DE GRUYTER OPEN

Acta Geophysica
vol. 63, no. 1, Feb. 2015, pp. 1-16
DOI: 10.2478/s11600-014-0227-4

The 2014 Kefalonia Doublet (M_w6.1 and M_w6.0), Central Ionian Islands, Greece: Seismotectonic Implications along the Kefalonia Transform Fault Zone

Vassilios KARAKOSTAS, Eleftheria PAPADIMITRIOU, Maria MESIMERI, Charikleia GKARLAOUNI, and Parthena PARADISOPOULOU

Geophysics Department, Aristotle University of Thessaloniki, Thessaloniki, Greece
e-mails: vkarak@geo.auth.gr, ritsa@geo.auth.gr (corresponding author), mmesimer@geo.auth.gr, hagarl@geo.auth.gr, ppara@geo.auth.gr

A b s t r a c t

The 2014 Kefalonia earthquake sequence started on 26 January with the first main shock (M_w 6.1) and aftershock activity extending over 35 km, much longer than expected from the causative fault segment. The second main shock (M_w 6.0) occurred on 3 February on an adjacent fault segment, where the aftershock distribution was remarkably sparse, evidently encouraged by stress transfer of the first main shock. The aftershocks from the regional catalog were relocated using a 7-layer velocity model and station residuals, and their distribution evidenced two adjacent fault segments striking almost N-S and dipping to the east, in full agreement with the centroid moment tensor solutions, constituting segments of the Kefalonia Transform Fault (KTF). The KTF is bounded to the north by oblique parallel smaller fault segments, linking KTF with its northward continuation, the Lefkada Fault.

Key words: 2014 Kefalonia earthquake, Kefalonia Transform Fault, aftershocks, step-over zone, seismotectonics.

1. INTRODUCTION

On 26 January 2014 an M_w 6.1 earthquake occurred in the western part of Kefalonia Island, the most seismically active region in the Aegean and surrounding areas (Fig. 1). The activated area is part of the Kefalonia Transform Fault Zone (KTFZ), ~100 km long, consisting of the Kefalonia and Lefkada Fault branches, and linking the continental collision with the oceanic subduction zones, in the Ionian Sea (Greece). Scordilis *et al*. (1985) first suggested that the 1983 Kefalonia earthquake (M7.0) had a dextral strike-slip mechanism. The mainly strike-slip motion of the Kefalonia Fault was confirmed by waveform modeling for the 1983 earthquake by Kiratzi and Langston (1991) and for the 17 September 1972 earthquake (M6.3) by

