RMS}}$ (hereafter referred to briefly as differences) of the fit between the tested models. For the 45-minute arcs, these differences are below one centimeter in the group of newer models (EIGEN-51C, EIGEN-5S, ITG-GRACE2010S, EGM2008, ULux_CHAMP2013S) and they do not exceed 0.4 cm − the largest of them, between EIGEN51-C and ULux_CHAMP 2013S, is equal to 0.39 cm. The EGM96 model clearly differs from the others, showing above a fifty times greater difference with the EIGEN51C model (21.02 cm). In the case of longer, 1-day arcs (Table 5 – differences in bold) differences remain at the centimeter level, reaching 13 cm for EIGEN51-C and ULux_CHAMP2013S (for the newer models). This time, the greatest difference with the EGM96 model is equal to 41 cm. This value is approximately three times greater than the largest difference in the group of newer models. It follows that the disproportion between the performance of newer models and the EGM96 model decreased significantly with the increasing length of the orbital arcs (from 45-min to 1-day), because the aforementioned ratio of differences decreased from about fifty to almost three.

In order to obtain a measure of the disproportion between the performance of newer models, two other ratios were computed for the 45-minute and 1-day orbital arcs by dividing the greatest difference in the group of newer models (Table 5) by the mean value of difference in this group of models. The ratios are equal to about 2.12 and 1.91 for the 45-minute and 1-day orbital arcs, respectively, which means that the disproportion between the performance of newer models is almost the same for both the 45-minute and 1-day orbital arcs. Strictly speaking, the disproportion slightly decreased for the longer orbital arc.

Table 5

Absolute values of differences between mean RMS of fit for all compared models, taking into account the estimated 45.0-minute orbital arcs and 1-day arcs (in bold) for the G-O mode (using the gravity field model and observations)

$\lvert\Delta\overline{\text{RMS}}\rvert$ [cm]	EIGEN-51C	EIGEN-5S	ITG-GRACE 2010S	EGM2008	ULux_CHAMP 2013S	EGM96
EIGEN-51C	–	0.16 / **4**	0.21 / **12**	0.30 / **7**	0.39 / **13**	21.02 / **41**
EIGEN-5S	–	–	0.05 / **8**	0.14 / **3**	0.23 / **9**	20.86 / **37**
ITG-GRACE2010S	–	–	–	0.09 / **5**	0.18 / **1**	20.81 / **29**
EGM2008	–	–	–	–	0.09 / **6**	20.72 / **34**
ULux_CHAMP2013S	–	–	–	–	–	20.63 / **28**

4. SUMMARY AND CONCLUSIONS

Six selected gravity field models were used in the process of GOCE satellite orbit determination. These models are involved in this process directly in the satellite motion model. For the computations, the ten GOCE initial state vectors from the reference orbit (reduced-dynamic PSO orbit for the GOCE satellite) were adopted. The sets of GPS pseudo-ranges simulated along the reduced-dynamic PSO orbit were the observations used in the adjustment. Different solution variants of the GOCE orbit have been computed, taking into account the data processed in the following modes: the geopotential model only, the geopotential model and the remaining dynamical models, the geopotential and the observations, the geopotential and the remaining dynamical models with the observations.

In order to compare the quality of determined solutions using the selected geopotential models, the RMS differences between the estimated orbits and the corresponding reference orbits were averaged in the frame of each aforementioned solution mode. Thus, the $\overline{\mathrm{RMS}}$ value for the given geopotential model is based on the RMS values computed for the three and ten orbit estimations with the same arc lengths and the different initial state vectors. Three groups of $\overline{\mathrm{RMS}}$ values were obtained. They refer to the length of the estimated orbital arcs of about 45 minutes, 90 minutes, and 1 day.

For tested arc lengths, the best results (the smallest $\overline{\mathrm{RMS}}$ values) were achieved especially for the EIGEN-51C model (for 45-minute arcs – in the G-O mode and for 1-day arcs – in the G-O and G-O-Gtr mode). The results in the G-DM-O mode depend to some extent on the orbital arc length. In the case of shorter, 45- and 90-minute arcs, the best solutions were obtained for the EIGEN-51C model and the EIGEN-5S model, whereas the ITG-GRACE2010S model is preferred for the 1-day arcs. However, the EIGEN-51C model was the second preferred model for these orbital arcs. The EIGEN-51C is a combined solution based on the GRACE and CHAMP mission data and terrestrial gravimetric data (Bruinsma *et al.* 2010). Similarly, the EIGEN-5S is based on the GRACE mission data and additionally on LAGEOS data (Förste *et al.* 2008). The results obtained in the G-DM-O mode also depend on the taken set of remaining dynamical models. Although these results both confirm good performance of the aforementioned EIGEN-51C model and the EIGEN-5S model and slightly worse performance the ULux_CHAMP2013S model, based on the CHAMP-only data (Weigelt *et al.* 2013). These results show clearly worse solutions for the older EGM96 model, which incorporates gravity anomalies derived from altimeter data, surface gravity data and, among others, such data as Doppler observations, SLR and optical observations (Lemoine *et al.* 1998).

After truncation of the geopotential models at the determined degree and order, the obtained results are slightly improved, which indicates a useful long and medium wavelength part of the assessed models in terms of determination of the GOCE orbit. In other words, the obtained degree and order values of truncation may indicate a sensitivity limit of GOCE orbit in terms of modeling geopotential. The mentioned part of the assessed models reaches the degree and order from 92 to 124 for the longer, 1-day arcs and from 93 to 126 for the shorter, 90-minute arcs. It seems that, for the shorter, 90-minute arcs, the useful part of the geopotential models is generally larger than for the longer, 1-day arcs. The mean degree and order of truncation is equal to about 99 for the longer, 1-day arcs whereas the corresponding mean value for the shorter, 90-minute arcs equals about 115. This might be related to smaller errors in the process of shorter arc determination. Additionally, in the case of the shorter, 90-minute arcs, the useful part of models is larger for the gravity models with the better results – smaller $\overline{RMS}$ values (EIGEN-51C, EGM2008, EIGEN-5S, ITG-GRACE2010S) than for the models with worse results – greater mean RMS values (ULux_CHAMP2013S, EGM96). The gravity field models lost the signal above determined values of degree and order, which means that a certain range of spherical harmonic coefficients is not useful in the considered orbital aspect. On the other hand, the presented truncation effect is not only due to the signal loss, but also due to the aforementioned regularization of the gravity field solutions.

An example increase of the time range of research and an example decrease of the length of orbital arc determined does not lead to significant changes in relative results – the order of five of six tested models in terms of performance in the GOCE orbit estimation did not change. The following sequential pattern is visible for all three cases of orbital arc lengths used (in the G-O mode, Tables 2-4): EIGEN-51C, EIGEN-5S, ITG-GRACE2010S, ULux_CHAMP 2013S, EGM96.

In comparing the results for the 45-minute and 1-day orbital arcs (Table 5), a significant decrease of disproportion of the performance between the newer models (EIGEN-51C, EIGEN-5S, ITG-GRACE2010S, EGM2008, ULux_CHAMP2013S) and the EGM96 model can be observed. It seems that larger errors which arise in the process of generating longer, 1-day orbital arcs can lead to a reduction of this disproportion with respect to the disproportion for the shorter, 45-minute arcs. In the case of the disproportion of the performance between the newer models, it is almost the same for both the 45-minute and 1-day orbital arcs (only slightly decreases for the longer arcs). The determined ratios measuring this disproportion are similar for both orbital arc lengths – 2.12 for the 45-minute arcs and 1.91 for the 1-day arcs.

Taking into account all obtained results (Tables 2-4), the generated GOCE orbital solutions particularly prefer such models as: EIGEN-51C and EIGEN-5S, EGM2008, ITG-GRACE2010S (especially this model in the G-DM-O mode for the 1-day arcs). All of these models are based on the GRACE mission data.

It is worth noting that the first model in the above list, *i.e.*, EIGEN-51C, is a combined solution that uses both satellite data and (especially) terrestrial data, which may have some importance for an orbit with such an extremely low altitude as the GOCE satellite orbit.

Acknowledgments. The author would like to thank Mehdi Eshagh and the unknown Reviewer for their valuable and constructive comments that helped improve the manuscript.

References

Bobojć, A., and A. Drożyner (2011), GOCE satellite orbit in aspect of selected gravitational perturbations, *Acta Geophys.* **59**, 2, 428-452, DOI: 10.2478/s11600-010-0052-3.

Bock, H., A. Jäggi, D. Švehla, G. Beutler, U. Hugentobler, and P. Visser (2007), Precise orbit determination for the GOCE satellite using GPS, *Adv. Space Res.* **39**, 10, 1638-1647, DOI: 10.1016/j.asr.2007.02.053.

Bock, H., A. Jäggi, U. Meyer, P. Visser, J. van den Ijssel, T. van Helleputte, M. Heinze, and U. Hugentobler (2011), GPS-derived orbits for the GOCE satellite, *J. Geod.* **85**, 11, 807-818, DOI: 10.1007/s00190-011-0484-9.

Bruinsma, S.L., J.C. Marty, G. Balmino, R. Biancale, C. Foerste, O. Abrikosov, and H. Neumayer (2010), GOCE gravity field recovery by means of the direct numerical method. **In:** *ESA Living Planet Symposium, 28 June – 2 July 2010, Bergen, Norway*.

Carrion, D., G. Vergos, A. Albertella, R. Barzaghi, I.N. Tziavos, and V.N. Grigoriadis (2015), Assessing the GOCE models accuracy in the Mediterranean area. **In:** *Assessment of GOCE Geopotential Models, Sp. Issue: Newton's Bull. N. 5*, 63-82.

Casotto, S., F. Gini, F. Panzetta, and M. Bardella (2013), Fully dynamic approach for GOCE precise orbit determination, *Bull. Geofis. Teor. Appl.* **54**, 4, 367-384; DOI: 10.4430/bgta0108.

Cheng, M., and J.C. Ries (2015), Evaluation of GOCE Gravity Models with SLR Orbit Tests. **In:** *Assessment of GOCE Geopotential Models, Sp. Issue: Newton's Bull. N. 5*, 187-192.

de Matos, A.C.O.C., D. Blitzkow, G. do Nascimento Guimarães, M.C.B. Lobianco, and I. de Oliveira Campos (2015), Evaluation of recent GOCE geopotential models in South America. **In:** *Assessment of GOCE Geopotential Models, Sp. Issue: Newton's Bull. N. 5*, 83-104.

Drewes, H. (2012), International Centre for Global Earth Models (ICGEM). **In:** The Geodesist's Handbook 2012, *J. Geod.* **86**, 10, 932-934, DOI: 10.1007/s00190-012-0584-1.

Drożyner, A. (1995), Determination of orbits with Toruń Orbit Processor system, *Adv. Space Res*. **16**, 12, 93-95, DOI: 10.1016/0273-1177(95)98788-P.

ESA (2010), GOCE Level 2 Product Data Handbook, European GOCE Gravity Consortium, ESA Tech. Note GO-MA-HPF-GS-0110, European Space Agency, Noordwijk.

ESA (2014), GOCE Flight Control Team; GOCE End-of-Mission Operations Report, Issue 1, July 2014.

Eshagh, M., and M. Najafi-Alamdari (2007), Perturbations in orbital elements of a low Earth orbiting satellite, *J. Earth Space Phys.* **33**, 1, 1-12.

Förste, Ch., F. Flechtner, R. Schmidt, R. Stubenvoll, M. Rothacher, J. Kusche, H. Neumayer, R. Biancale, J.-M. Lemoine, F. Barthelmes, S. Bruinsma, R. Koenig, and U. Meyer (2008), EIGEN-GL05C – A new global combined high-resolution GRACE-based gravity field model of the GFZ-GRGS cooperation, *Geophys. Res. Abstr.* **10**, EGU2008-A-03426.

Förste, Ch., S.L. Bruinsma, F. Flechtner, J.Ch. Marty, Ch. Dahle, O. Abrikosov, J.M. Lemoine, H. Neumayer, F. Barthelmes, R. Biancale, and R. König (2014), EIGEN-6C4 – The latest combined global gravity field model including GOCE data up to degree and order 1949 of GFZ Potsdam and GRGS Toulouse, *Geophys. Res. Abstr.* **16**, EGU2014-3707.

Gruber, Th., P.N.A.M. Visser, Ch. Ackermann, and M. Hosse (2011), Validation of GOCE gravity field models by means of orbit residuals and geoid comparisons, *J. Geod.* **86**, 807-818, DOI: 10.1007/s00190-011-0484-9.

Heiskanen, W., and H. Moritz (1967), *Physical Geodesy*, W.H. Freeman and Co., San Francisco.

Hirt, C., M. Rexer, and S. Claessens (2015), Topographic evaluation of fifth-generation GOCE gravity field models globally and regionally. **In:** *Assessment of GOCE Geopotential Models, Sp. Issue: Newton's Bull. N. 5*, 163-186.

Jäggi, A., U. Hugentobler, and G. Beutler (2006), Pseudo-stochastic orbit modeling techniques for low-Earth orbiters, *J. Geod.* **80**, 1, 47-60, DOI: 10.1007/s00190-006-0029-9.

Jäggi, A., H. Bock, U. Meyer, G. Beutler, and J. van den Ijssel (2015), GOCE: assessment of GPS-only gravity field determination, *J. Geod.* **89**, 1, 33-48, DOI: 10.1007/s00190-014-0759-z.

Lejba, P., S. Schillak, and E. Wnuk (2007), Determination of orbits and SLR stations' coordinates on the basis of laser observations of the satellites Starlette and Stella, *Adv. Space Res.* **40**, 1, 143-149, DOI: 10.1016/j.asr.2007.01.067.

Lemoine, F.G., S.C. Kenyon, J.K. Factor, R.G. Trimmer, N.K. Pavlis, D.S. Chinn, C.M. Cox, S.M. Klosko, S.B. Luthcke, M.H. Torrence, Y.M. Wang, R.G. Williamson, E.C. Pavlis, R.H. Rapp, and T.R. Olson (1998), The Development of the Joint NASA GSFC and the National Imagery and Mapping Agency (NIMA) Geopotential Model EGM96, NASA Technical Paper NASA/TP1998206861, Goddard Space Flight Center, Greenbelt, USA.

Mayer-Gürr, T., E. Kurtenbach, A. Eicker, and J. Kusche (2011), The ITG-1Grace 2010 gravity field model, Institute of Geodesy and Geoinformation, Bonn University, Bonn, Germany, available from: http://www.igg.uni-bonn.de/apmg/index.php?id=itg-grace2010.

Melbourne, W., R. Anderle, M. Feissel, R. King, D. McCarthy, D. Smith, B. Tapley, and R. Vincente (1983), Project MERIT Standards, Circ. 167, U.S. Naval Observatory, Washington, D.C., U.S.A.

Pail, R., S. Bruinsma, F. Migliaccio, Ch. Förste, H. Goiginger, W.D. Schuh, E. Höck, M. Reguzzoni, J.M. Brockmann, O. Abrikosov, M. Veicherts, T. Fecher, R. Mayrhofer, I. Krasbutter, F. Sansò, and C.Ch. Tscherning (2011), First GOCE gravity field models derived by three different approaches, *J. Geod.* **85**, 819-843, DOI: 10.1007/s00190-011-0467-x.

Papanikolaou, Th.D., and D. Tsoulis (2014), Dynamic orbit parameterization and assessment in the frame of current GOCE gravity models, *Phys. Earth Planet. In.* **236**, 1-9, DOI: 10.1016/j.pepi.2014.08.003.

Pavlis, N.K., S.A. Holmes, S.C. Kenyon, and J.K. Factor (2012), The development and evaluation of the Earth Gravitational Model 2008 (EGM2008), *J. Geophys. Res.* **117**, B04406, DOI: 10.1029/2011JB0010.1029/2011JB008916.

Reigber, Ch., H. Jochmann, J. Wünsch, S. Petrovic, P. Schwinzer, F. Barthelmes, K.H. Neumayer, R. König, Ch. Förste, G. Balmino, R. Biancale, J.M. Lemoine, S. Loyer, and F. Perosanz (2005), Earth gravity field and seasonal variability from CHAMP. **In:** *Earth Observation with CHAMP – Results from Three Years in Orbit*, Springer, Berlin, 25-30.

Rummel, R., D. Muzi, M. Drinkwater, R. Floberghagen, and M. Fehringer (2009), GOCE: Mission overview and early results. **In:** *The 2009 American Geophysical Union Fall Meeting, 14-18 December 2009*, *San Francisco, USA.*

Sośnica, K. (2014), *Determination of Precise Satellite Orbits and Geodetic Parameters using Satellite Laser Ranging*, Astronomical Institute, Faculty of Science, University of Bern, Switzerland.

Sośnica, K., D. Thaller, A. Jäggi, R. Dach, and G. Beutler (2012), Sensitivity of LAGEOS orbits to global gravity field models, *Artif. Sat.* **47**, 2, 47-65, DOI: 10.2478/v10018-012-0013-y.

Šprlák, M., C. Gerlach, and B.R. Pettersen (2015), Validation of GOCE global gravitational field models in Norway. **In:** *Assessment of GOCE Geopotential Models, Sp. Issue: Newton's Bull. N. 5*, 13-24.

Standish, E.M., X.X. Newhall, J.G. Williams, and D.K. Yeomans (1992), Orbital ephemerides of the sun, moon and planets. **In:** P.K. Seidelmann (ed.), *Explanatory Supplement to the Astronomical Almanac*, University Science Books, Mill Valley, 279-323.

Tapley, B., S. Bettadpur, M. Watkins, and C. Reigber (2004), The gravity recovery and climate experiment: Mission overview and early results, *Geophys. Res. Lett.* **31**, L09607, DOI: 10.1029/2004GL019920.

Voigt, C., and H. Denker (2015), Validation of GOCE gravity field models in Germany. **In:** *Assessment of GOCE Geopotential Models, Sp. Issue: Newton's Bull. N. 5*, 37-48.

Weigelt, M., T. van Dam, A. Jäggi, L. Prange, M.J. Tourian, W. Keller, and N. Sneeuw (2013), Time-variable gravity signal in Greenland revealed by high-low satellite-to-satellite tracking, *J. Geophys. Res.* **118**, 7, 3848-3859, DOI: 10.1002/jgrb.50283.

Yi, W. (2012), An alternative computation of a gravity field model from GOCE, *Adv. Space Res.* **50**, 3, 371-384, DOI: 10.1016/j.asr.2012.04.018.

Yi, W., R. Rummel, and T. Gruber (2013), Gravity field contribution analysis of GOCE gravitational gradient components, *Stud. Geophys. Geod.* **57**, 174-202, DOI: 10.1007/s11200-011-1178-8.

Received 30 September 2015
Received in revised form 4 July 2016
Accepted 18 October 2016

DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 6, Dec.2016, pp. 2781-2793
DOI: 10.1515/acgeo-2016-0061

The Relationship of Stratospheric QBO with the Difference of Measured and Calculated NmF2

Kadri KURT[1], Ali YEŞIL[1], Selçuk SAĞIR[2], and Ramazan ATICI[2]

[1]Department of Physics, Faculty of Sciences, Firat University, Elazig, Turkey
e-mails: kadrikurtt@hotmail.com (corresponding author), ayesil@firat.edu.tr

[2]Department of Physics, Faculty of Arts and Sciences, Mus Alparslan University, Mus, Turkey; e-mails: s.sagir@alparslan.edu.tr, r.atici@alparslan.edu.tr

Abstract

The relationship between stratospheric QBO and the difference (ΔNmF2) between NmF2 calculated with IRI-2012 and measured from ionosondes at the Singapore and Ascension stations in the equatorial region was statistically investigated. As statistical analysis, the regression analysis was used on variables. As a result, the relationship between QBO and ΔNmF2 was higher for 24:00 LT (local time) than 12:00 LT. This relationship is positive in the solar maximum epoch for both stations. In the solar minimum epoch, it is negative at 24:00 LT for Ascension and at 12:00 LT for Singapore. Furthermore, it was seen that the relationship of the ΔNmF2 with both the easterly and westerly QBO was negative for all solar epochs and every LT, at Ascension station. This relationship was only positive for solar maximum epoch and 12:00 LT, at Singapore station.

Key words: International Reference Ionosphere, QBO, NmF2, regression analysis.

1. INTRODUCTION

The empirical International Reference Ionosphere (IRI) model (Bilitza 2001) is actively used in a great variety of applied and research projects. It is well

Ownership: Institute of Geophysics, Polish Academy of Sciences;

known that IRI is an empirical ionospheric model based on experimental observations of the ionospheric plasma, either ground-based or *in situ* measurements. In particular, IRI provides a basis for the simulation and prediction of the ionospheric radio wave propagation. The main purpose of IRI is to provide reliable ionospheric densities, composition, and temperatures (*e.g.*, Bilitza 2001, Bilitza *et al.* 1979). The IRI can be used as the "quiet ionosphere" reference in applications that detect and study ionospheric disturbances. The model takes into account daily and seasonal variations, as well as the impact of solar activity on ionospheric conditions.

The terrestrial ionosphere is a layer that starts at approximately 60 km and extends to about 1000 km above the planet's surface. This layer is generated due to the interaction between solar radiation and the atmospheric constituents (Rishbeth and Garriott 1969, Schunk and Nagy 2009). The electron density is one of parameters characterizing the ionospheric F region. The electron density at heights near the F2 layer could be more sensitive to changes of neutral composition, temperature, and horizontal winds (Buresova *et al.* 2014). One of the sources that affect the electron density are the dynamic processes in the lower atmosphere. The dynamic processes in the lower atmosphere affect the ionosphere through the electrical and electromagnetic waves and upward propagating waves in the neutral atmosphere. Upward propagating waves in the neutral atmosphere are the most important because they can store energy and have atmospheric modifications. The metrological effects in the ionosphere can be caused by upward propagating gravity waves, and tidal and planetary waves (Lastovicka 2006, Kazimirovsky *et al.* 2003).

The Quasi-Biennial Oscillation (QBO) is a quasi-periodic interannual oscillation of the tropical stratospheric zonal winds between easterlies and westerlies with a mean period of 28-29 months (Heaps *et al.* 2000, Baldwin *et al.* 2001). Easterly and westerly phase of QBO develops at the top of the stratosphere and propagates downward at ~1 km per month until they dissipate at the tropical tropopause (Lindzen 1987). The west phase of QBO whose amplitude is 10-20 m/s descends faster than the east phase of QBO whose amplitude is 20-30 m/s. The phases are coherent through the whole equatorial belt at any given time. Peak amplitudes are over the equator at an altitude of 24 km (~30 mb) and amplitudes decrease away from the equator. The QBO by means of the waves, as seen in Fig. 1, can influence the Mesosphere Lower Thermosphere (MLT) beyond the stratopause (Baldwin *et al.* 2001, Mohanakumar 2008). Then, it may affect the electrical field of ionospheric E region and, thus, QBO can reach up to the F layer along the geomagnetic field lines from the E region (Chen 1992).