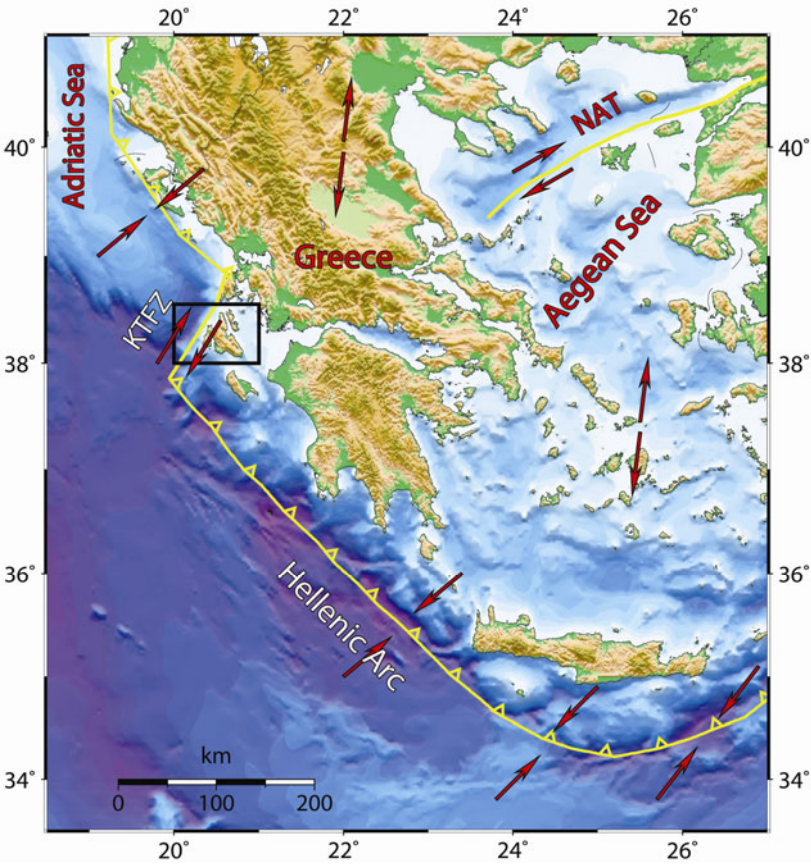


Fig. 1. The main geodynamic features of the Aegean and surrounding areas shown on a relief map. The active boundaries are shown as solid lines. The arrows indicate the approximate direction of relative plate motion. The study area is denoted by the square. KTFZ – Kefalonia Transform Fault Zone, NAT – North Aegean Trough.

Papadimitriou (1993). Later on, Louvari *et al*. (1999) investigated the strike slip nature of both Kefalonia and the adjacent Lefkada Fault branches, with maximum reported earthquake magnitudes of 7.4 and 6.6, respectively (Papazachos and Papazachou 2003). The fault zone follows the submarine Kefalonia valley, west of the island chain from Lefkada to Kefalonia. The southern prolongation of the KTFZ was shown by seismic line data and fault plane solutions located south of Kefalonia in the Ionian abyssal plain (Kokinou *et al*. 2006). For this region complete historical information exists for strong ($M \geq 6.5$) earthquakes in the last five centuries, revealing an average of about one such shock per decade (Papadimitriou and Papazachos 1985). These events have repeatedly destroyed urban areas, producing extensive damage and loss of life, with the most severe one being the 1953 paroxysm with four events (9 August, $M6.4$; 11 August, $M6.8$; 12 August, $M7.2$; 21 October, $M6.3$) that almost completely destroyed structures on Kefalonia Island.

The 2003 Lefkada sequence was the key event for the installation of a local network that provided for the first time the proper data for a detailed investigation of the activated main rupture (Karakostas *et al*. 2004), and the activation of secondary structures, which are capable to produce moderate to major earthquakes, thus necessarily to be taken into account in the seismic hazard assessment (Karakostas 2008, Karakostas and Papadimitriou 2010). Accurate microseismicity locations, derived from the recordings of a local network installed and operated on the two Islands (Kefalonia and Lefkada) in 2007-2008, clearly agree with the historical seismicity distribution and extent along the western coasts of both Islands, being located, however, closer to coastlines or onshore (Karakostas *et al*. 2010). Seismicity relocation performed for the purpose of the current study, further confirms this activity location. It is worth to note that during the last period (2007-2014, epicenters shown in red in Fig. 2) the area hosting the 2014 sequence was free of epicenters.

The current seismic excitation is the result of right lateral shear strain accumulated on a zone of weakness, which abuts and slightly overlaps the rupture area of the 1983 main shock ($M7.0$). Both 2014 main shocks and the activated area in general, are located inside stress enhanced areas revealed by the application of the stress evolutionary model (Papadimitriou 2002). The spatio-temporal properties of the sequence, giving more insight to the seismotectonics of this part of the active boundary, reveal the activation of three separate fault segments: the two dextral fault segments almost north-south striking and steeply east-dipping being associated with the two main shocks, along with the ENE-WSW striking also dextral lineaments that form a step-over transfer zone.