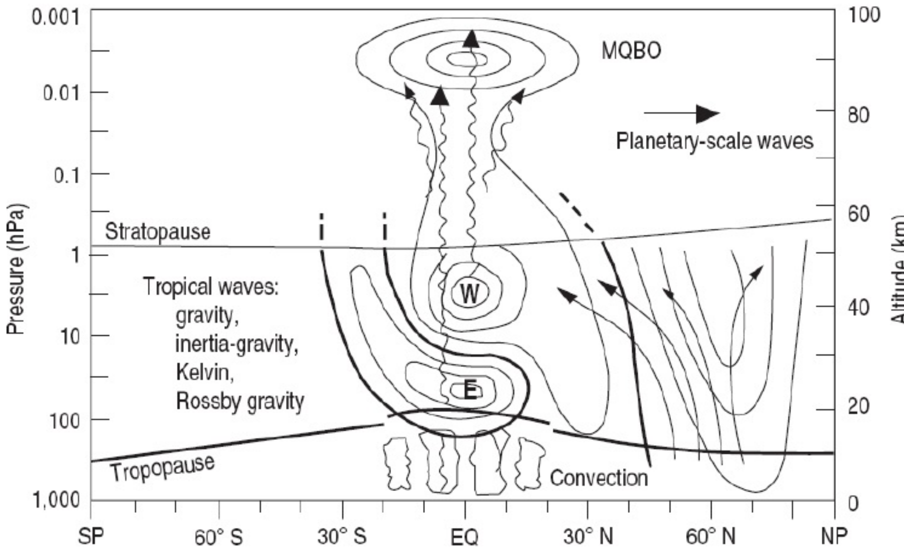


Fig. 1. The atmospheric waves affecting the spread of the QBO (Baldwin *et al.* 2001, Mohanakumar 2008).

In the present study, the relationship between QBO and difference (ΔNmF2) between NmF2 (maximum electron density of the ionospheric F2 region) calculated with IRI-2012 and measured at Singapore (01.22 °N, 103.55 °E) and Ascension (7.9 °S, 14.4 °W) stations in the equatorial region is statistically investigated. In this context, in order to study the relationship between variables, a multiple regression analysis is utilized.

2. THE STATISTICAL ANALYSES METHOD

A multiple regression analysis (Enders 2008, Sagir *et al.* 2015a, b) is used for processing of data measured at certain time intervals (Yadav *et al.* 2011). A prerequisite for this regression analysis is to determine the stability of variables. The stability of variables is determined by the Unit Root Test. If the series are not stable, with the mean and the variance changing with time, then these series are made stationary by calculating the first order of difference of variables (D(QBO) and D(ΔNmF2)). Next stage is to determine whether there is a relationship between the dependent variable (foF2) and independent variable (QBO) or not. In this study, this condition is provided by co-integration test. If there is a long-term relationship between the variables, then the last stage is to establish the regression model to determine the coefficients of the variables (Sagir *et al.* 2015a, b). For more detail information associated with these tests, see Sagir *et al.* (2015a, b). Thus, the statistical analysis model is described by the following formula:

$$D(\Delta \mathrm{NmF2}) = \mathrm{c} + \beta_0 \left(\mathrm{D(QBO)}\right) + \beta_1 \,\mathrm{Dummy\,Western} + \beta_2 \,\mathrm{Dummy\,Eastern} \quad , \quad (1)$$

where c is a constant, and β_0, β_1, and β_2 denote the variable coefficients. Dummy Western representing the western direction of QBO, and Dummy Eastern representing the eastern direction of QBO, are also included in the model (Sagir *et al*. 2015a, b).

To obtain NmF2 values, firstly, the F2 region critical frequency (foF2) data were taken from associated stations (SPIDR (data available at http://spidr.ngdc.noaa.gov)). foF2 is defined as the highest frequency at which radio waves can be transmitted vertically to the ionosphere and reflected by an ionospheric F region. It is denoted by f_c and defined as given in Hz, where NmF2 is the electron density in m^3 in the F2 region (Yesil *et al*. 2009). Then, these data were converted to NmF2 values through the formula $f_c = 9 \times 10^{-3} \sqrt{N_e}$. The obtained data were labeled "measured" data. Then, NmF2 values were calculated using the IRI-2012 model (URSI model was used as F-peak model) (http://omniweb.gsfc.nasa.gov/vitmo/iri2012) for the same stations at 12:00 LT and 24:00 LT. The obtained values were also called the "calculated" values. The difference (measured-IRI, ΔNmF2), which was not included in the IRI model, was calculated by subtracting the calculated NmF2 values from the measured NmF2 values. The relationship between determined ΔNmF2 values with QBO measured at 10 hPa height in the Singapore (for 1987 and the following years) and Canton Island (02.46 °S, 171.43 °W) (for the years between 1953 and 1965) stations (data available at the http://strat-www.met.fu-berlin.de/en/met) have been statistically investigated in solar maximum and solar minimum epochs.

3. RESULTS AND DISCUSSION

In this study, stations in the equatorial region, where the presence of QBO is clearly observed, were selected. Equation 1 given in Section 2 was applied to the data sets to investigate the effect of the stratospheric QBO on ΔNmF2. The monthly mean values of QBO and NmF2 data were used for the study. The monthly medians (Pancheva and Mukhtarov 1996, Ikubanni *et al*. 2014) of NmF2 values were adapted to the QBO data.

3.1 Results obtained for the Ascension station

The variations of ΔNmF2 and QBO for 24:00 LT and 12:00 LT have been indicated in Fig. 2 with graphs (a) and (c) for solar maximum and graphs (b) and (d) for solar minimum, respectively.

The time period from January 2001 to December 2003 was used for the solar maximum condition, and the period from January 2006 to December 2008 was used for the solar minimum condition (Kirov *et al.* 2014). In

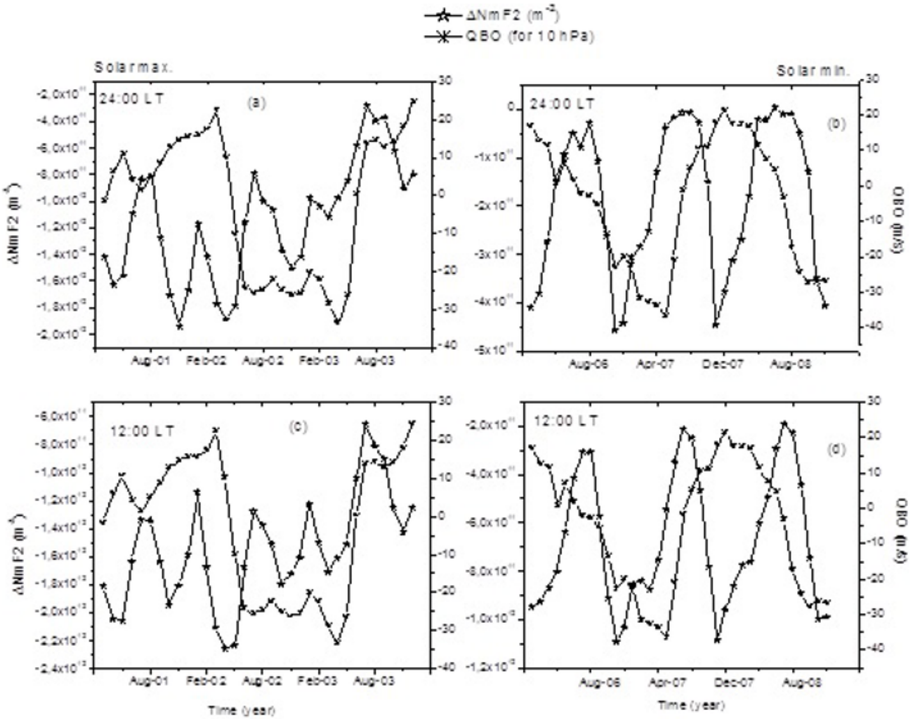


Fig. 2. The variation according to solar maximum and minimum cases of the relationship between QBO measured at 10 hPa altitude and ΔNmF2 for the Ascension station.

Fig. 2, for both times periods in the solar maximum epoch, while ΔNmF2 demonstrates a negative relationship for the easterly QBO (QBO values are less than zero), a positive relationship was observed for westerly QBO (QBO values are greater than zero). In the solar minimum epoch, a positive relationship was observed between westerly QBO and ΔNmF2. With regards to easterly QBO, a regular relationship was not observed at 24:00 LT; however, a positive relationship was observed for 12:00 LT in certain years.

Tables 1 and 2 give the unit root test results at Ascension in the solar maximum epoch for 24:00 LT and 12:00 LT, respectively. For the values to be stationary, each value given in the top section of the table should be larger in absolute values than the corresponding McKinnon (1996) critical values found at the bottom.

Table 1 shows that Δ(NmF2) and QBO contains the unit root in all three tests and the PP test with respect to their levels, respectively. So they are not stationary. Table 2 shows that Δ(NmF2) and QBO contain the unit root in Phillips–Perron Test (PP) and Kwiatkowski–Phillips–Schmidt–Shin Test

Table 1

The unit root test results at Ascension
in the case of solar maximum for 24:00 LT

Variables	for 24:00 LT		
	ADF	PP	KPSS
QBO	–4.25	–1.18	0.14
Δ(NmF2)	–0.48	–1.07	0.13
D(QBO)	–4.47	–2.49	0.20
D(Δ(NmF2))	–5.36	–5.60	0.50
The level of significance	McKinnon (1996) critical values		
1%	–4.27	–4.26	0.21
5%	–3.55	–3.55	0.14
10%	–3.21	–3.20	0.11

Table 2

The unit root test results at Ascension
in the case of solar maximum for 12:00 LT

Variables	for 12:00 LT		
	ADF	PP	KPSS
QBO	–0.48	–1.07	0.13
Δ(NmF2)	–5.08	–2.11	0.10
D(QBO)	–4.47	–2.49	0.20
D(Δ(NmF2))	–6.33	–5.63	0.38
The level of significance	McKinnon (1996) critical values		
1%	–4.27	–4.26	0.21
5%	–3.55	–3.55	0.14
10%	–3.21	–3.20	0.11

(KPSS) and Augmented–Dickey Fuller (ADF) and PP tests with respect to their levels (Kwiatkowski *et al*. 1992), respectively. So they are not stationary. The series become stationary in the first differences of QBO and Δ(NmF2) [D(QBO), D(Δ(NmF2))] in both tables.

Table 3 shows the co-integration test results obtained for the model in Eq. 1. Because the probability p values are smaller than 0.05 in the model and the ADF value is greater in absolute values than the McKinnon (1996)

Table 3

The co-integration test results for Ascension

Regression model	for 12:00 LT		for 24:00 LT	
	ADF	*p* value	ADF	*p* value
Model	–5.30	0.00	–6.20	0.00
The level of significance	McKinnon (1996) critical values			
1% 5% 10%		–2.65 –1.95 –1.60		

critical values at the bottom section of the table, ($|-5.30|>|-2.65|$) for 12:00 LT and ($|-6.20|>|-2.65|$) for 24:00 LT, there is a relationship between the variables. Furthermore, the level of statistical significance is at a rate of 1%.

Table 4 represents a multiple regression analysis results estimated by the model giving the relationship between QBO values at 10 hPa altitude and Δ(NmF2) for the Ascension in the solar maximum and minimum epochs.

Tables 4 and 5 list the results of the regression analysis. Ordinary Least Square (OLS) method forecasts are coherent in the presence of heteroskedasticity, but the standard errors are no longer valid. The White Heteroskedasticity (White Het.) Test is a test for heteroskedasticity in OLS residuals. The null hypothesis of the White Test is that there is no heteroskedasticity, and the value of this mutable has to also be greater than 0.05. The Durbin–Watson Test for serial correlation assumes that ε is stationary and normally perturbed with mean as zero. It tests the null hypothesis that the errors are not correlated and the values of variables require to be between 1.5 and 2.5. Probability (F statistics) (Prob. (F statistic)) tests the whole significance of the regression model and the value of this parameter has to be smaller than 0.05. Autoregressive (AR) processes have theoretical autocorrelation functions (ACFs) that decay toward zero, instead of cutting off to zero (Enders 2008, Sagir *et al.* 2015a, b).

An increase of 1 m/s in QBO measured at an altitude of 10 hPa in the solar maximum case for the Ascension station statistically leads to an increase of 7.6×10^{10} m^{-3} at 12:00 LT and a decrease of 1.64×10^{12} m^{-3} at 24:00 LT in Δ(NmF2). Furthermore, it is observed that easterly and westerly QBO causes a decrease in Δ(NmF2) for both LT. Similarly, QBO is more effective on Δ(NmF2) at 24:00 LT compared to 12:00 LT. The statistical relationship coefficient (R^2) is higher at night compared to the day.

Table 4

Regression analysis results for the Ascension station
in the solar minimum and maximum epochs

Coefficient	Solar minimum epoch		Solar maximum epoch	
	for 12:00 LT	for 24:00 LT	for 12:00 LT	for 24:00 LT
c	–7.67E+11 (0.00)*	–2.51E+11 (0.00)*	–1.52E+12 (0.00)*	–6.8E+10 (0.74)
β_0	6.36E+9 (0.01)*	–3.24E+9 (0.05)**	7.6E+10 (0.00)*	–1.64E+12 (0.05)**
β_1	–8.13E+11 (0.00)*	–2.51E+11 (0.01)*	–1.52 E+12 (0.00)*	–1.64 E+12 (0.00)*
β_2	–7.67E+11 (0.00)*	–2.51E+11 (0.00)*	–1.51 E+12 (0.01)*	–1.71 E+12 (0.00)*
AR (1)	0.06 (0.59)	0.34 (0.24)	8.17E-6 (0.03)**	8.11E-6 (0.04)**
R^2	0.91	0.85	0.71	0.83
Adj. R^2	0.89	0.79	0.62	0.79
Durbin–Watson	1.66	1.77	2.06	2.14
Prob. (F statistics)	0.00	0.00	0.00	0.00
Serial Cor. LM	0.00	0.92	0.10	0.44
White Het.	0.07	0.15	0.08	0.41

*, **, *** represents the significant level at 1, 5, and 10%, respectively.

An increase of 1 m/s in QBO measured at an altitude of 10 hPa in the solar minimum case for the Ascension station statistically leads to an increase of 6.36×10^9 m^{-3} at 12:00 LT and a decrease of 3.24×10^9 m^{-3} at 24:00 LT on Δ(NmF2). The easterly and westerly QBO lead to a decrease in both LT on Δ(NmF2). It is observed that QBO is more influential at 12:00 LT than at 24:00 LT on Δ(NmF2). The statistical relationship coefficient (R^2) is higher during the day compared to the night in a manner different to that in solar maximum. The relationship coefficient in the solar minimum epoch is greater than that of the solar maximum epoch.

The value of the Serial Correlation LM (Serial Cor. LM) test and the White Heteroskedasticity (White Het.) Test must be greater than 0.05. The Durbin–Watson Test value must be between 1.5 and 2.5. Probability (F-statistics) (Prob. (F-statistic)) must be smaller than 0.05 (Sagir *et al.* 2015a, b). Results of all these tests, given at the bottom of Table 4, indicate the accuracy of our model.

Table 5

Regression analysis results in the solar maximum and minimum epochs for the Singapore station

Coefficient	Solar minimum epoch		Solar maximum epoch	
	for 12:00 LT	for 24:00 LT	for 12:00 LT	for 24:00 LT
c	–6.68E+11 (0.00)*	–2.92E+11 (0.00)*	–4.24E+10 (0.09)***	–1.87E+12 (0.00)*
β_0	–6.60E+9 (0.03)**	1.20E+9 (0.09)***	4.48E+9 (0.06)**	–5.64E+9 (0.05)**
β_1	–7.12E+11 (0.00)*	–4.11E+11 (0.01)*	7.09E+10 (0.84)*	–1.75E+12 (0.00)*
β_2	–6.68E+11 (0.00)*	–2.92E+11 (0.00)*	4.24E+10 (0.91)*	–1.73E+12 (0.00)*
AR (1)	0.64 (0.54)	0.30 (0.24)	0.59 (0.54)	0.515 (0.01)*
R^2	0.64	0.79	0.75	0.89
Adj. R^2	0.58	0.69	0.69	0.84
Durbin–Watson	2.06	1.68	1.80	1.72
Prob. (F-statistics)	0.00	0.00	0.00	0.02
Serial Cor. LM	0.51	0.06	0.16	0.66
White Het.	0.46	0.91	0.12	0.07

*, **, *** represents the significant level at 1, 5, and 10%, respectively.

Because the value of Prob. (F-statistic) in Table 4 is less than 0.05, it is shown that the model is meaningful. The *p* values in the model that are less than 0.1 (indicated by parenthesis) also show that the model is meaningful according to the variables. Since the Durbin–Watson value varies between 1.5 and 2.5 and the values of Serial Cor. LM and White Het. are greater than 0.05, as given in the bottom section of Table 4, these statistical parameters also support the accuracy of the model (Sagir *et al*. 2015a, b).

3.2 Results obtained for the Singapore station

The variation of ΔNmF2 and QBO for 24:00 LT and 12:00 LT, respectively, have been indicated in Fig. 3 with graphs (a) and (c) for solar maximum and graphs (b) and (d) for solar minimum. Data pertaining to the time periods from July 1958 to June 1961 and from January 1963 to December 1965 were used for the solar maximum and solar minimum epochs, respectively. In the

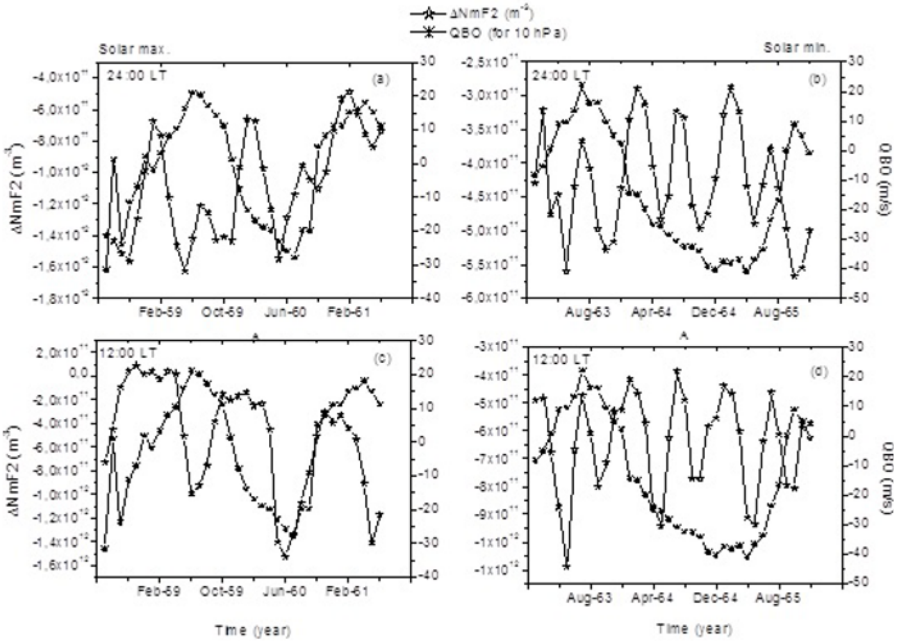


Fig. 3. The variation according to the solar maximum and minimum cases of the relationship between QBO measured at 10 hPa altitude and ΔNmF2 for the Singapore station.

solar maximum epoch, while QBO appeared to have a positive relationship for easterly 24:00 LT and a negative relationship for 12:00 LT, a negative relationship was observed for westerly QBO in both LT. With regards to the solar minimum epoch, a negative relationship is observed in both LT between both westerly and easterly QBO and ΔNmF2.

An increase of 1 m/s in QBO measured at an altitude of 10 hPa in the solar maximum case for the Singapore station statistically leads to an increase of 4.48×10^9 m^{-3} at 12:00 LT and a decrease of 5.64×10^9 m^{-3} at 24:00 LT in Δ(NmF2). Furthermore, it is observed that both the easterly and westerly QBO caused an increase in Δ(NmF2) for 12:00 LT and a decrease in Δ(NmF2) for 24:00 LT. QBO is more influential on Δ(NmF2) at 24:00 LT compared to 12:00 LT. The statistical relationship coefficient (R^2) between variables was higher at night compared to the day.

When the effect of QBO on Δ(NmF2) in the solar maximum case is compared between stations, it may be said that the QBO has a greater effect on electron density at the Ascension station.

For solar minimum epoch, an increase of 1 m/s in QBO for the Singapore station leads statistically to a decrease of 6.6×10^9 m^{-3} at 12:00 LT and an increase of 1.2×10^9 m^{-3} at 24:00 LT in Δ(NmF2). It is observed that the easterly and westerly QBO lead to a decrease on Δ(NmF2) for both LT. Similarly, QBO has a greater effect on Δ(NmF2) at 12:00 LT compared to 24:00 LT. The statistical relationship coefficient (R^2) is higher at night compared to daytime. The relationship coefficient in the solar minimum epoch is lower than that of the solar maximum epoch.

The previous study (Lühr and Xiong 2010), that investigated the differences between the ionosonde data with the IRI-model, emphasized that the IRI model overestimated the electron density. The reason for this overestimation could be lower atmospheric effects, such as QBO, that were not induced to the IRI model, in the equatorial region.

4. CONCLUSION

In order to emphasize the need to include QBO in the IRI model, the relationship between Δ(NmF2) and QBO was statistically investigated in this study. Within this context, a multiple regression analysis was performed in order to investigate the relationship between variables. Obtained results have demonstrated a statistically high relationship between QBO with Δ(NmF2) in the solar maximum and minimum epochs for both LT. This statistical relationship is greater at night in both the Ascension and Singapore stations at solar maximum epoch. Under solar minimum epoch, this relationship is higher during the day at the Ascension station and during the night at the Singapore station. The overestimation of electron density, in the equatorial region of the semi-empirical IRI model, is an issue addressed by some authors (*e.g.*, Lühr and Xiong 2010). Even though these obtained statistical results do not propose any physical mechanism concerning the relationship between variables, the difference of the electron density between the values obtained with the IRI model and the values obtained by the ionosonde may be said to be due to QBO, especially in the equatorial regions. This case indicates that QBO needs to be added to IRI-model.