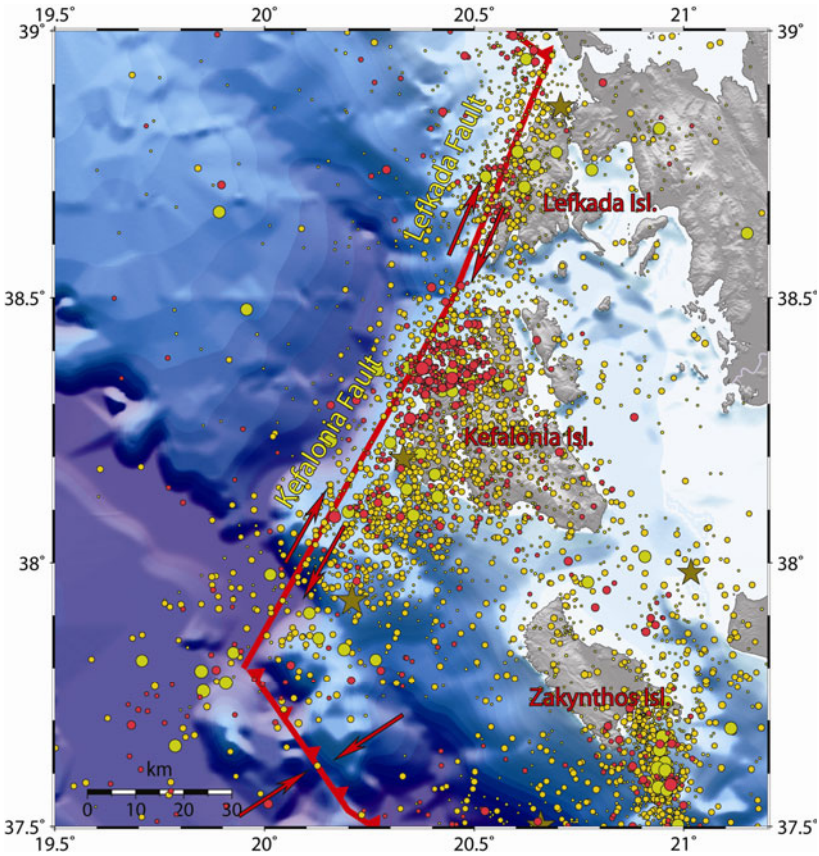


Fig. 2. Relocated seismicity for the period 1983-2013, along with the major active boundaries, the subduction front in the south, the Kefalonia Transform Fault Zone, with the distinctive Kefalonia and Lefkada branches, and the collision boundary north of Lefkada Island. Earthquakes of $M \geq 6.0$, 5.0, 4.0, and 3.0, are depicted by stars, circles, diamonds and squares, respectively, with light grey for the period 1983-2006 and red for 2007 – January 2014.

2. RELOCATION – MAIN SHOCKS AND AFTERSHOCKS SOURCE PARAMETERS

The earthquakes of the 2014 Kefalonia seismic sequence were located using waveform data from stations of the Hellenic Unified Seismological Network (HUSN) and accelerometers operated in the epicentral area by the Institute of Engineering Seismology and Earthquake Engineering (Fig. 3), whereas seismograms readings were carried out by the staff of the Geophysics Department of the Aristotle University of Thessaloniki. The *P*-waves velocity model (Haslinger *et al*. 1999) was used for this purpose (Table 1) – Wadati