References

Baldwin, M., L. Gray, T. Dunkerton, K. Hamilton, P. Haynes,W. Randel, J. Holton, M. Alexander, I. Hirota, and T. Horinouchi (2001), The quasi-biennial oscillation, *Rev. Geophys.* **39**, 2, 179-229, DOI: 10.1029/1999RG000073.

Bilitza, D. (2001), International reference ionosphere, *Radio Sci.* **36**, 2, 262-275, DOI: 10.1029/2000RS002432.

Bilitza, D., N.M. Sheikh, and R.A. Eyfrig (1979), Global model for the height of the F2-peak using M(3000)F2 values from the CCIR numerical map, *Telecommun. J.* **46**, 549-553.

Buresova, D., J. Lastovicka, P. Hejda, and J. Bochnicek (2014), Ionospheric disturbance sunder low solar activity conditions, *Adv. Space Res.* **54**, 1, 185-196, DOI: 10.1016/j.asr.2014.04.007.

Chen, P.R. (1992), Evidence of the ionospheric response to the QBO, *Geophys. Res. Lett.* **19**, 11, 1089-1092, DOI: 10.1029/91GL01564.

Enders, W. (2008), *Applied Econometric Time Series*, 2nd ed., John Wiley & Sons, 480 pp.

Heaps, A., W. Lahoz, and A. O'Neill (2000), The Quasi-Biennial zonal wind Oscillation (QBO), Centre for Global Atmospheric Modelling, Department of Meteorology, University of Reading, UK.

Ikubanni, S.O., J.O. Adeniyi, and O.K. Obrou (2014), Monthly mean foF2 model for an African low-latitude station and comparison with IRI, *Adv. Space Res.* **53**, 4, 635-646, DOI: 10.1016/j.asr.2013.12.016.

Kazimirovsky, E., H. Herraiz, and B.A. De La Morena (2003), Effects on the ionosphere due to phenomena occurring below it, *Surv. Geophys.* **24**, 2, 139-184, DOI: 10.1023/A:1023206426746.

Kirov, B., S. Asenovski, K. Georgieva, and V.N. Obridko (2014), What causes geomagnetic activity during sunspot minimum? *Geomag. Astron.* **55**, 8, 1033-1038, DOI: 10.1134/S0016793215080149.

Kwiatkowski, D., P.C.B. Phillips, P. Schmidt, and Y. Shin (1992), Testing the null hypothesis of stationary against the alternative of a unit root: How sure are we that economic time series have a unit root? *J. Econometrics* **54**, 1-3, 159-178, DOI: 10.1016/0304-4076(92)90104-Y.

Lastovicka, J. (2006), Forcing of the ionosphere by waves from below, *J. Atmos. Sol.-Terr. Phys.* **68**, 3-5, 479-497, DOI: 10.1016/j.jastp.2005.01.018.

Lindzen, R.S. (1987), The development of the theory of the QBO, *Bull. Am. Meteorol. Soc.* **68**, 4, 329-337, DOI: 10.1175/1520-0477(1987)068<0329:OTDOTT>2.0.CO;2.

Lühr, H., and C. Xiong (2010), IRI-2007 model over estimates electron density during the 23/24 solar minimum, *Geophys. Res. Lett.* **37**, 23, L23101, DOI: 10.1029/2010GL045430.

MacKinnon, J.G. (1996), Numerical distribution functions for unit root and cointegration tests, *J. Appl. Econom.* **11**, 6, 601-618, DOI: 10.1002/(SICI)1099-1255(199611)11:6<601::AID-JAE417>3.0.CO;2-T.

Mohanakumar, K. (2008), *Stratosphere Troposphere Interaction: An Introduction*, Springer Science + Business Media.

Pancheva, D.V., and P.Y. Mukhtarov (1996), A single-station spectral model of the monthly median F-region critical frequency, *Ann. Geophys.* **39**, 4, 807-818.

Rishbeth, H., and O.K. Garriott (1969), *Introduction to Ionospheric Physics*, Academic Press, New York.

Sagir, S., S. Karatay, R. Atici, A. Yesil, and O. Ozcan (2015a), The relationship between the Quasi Biennial Oscillation and Sunspot Number, *Adv. Space Res.* **55**, 1, 106-112, DOI: 10.1016/j.asr.09.035.

Sagir, S., R. Atici, O. Ozcan, and N. Yuksel (2015b), The effect of the stratospheric QBO on the neutral density of the D region, *Ann. Geophys.* **58**, 3, A0331, DOI: 10.4401/ag-6491.

Schunk, R.W., and A.F. Nagy (2009), *Ionospheres: Physics, Plasma Physics, and Chemistry*, 2nd ed., Cambridge University Press, Cambridge.

Yadav, S., R.S. Dabas, R.M. Das, A.K. Upadhayaya, S.K. Sarkar, and A.K. Gwal (2011), Variation of F-region critical frequency (foF2) over equatorial and low-latitude region of the Indian zone during 19th and 20th solar cycle, *Adv. Space Res.* **47**, 1, 124-137, DOI: 10.1016/j.asr..09.003.

Yesil, A., S. Sagir, and O. Ozcan (2009), Comparison of maximum electron density predicted by IRI-2001 with that measured over Chilton station, *E.-J. New World Sci. Acad.* **4**, 3, 92-99.

Received 4 December 2015
Received in revised form 28 March 2016
Accepted 15 June 2016

DE GRUYTER
OPEN

Acta Geophysica
vol. 64, no. 6, Dec. 2016, pp. 2794-2809
DOI: 10.1515/acgeo-2016-0114

Ranking of Sudden Ionospheric Disturbances by Means of the Duration of Vlf Perturbed Signal in Agreement with Satellite X-Ray Flux Classification

Ahmed AMMAR and Hassen GHALILA

Laboratoire de Spectroscopie Atomique, Moléculaire et Applications (LSAMA),
Faculty of Science, University of Tunis El Manar, Tunis, Tunisia;
e-mails: ammarahmed.ph@gmail.com (corresponding author),
ghalila.sevestre@wanadoo.tn

Abstract

Ionosphere undergoes permanently solar flares that quickly change its properties inducing sometime unwanted effects. These changes, or events, are known as Sudden Ionospheric Disturbances (SIDs) and the knowledge of their magnitude may be of great interest to anticipate probable damages. Currently, there does not exist any classification of these ionospheric changes based on their amplitude due to the wide variability of its responses. The only way to surmise their importance is to study them indirectly, throughout the classification of the X-ray flux intensity recorded by satellites. An attempt of classification based on their duration was proposed by the American Association of Variable Star Observers (AAVSO) but it is not very accurate because SID's duration is measured directly from the raw signal of the Very Low Frequency (VLF) signal and/or the Low Frequency (LF) signal. The aim of this work is to investigate, through a set of simple mathematical techniques applied to VLF/LF signals recorded by ground based receivers, the best method to estimate SIDs durations and then propose a new classification based on these durations.

Key words: VLF/LF, SIDs' classification, statistical ACP method, EMD method.

Ownership: Institute of Geophysics, Polish Academy of Sciences;

1. INTRODUCTION

The study of the Sun and its effects on the ionosphere is one of the many themes studied by the United Nations International Space Weather Initiative (ISWI). Among the many devices used for this topic, the Sudden Ionospheric Disturbance (SID) monitors (Scherrer *et al.* 2008) are a ground based VLF (Very Low Frequency) receivers dedicated to study Sudden Ionospheric Disturbances (SIDs) (Dellinger 1937). Those SIDs occur in association with solar flares and may have a very strong and relatively long-lasting effect on the ionosphere (Thomson and Clilverd 2001, Grubor *et al.* 2008). Indeed, the more intense the solar flares, the more important are their effects on the Earth. Therefore, the knowledge of the magnitude of a solar flare is crucial to prevent unwanted problems. This is the reason why these flares are closely monitored and classified with respect to X-ray fluxes intensity recorded by GOES satellites since the 1970s by the National Oceanic and Atmospheric Administration (NOAA) Space Weather Scales (SWPC/NOAA 2015).

Strong solar storms have been known as the most devastating solar effects that can cause satellites' entire failure or lead to communications blackout. We can mention as an example the "Halloween Storm" of October-November 2003, which is the strongest solar activity, the near-Earth space environment ever experienced in recent history (Rosen and Johnson 2004). Furthermore, during that period one of the solar flares had saturated the GOES's X-ray sensors, which made the estimation of its intensity difficult and consequently it was not classified. To overcome this problem, a study of the amplitude and phase of SID recorded by a ground VLF radio receiver (Thomson *et al.* 2005) succeed to give a reasonable estimation to the real intensity of this great solar flare. This example reveals the need to find another reliable method to classify solar events as an alternative to the NOAA one.

Due to the many parameters that govern the behaviour of the ionosphere (*e.g.*, solar cycle activity, seasonal period, day-night variation, path of the VLF from the transmitter to the receiver, chemical processes, *etc.*) (Mitra 1975, Dahlgren *et al.* 2011, Correia *et al.* 2011), the lasting effect of the flares is highly variable. Until the current day, there is only one rough technique used to classify the SIDs, which is based on their duration (AAVSO 2015). In other words, instead of asking the classical question: How strong is the effect of solar flare on the ionosphere? We ask the question: How much time does the effect of solar flare remain on the ionosphere? We propose, in the present work, new approaches to improve this method. To do this, we compare SIDs events observed from a ground VLF/LF recording (SID monitor) with solar flares observed from out space by NOAA's GOES satellite sensors. A special care is given for the treatment of the VLF recording and the assessment of the event's duration. In Section 1 we present the typical VLF signals and the way to identify a SID. In the Section 2 we present the different methods used to measure the event's duration on VLF recording obtained during June 2012. The last Section is dedicated to the discussion and the description of the criteria chosen for the classification.

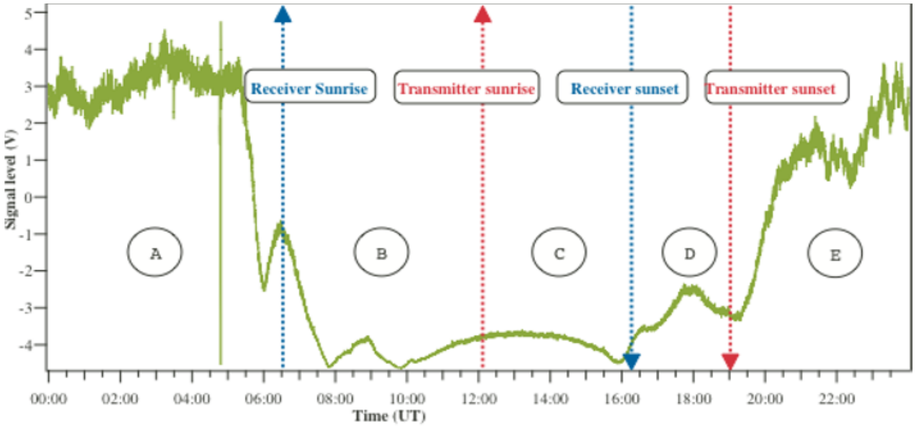
2. VLF RECORDING AND SIDs

SID monitors daily VLF by recording, every five seconds, VLF radio waves propagating in the earth-ionosphere waveguide. It is worth noting that the SID monitor records only the amplitude of the VLF/LF signal (no phase information is recorded). In the present work we have used data recorded by the SID monitor installed at the Faculty of Sciences of Tunis (FST) (36.833° N, 10.133° E), which has been tuned to the radio-VLF transmitter NAA operating at 24.00 kHz in Maine, USA (44.633° N, 67.267° W). Figure 1 shows the Great Circle Path (GCP) between the VLF transmitter station (Maine) and the VLF receiver station (Tunis) which is ~ 6326 km.

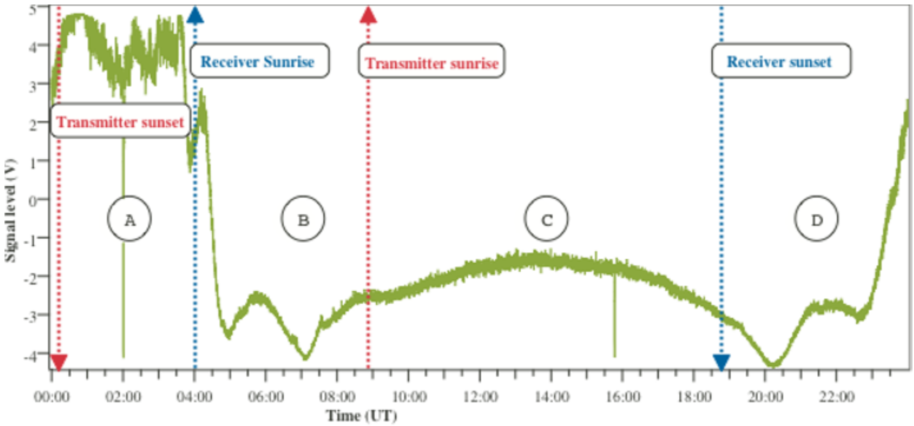
Figure 2 illustrates the diurnal and seasonal changes of amplitude of VLF signals; Fig. 2a corresponds to winter condition and Fig. 2b corresponds to summer condition. We can see from both figures, 2a and 2b, that the most stable part of the data is the part (C), when the GCP of the VLF wave is totally in the Earth's day side. This part of the signal will be referred to as "quiet diurnal level" for the remaining of the document.

We limit all record's analyzes to the diurnal zone, when the majority of the GCP is in the Earth's day-side. The detection of SID events is to track down

Fig. 1. Great Circle Path (GCP) between the NAA transmitter (24 kHz) and receiver at Tunis over a distance of ~6326 km.



(a) Tunis/NAA VLF signal recorded on 1 January 2012



(b) Tunis/NAA VLF signal recorded on 20 June 2012

Fig. 2. Variation of the data level of NAA (USA)-LSAMA (Tunisia) for a day without solar flares. We can see that the signal is divided to four parts depending on the elevation of the sun on both receiver and transmitter locations. Parts (A) and (E) are when the GCP of VLF wave is totally in the Earth's night-side, parts (B) and (D) are when the GCP is partially in the Earth's day-side, and part (C) is when GCP is totally in the Earth's day-side.

any disturbances from the quiet diurnal level. Then we systematically compare VLF recording with GOES X-ray Flux to make sure that these disturbances are induced by solar flares.

The shape of SID events depends on some factors, such as the distance between the transmitting and receiving stations, as well as the temporal variation of ionospheric parameters. They could have an increase in amplitude, some have an amplitude decrease, some have complex shapes in time (*e.g.*, a decrease below the unperturbed values that followed some of the strong increases in VLF signal) and in some cases the amplitude can be saturated. This variation in SID's shape could be investigated by comparing it with signals produced by a waveguide propagation model (Dahlgren *et al.* 2011). Analyses presented here are limited to SID events detected on signal coming from NAA transmitter and received at Tunis. All the SIDs have similar shapes with a positive increase in amplitude without any saturation. The module can be applied directly to SIDs with negative variation as well. As an example, Fig. 3b shows the signal of the 9 June 2012, involving three events close to 10:29, 11:30, and 16:48 UT. Comparing this signal to the X-ray Flux recorded by the GOES satellite (Fig. 3a) for the same day we can note the almost perfect synchronization with measured solar flares.

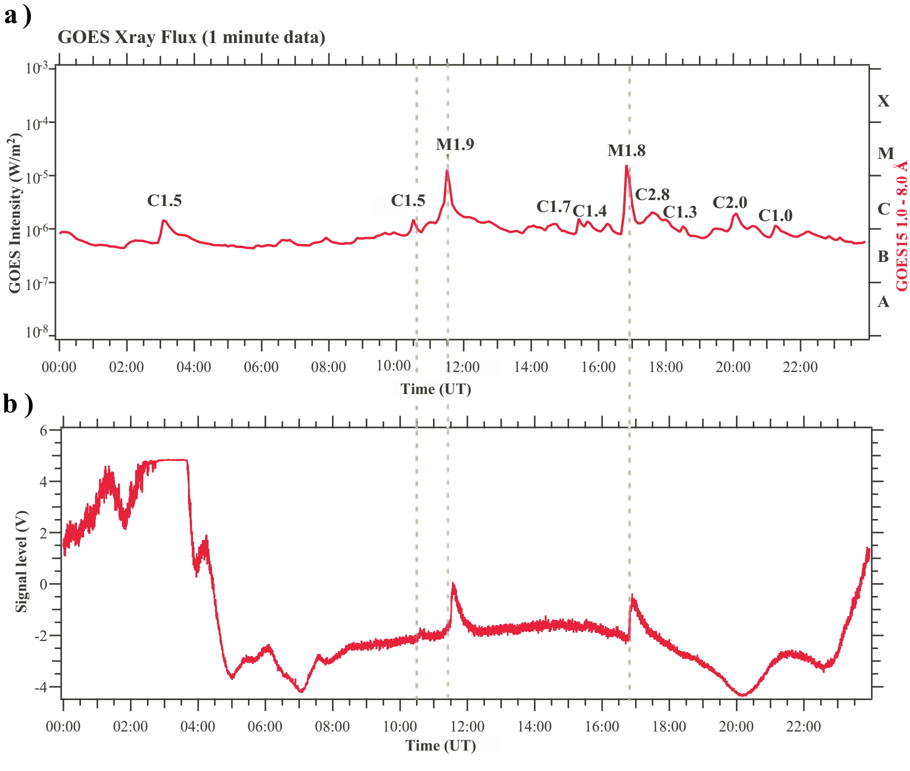


Fig. 3. Identification of the SID events with the GOES Flux. Dashed lines show how the identification is done. (a) X-ray Flux from the Sun observed from GOES satellite during 9 June 2012, and (b) Ground VLF amplitude observed from Tunis during 9 June 2012.

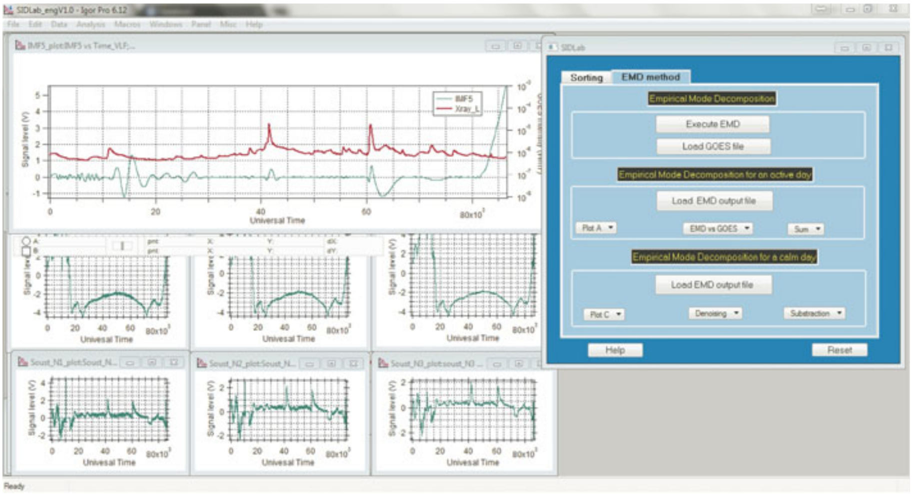


Fig. 4. Screen capture of SIDLab application developed under Igor Pro software (https://www.wavemetrics.com/).

This ascertainment allows us to conclude that these three events correspond to SIDs events. To make all these processes faster and also the methods applied in this work more efficient, we gathered all in one application called "SIDLab", whose interface is shown in Fig. 4.

3. SIGNAL PROCESSING OF JUNE 2012 VLF RECORDING

In this section, we describe and test several methods in order to estimate the length of SIDs. These methods are applied to data recorded during the month of June 2012 with our SID monitor. During this month we have detected 24 events. We believe that this number of events is sufficient to select a good method among those proposed for SIDs classification. Indeed, longer periods of VLF recording should be more rigorous.

3.1 Measure of SIDs durations from raw VLF data

We take directly from the raw data the chronology of the event, *i.e.*, the starting time, the time corresponding to the maximum and the end of the SID event using the methodology indicated by the AAVSO's report (AAVSO 2015) and as indicated by the methodology document SWPC (SWPC/NOAA 2015). This method is effective only near solar noon because during this period the VLF record is relatively flat and stable, being distant from the sunrise and sunset part of VLF data.

As we can see in Fig. 3, it is difficult to accurately collect these data for the first and the third SID. In fact, the first SID is almost completely hidden by the background noise while the third one performs its decay near the period of the sunset. These two peaks require more data processing in order to provide accurate values for their chronology.

3.2 Subtraction method: quiet day-active day

The first step in the signal processing is to eliminate the diurnal variation of the signal obtained without any SID. We chose a signal recorded in a day preferably with little noise and should be the nearest one to the active day of interest (which have signatures of SIDs) in order to keep identical conditions and better isolate the direct effects of solar flares. This day can be referred as a reference day (or control day). As we mentioned before, we only focus on the stable diurnal period, the time interval between sunset and sunrise. So we apply a cubic spline interpolation (the yellow line in Fig. 5) to this part of signal. This later will be used as a baseline that we will subtract from each active day (a day with SIDs, red line in Fig. 5). This method allows us to better isolate SIDs events and reduces the noise for a better detection of weak SIDs (*e.g.*, after applying this method the SID induced by solar flare C3.5 will be much clearer).

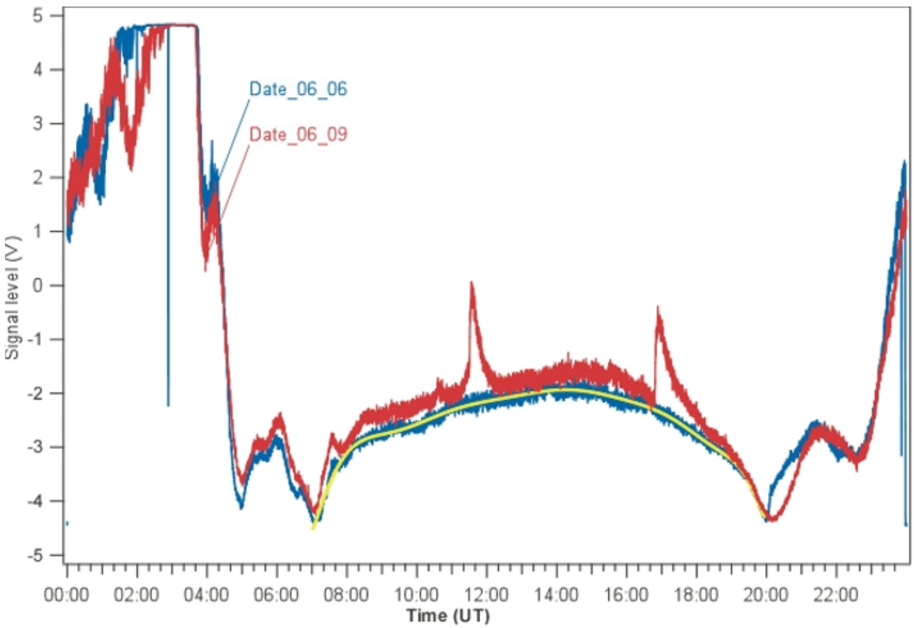


Fig. 5. Visualization of the two recording dates of 6 June 2012 (blue) and 9 June 2012 (red). The yellow line corresponds to a cubic spline of the selected reference day 6 June 2012, between the time 07:00 and 20:00 UT.