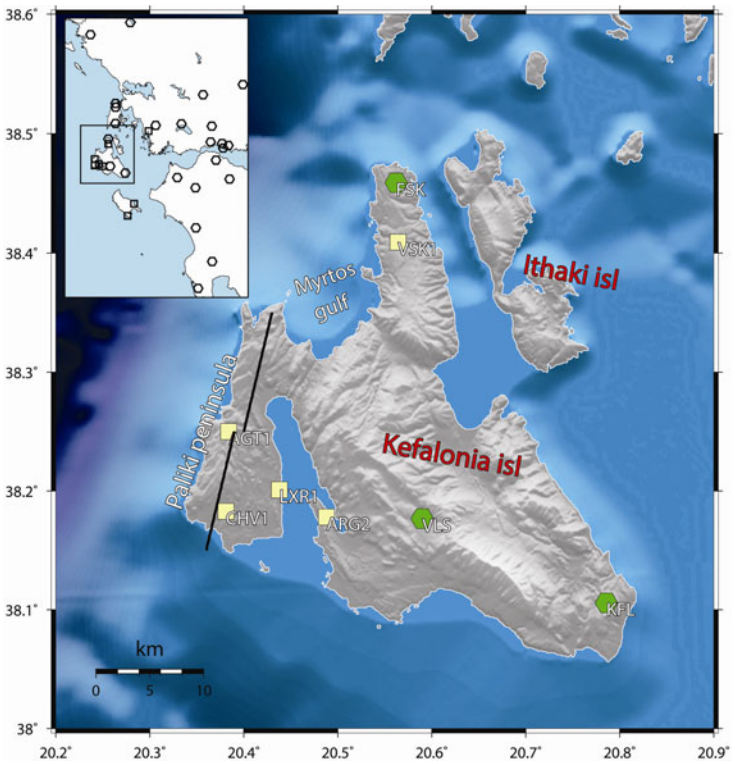


Fig. 3. Locations of the seismological stations (green polygons) of the Unified Hellenic Seismological Network and accelerometers (yellow squares) of the Institute of Engineering Seismology and Earthquake Engineering, installed and operating on Kefalonia Island, and used for aftershock relocation, near the activated fault segments which are shown by black line. The inset map shows the additional stations of the regional seismological network, recordings of which are also used.

Table 1

P-wave velocity model by Haslinger *et al*. (1999) used for the earthquake relocation

Velocity [km/s]	Width [km]
5.47	2.0
5.50	3.0
6.00	5.0
6.20	5.0
6.48	5.0
6.70	10.0
6.75	10.0
8.00	half space

plots based on the recordings of the first days show a *P*- to *S*-wave velocity ratio of $v_p/v_s = 1.78$. Thirty-three seismological stations in epicentral distances less than 150 km were employed for the earthquake location, performed with the HYPOINVERSE computer program (Klein 2000). The azimuthal coverage and the density of the seismological network is satisfactory, controlling adequately the calculated epicenters and focal depths. Several of the stations are inside or very close to the epicentral area, ensuring significantly accurate locations. Time corrections were calculated for all seismological stations, following a procedure of successive iterations until the changes in the calculated time corrections become negligible. In this way lateral variations, which are not included in the 1D velocity model, are taken into account (Karakostas *et al*. 2012, 2014). The data obtained were relocated using the double difference technique (Waldhauser and Ellsworth 2000, Waldhauser 2001). From 1180 earthquakes analyzed until 16 February 2014, 1150 (97%) were relocated, using the catalog data for both *P*- and *S*-phases applying the conjugate gradients method (LSQR, Paige and Saunders 1982). Results were obtained performing 25 iterations by applying distance and misfit weighting after the fifth and tenth iterations, respectively.

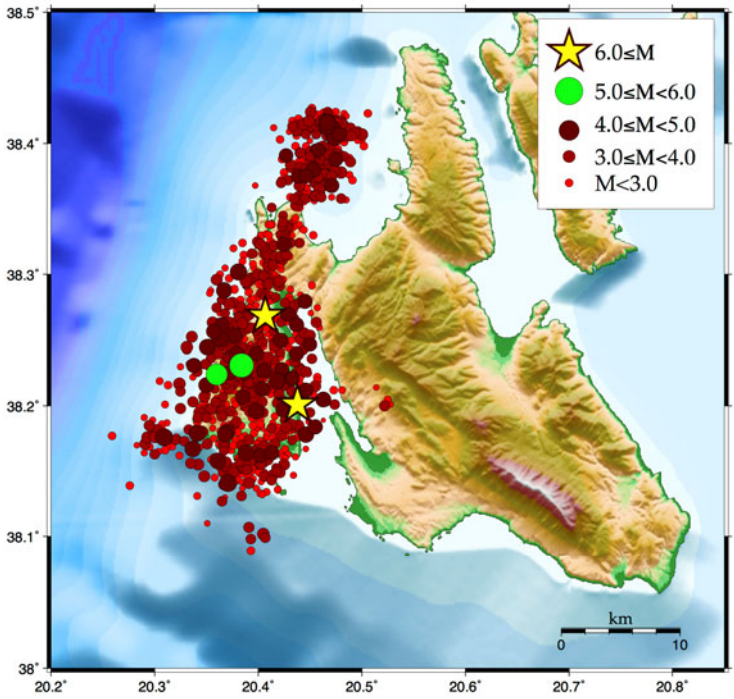


Fig. 4. Relocated aftershock activity of the first 22 days of the sequence (26 January – 16 February 2014).