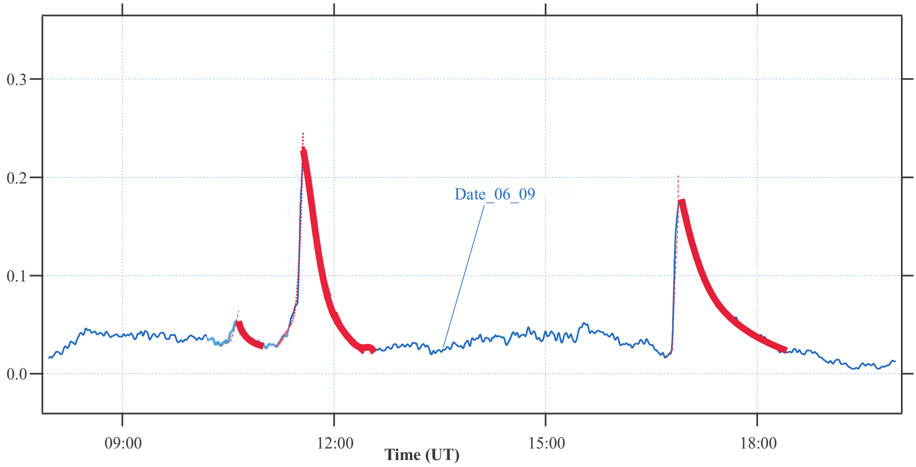


Fig. 6. Signal obtained after subtracting the cubic spline fit of 6 June 2012 from the record of the 9 June 2012 for the period between 07:00 and 20:00 UT data.

In Fig. 6, after applying this subtraction method, we find that AAVSO's methodology to classify SIDs is much easier than before and appreciably improves our estimates. Also, the small peak at 10:32 UT is much clearer and allowing the better determination of its duration.

3.3 Method of exponential decay

This method is combined with the subtraction method described in the previous section. As we have already mentioned in Section 2, the shape of recorded SIDs in present work shows two parts, an increase followed by a decrease in amplitude. The first part, which is very brief, corresponds approximately to the ionization phase of the ionosphere by X-ray radiations and the second part, with longer duration, corresponds approximately to the phase of recombination of the ionosphere and characterized by an exponential decay. The limit between these two parts is not always precisely defined because the time when ionization phase passes in recombination phases is not exactly amplitude independent. Also, the amplitude is not necessarily a monotonic function of the electron density and we cannot generally say that the time when amplitude finishes its increase is the same as the time of electron density maximum (Nina *et al.* 2012). We focus here our attention on the second part of the signal and we propose a reliable method for the determination of the recombination time for each detected SID.

The instant selected for characterizing the end time is taken as the exponential decay rate (Fig. 6, red line). This solution avoids the difficulty we encounter when the recovery time is very long and no end time is clearly distinguishable. Measurements of the duration of the SIDs by this method show a remarkable reduction of the duration (about half) in comparison with AAVSO's methods.

3.4 Empirical Mode Decomposition method

There are many methods of data processing (Fourier, wavelets), which are based on the decomposition of signals formed on the basis of the Eigen modes. The Empirical Mode Decomposition (EMD) is a data analysis method developed by Huang (Huang *et al.* 1998) for the study of oceanographic data. Subsequently, it has been introduced in other areas. The main objective of the EMD is to define decompositions that do not depend on the choice of a particular base. In addition, the EMD method is particularly well suited for the study of non-stationary signals, as is the case for our signals. In contrary, this method is not well suited for events that encompass several perturbations, as it is sometimes the case here.

We use the EMD software, a C compiled program, developed by Loudet (2009). The graphical user interface developed in our application (SIDLab) calls this program out and this greatly facilitates the comparison between various combinations of Intrinsic Mode Functions (IMF) and the GOES data.

Figure 7 shows again the three SIDs of 9 June 2012, analyzed with EMD method. We have selected the fifth mode (IMF5) to assess the duration of the three SIDs resulting from solar flares C1.6, M1.8, and M1.9, respective-

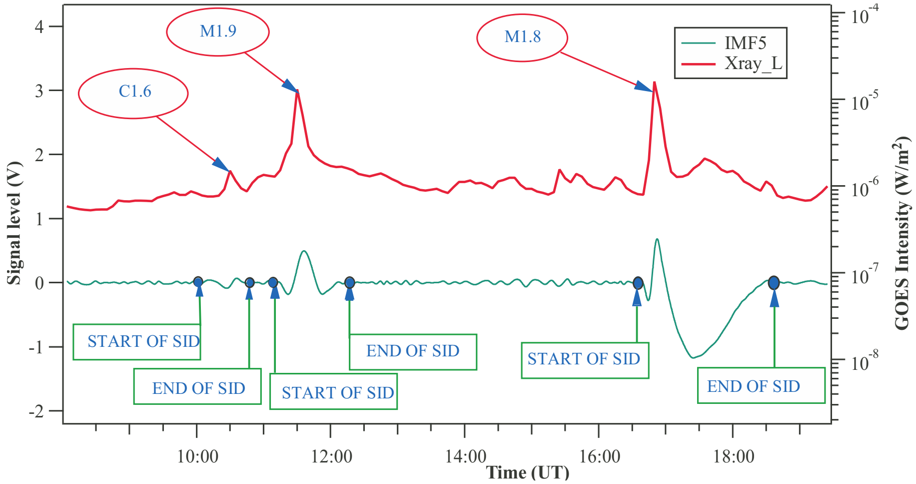


Fig. 7. Plot of GOES spectrum with the IMF5 component of the VLF data. The SIDs are interpreted by a strong oscillation around the zero of the data amplitude. The return to zero fixes the end of each SID event.

ly. By observing carefully Fig. 7, we can notice the improvement obtained by the EMD analysis regarding the temporal analysis of the events. Return to zero in the IMF component is more obvious to detect than the return to a level corresponding to normal conditions of propagation in the raw data. In contrary the determination of the exact position of the starting point of the SID is not always obvious. So basing on some additional criteria, *e.g.*, that the start of the SID must come after the start of the flare, and also taking into account the change in tendency which must be significant, we can better fix this position.

3.5 Noise reduction of the VLF signals

In order to reduce the noise we got after the subtract method (Fig. 6), we applied a noise reduction to the VLF raw data based on the EMD technique. Again, a special function is built in our application to easily achieve this operation. In this case, it eliminates the first three IMF (Loudet 2009) from raw VLF signal, getting a clearer, signal such as the signal in Fig. 8 (green line).

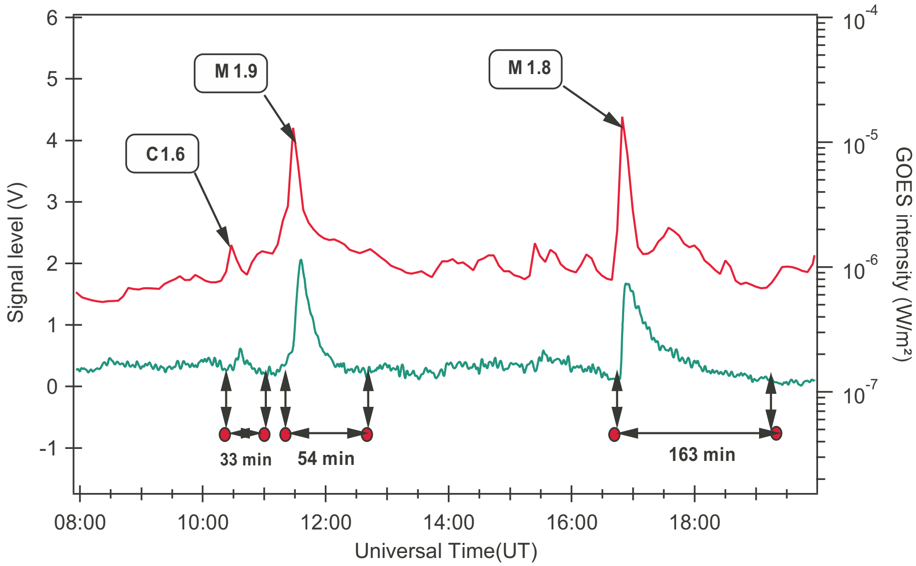


Fig. 8. Correlation between VLF recording (green) with that of GOES15 satellite (red), for the day 9 June 2012. The classification is indicated on the curves.

4. DISCUSSION

In Table 1, we assigned to each event occurred during June 2012 the durations measured by the different methods. We referred to each of these methods as follows: τ_{M1} for duration measured from the raw signal, τ_{M2} for the subtraction method, τ_{M3} for the exponential decay, τ_{M4} for the EMD method, and τ_{M5} for the noise reduction method.

Table 1

Estimated durations for each method

Date	Number of SID	M1[a)]	M2[b)]	M3[c)]	M4[d)]	M5[e)]	GOES[f)]	GOES class[g)]
7 June 2012	1	38	60	38	–	65	18	C2.7
7 June 2012	2	98	117	65	42	42	26	C9.1
9 June 2012	3	54	89	45	47	58	15	M1.9
9 June 2012	4	43	97	43	41	69	11	M1.8
13 June 2012	5	39	108	34	16	44	5	C1.3
13 June 2012	6	153	404	95	71	82	122	M1.2
13 June 2012	7	31	49	42	48	57	10	C2.7
14 June 2012	8	41	61	39	27	45	17	C5.0
14 June 2012	9	166	328	75	97	71	124	M1.9
15 June 2012	10	62	146	40	65	50	26	C3.4
15 June 2012	11	–	82	55	–	35	22	C2.0
16 June 2012	12	66	87	60	52	73	15	C1.8
17 June 2012	13	53	109	36	42	28	10	C3.9
20 June 2012	14	51	119	44	47	77	23	C3.1
27 June 2012	15	34	38	22	28	31	7	C3.4
28 June 2012	16	56	63	31	39	53	8	M2.4
29 June 2012	17	74	79	29	38	43	5	C2.9
30 June 2012	18	80	73	37	–	44	9	M2.2
30 June 2012	19	48	44	34	34	47	15	C4.4
30 June 2012	20	58	69	37	39	45	54	C2.7
30 June 2012	21	54	49	25	27	52	6	M1.0
30 June 2012	22	37	91	47	92	70	8	M1.6

Explanations: all the values are in minutes; (a) raw VLF signal, (b) raw minus reference, (c) exponential decay method, (d) EMD, (e) noise reduction method, (f) GOES duration, and (g) solar flare class associated with each SID.

The question that comes out from all these measurements is: which of these methods is the most relevant to classify SIDs according to their duration? Part of the answer can be formulated by the question: which are the best criteria we have to define in order to give a reliable classification. In our opinion, reliable classification should satisfy the following two criteria:

(i) Which method gives the best correlation to GOES classification?

(ii) Which method reveals less dispersion?

These two criteria act in the same direction, which consist to circumvent the overall terrestrial conditions (climate, ionospheric variability, noises, *etc*.). To fulfil these two criteria, we define two variables:

(i) Variable x_{Gi} measuring the discard between the duration of the proposed methods τ_{Mi} with the duration obtained by GOES Flux τ_G:

$$x_{Gi} = \tau_G - \tau_{Mi} \tag{1}$$

(ii) Variable x_{Mi} measuring the discard between the averages of the methods $\tau_{\overline{Mi}}$ (without GOES) with each method described in Section 4:

$$x_{Mi} = \tau_{Mi} - \tau_{\overline{Mi}} \tag{2}$$

Individual Standardized Deviation that takes into account the dependency of the error on the size of flare is defined for each proposed method. This parameter, named G_{Mi} and defined below, binds the two criteria in a unique quantity:

$$G_{Mi} = \frac{|x_{Gi} * x_{Mi}|}{d^2} \tag{3}$$

where $d = \tau_G - \tau_{\overline{Mi}}$ is a normalization factor. We analyse the duration of all the SIDs (21 events) measured during one month by means of the Principal Component Analysis (PCA). The aim of this technique is to reduce a large number of variables to a much smaller number of Principal Components (PCs) that best describe the data in terms of the variance. These PCs are a linear combination of the physical set of variables (SIDs events). This approach enables effective visualization, regression and classification of the multivariate data (Brereton 1992). In our case, it will help us to easily show the correlation that may exist between the different methods by clustering them. In our study, the binding parameters G_{Mi} are considered as the individual and the SID events as the variables. In this way we built a matrix of dimension 5*21 parameters, say GM_{ij}, where $i = 1 \ldots 5$ corresponds to the columns of Table 1 and $j = 1 \ldots 21$ to the lines. This statistical analysis is done using 'R' software (https://www.r-project.org/).

As shown in Fig. 9, the first two principal components (Dim1 + Dim2) explain 86.98% of the variance in the data matrix and the first one (Dim1) alone explains more than 73% of the variance. According to Cattell's scree test, the first two components are sufficient to get the maximum dispersion of the data. We can clearly see here that M_{G4}, M_{G5}, and M_{G3} are close both to each other and to the centre (0, 0) of the diagram relative to the M_{G1} and M_{G2}, which are more dispersed and farther from the centre. We recall here that we use the AAVSO method to define the duration of the SID in M_{G1}.

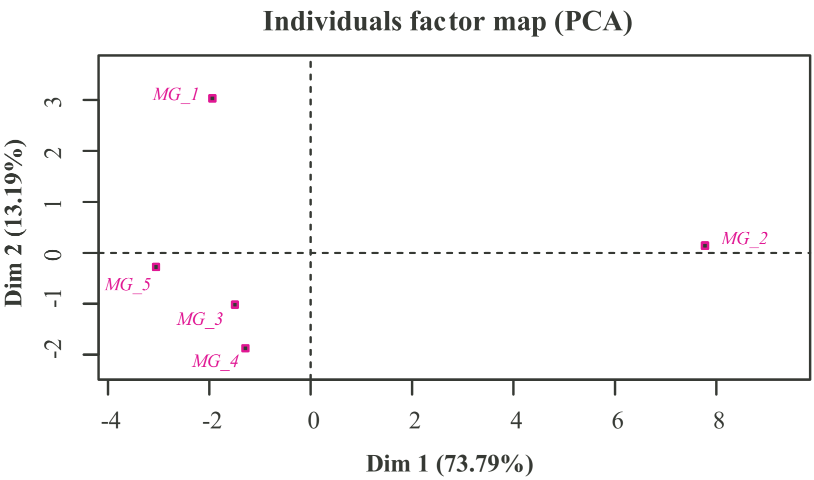


Fig. 9. Scores plots of the methods from PCA: first plane (Dim1, Dim2) with cumulative percentage of variance of 86.98%.

Those results indicate that the M_{G4}, M_{G5}, and M_{G3} methods based on EMD technique, noise reduction and exponential decay, respectively, are closer to the GOES duration and less dispersed than the last two methods based on direct measure of the duration. In other words, this means that these three methods better satisfy the two criteria. Furthermore, the parameter M_{G3} which is the closest to the centre (0, 0) of the diagram indicates that it is least dispersed among the three and thus the exponential decay is the most recommended method to determine the duration of the SID and accordingly its classification.

In the light of those results, we can assume that classification based on the exponential decay method will better correlate the GOES classification. To confirm this point, we use the Spearman's rank correlation (Gibbons and Chakraborti 2011) which seeks correlations between non-parametric data as our own. The results of the test are listed in Table 2.

Table 2

Non-parametric Spearman correlation rank

Method	p value	Rho
M3	6.0410^{-5}	0.741
M4	0.0043	0.593
M2	0.0121	0.483
M1	0.0281	0.425
M5	0.0503	0.359

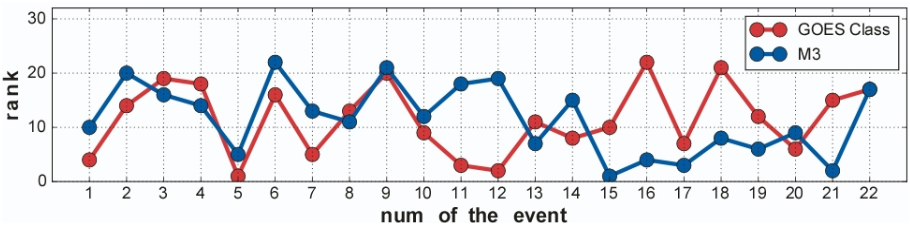


Fig. 10. Ranking of the SIDs based on their duration (in blue) and the ranking of the flares based on GOES intensity (in red).

We note the first rank of the method 3 with the smallest p value indicating a ”strong evidence against null hypothesis” attesting a very significant test and a correlation coefficient ”Rho” of around 0,74 which indicate a ”strong” correlation according to the hierarchy established by the Spearman statistics. We give in Fig. 10 the SIDs ranking based on their duration together with the flares ranking based on their GOES intensities. For example, for the first event (7 June 2012) the GOES flare is at the 4th rank (last column in Table 1) whereas the M3 method places it at the 10th rank. We can notice the good correlation between these two line graphs based on their similar behaviour. The same comparisons are done with the other methods but none of them showed such a good agreement with the GOES ranking which is in accordance with the values of Table 2.

5. CONCLUSION

The main goal of this work is to define objective criteria that allow us to create a stable and reliable classification of the SID on the basis of their duration. We investigate for this purpose different methods that minimize the dispersion of the values independently of the activity and noise present on the VLF recording. The exponential decay method (M3) seems to give a satisfactory result for the classification, showing a strong correlation with the satellite intensity classification. Moreover, this method is well adapted to analyze a weak signal (like the C1.6 present in our VLF recording). We plan to experiment with all these methods using a greater amount of VLF and LF recording and with all different shapes of SIDs described in Section 2 in order to improve and confirm this result. Future developments of the ”SIDLab” application will help us to get this goal.

Acknowledgments. To the support and the encouragements of Prof. Deborah Scherer from Stanford University, USA. The authors also thank STO (Société Tunisienne d’Optique) and ICTP through the Affiliated Center, which supported this work.

References

AAVSO (2015), Reducing data gathered by VLF monitoring systems, American Association of Variable Star Observers, Cambridge, MA, USA, available from: https://www.aavso.org/sid-reducing-data (accessed: 11 June 2015).

Brereton, R.G. (ed.) (1992), *Multivariate Pattern Recognition in Chemometrics: Illustrated by Case Studies*, Elsevier, Amsterdam.

Correia, E., P. Kaufmann, J.-P. Raulin, F. Bertoni, and H.R. Gavilán (2011), Analysis of daytime ionosphere behavior between 2004 and 2008 in Antarctica, *J. Atmos. Sol.-Terr. Phys.* **73**, 16, 2272-2278, DOI: 10.1016/j.jastp.2011.06.008.

Dahlgren, H., T. Sundberg, A.B. Collier, E. Koen, and S. Meyer (2011), Solar flares detected by the new narrowband VLF receiver at SANAE IV, *S. Afr. J. Sci.* **107**, 9/10, 40-47, DOI: 10.4102/sajs.v107i9/10.491.

Dellinger, J.H. (1937), Sudden disturbances of the ionosphere, *Proc. Inst. Radio Eng.* **25**, 10, 1253-1290, DOI: 10.1109/JRPROC.1937.228657.

Gibbons, J.D., and S. Chakraborti (2011), *Nonparametric Statistical Inference*, Springer, Berlin Heidelberg, DOI: 10.1007/978-3-642-04898-2_420.

Grubor, D.P., D.M. Šulić, and V. Žigman (2008), Classification of X-ray solar flares regarding their effects on the lower ionosphere electron density profile, *Ann. Geophys*. **26**, 7, 1731-1740, DOI: 10.5194/angeo-26-1731-2008.

Huang, N.E., Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung, and H.H. Liu (1998), The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, *Proc. Roy. Soc. A* **454**, 1971, 903-995, DOI: 10.1098/rspa.1998.0193

Loudet, L. (2009), Application of Empirical Mode Decomposition to the detection of Sudden Ionospheric Disturbances by monitoring the signal of a distant Very Low Frequency transmitter, available from: http://sidstation.loudet.org/emd-en.xhtml (accessed: 11 June 2015).

Mitra, A.P. (1975), D-region in disturbed conditions, including flares and energetic particles, *J. Atmos. Terr. Phys*. **37**, 6, 895-913, DOI: 10.1016/0021-9169(75)90005-7.

Nina, A., V. Čadež, V. Srećković, and D. Šulić (2012), Altitude distribution of electron concentration in ionospheric D-region in presence of time-varying solar radiation flux, *Nucl. Instrum. Meth. Phys. Res. B* **279**, 110-113, DOI: 10.1016/j.nimb.2011.10.019.

Rosen, R.D., and D.L. Johnson (2004), Service assessment. Intense space weather storms October 19 – November 07, 2003, U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service, Silver Spring, MD, USA, available from: http://www.nws.noaa.gov/os/assessments/pdfs/SWstorms_assessment.pdf (accessed: 30 September 2015).

Scherrer, D., M. Cohen, T. Hoeksema, U. Inan, R. Mitchell and P. Scherrer (2008), Distributing space weather monitoring instruments and educational materials worldwide for IHY 2007: The AWESOME and SID project, *Adv. Space Res.* **42**, 11, 1777-1785, DOI: 10.1016/j.asr.2007.12.013.

SWPC/NOAA (2015), GOES X-ray Flux, Space Weather Prediction Center, National Oceanic and Atmospheric Administration, Boulder, USA, available from: http://www.swpc.noaa.gov/products/goes-x-ray-flux (accessed: 11 June 2015).

Thomson, N.R., and M.A. Clilverd (2001), Solar flare induced ionospheric D-region enhancements from VLF amplitude observations, *J. Atmos. Sol.-Terr. Phys.* **63**, 16, 1729-1737, DOI: 10.1016/S1364-6826(01)00048-7.

Thomson, N.R., C.J. Rodger, and M.A. Clilverd (2005), Large solar flares and their ionospheric D region enhancements, *J. Geophys. Res*. **110**, A6, DOI: 10.1029/2005JA011008.

Received 3 April 2016
Received in revised form 13 August 2016
Accepted 18 October 2016