The aftershock distribution (Fig. 4) for a 22-day period (26 January – 16 February 2014, 1180 earthquakes) reveals a seismic zone trending almost NNE-SSW, parallel to the main axis of the Paliki peninsula, as well as to the north offshore area in Myrtos gulf. The two main earthquakes of the sequence (M_w 6.1 and M_w 6.0, shown as stars) occurred in the southern and the central part of the peninsula, respectively. The seismic sequence started at the southern part of Paliki with the first strong earthquake (M_w 6.1, GCMT solution: strike = 20°/dip = 65°/rake = 177°). In the first 24 hours aftershock activity, particularly earthquakes with $M > 4.0$, is concentrated on an area of about 13 km long, starting from the southern coasts of Paliki and going to the north (Fig. 5a). Further north of this area, up to the northern coasts of the peninsula, seismicity is comparatively low with a lack of $M > 4.0$ events (inside the ellipse of Fig. 5a). In the southernmost part of the low activity area the second main earthquake occurred on 3 February 2014 (Fig. 5b). The aftershocks in the first 24 hours after the second strong event (encompassing

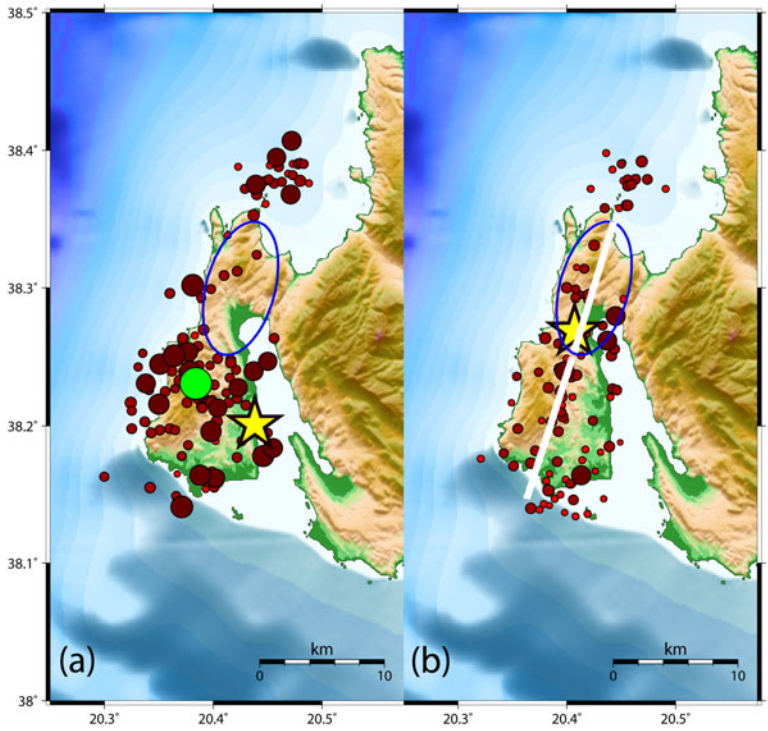


Fig. 5: (a) Aftershock activity for one day after the occurrence of the 26 January 2014 main shock, and (b) one day after the 3 February 2014 main shock. Symbols are the same as in Fig. 4. The ellipse indicates a comparatively low activity area.

in the ellipse in Fig. 5b) are less in number and of smaller magnitudes, in comparison with the first 24 hours activity following the first main shock. However, the seismic zone maintains its dimensions as in the first day of activity, which along with the GCMT focal mechanism (strike = 12°/dip = 45°/rake = 154°) indicate a 10 km-long rupture just north of the first one.

The low activity observed in the rupture area of the second main event soon after the initiation of the sequence, persisted even after the occurrence of the second main shock, implying an asperity that remained locked after the first main shock and that was broken completely during the second main shock. Low seismicity provides evidence for the lack of other smaller faults in this area. The high seismicity in the southern zone, the area associated with the first rupture and the strong aftershock of M5.5 (occurred on 26 January at 18:45, GCMT solution: strike = 11°/dip = 45°/rake = 120°), could be explained as a result of the main rupture, the existence of smaller faults in the area and the stress transfer after the second strong earthquake. Seismicity in the northernmost (offshore) part of the zone is continuous since the beginning of the sequence, which could be considered as a consequence of triggering of smaller faults in this area.

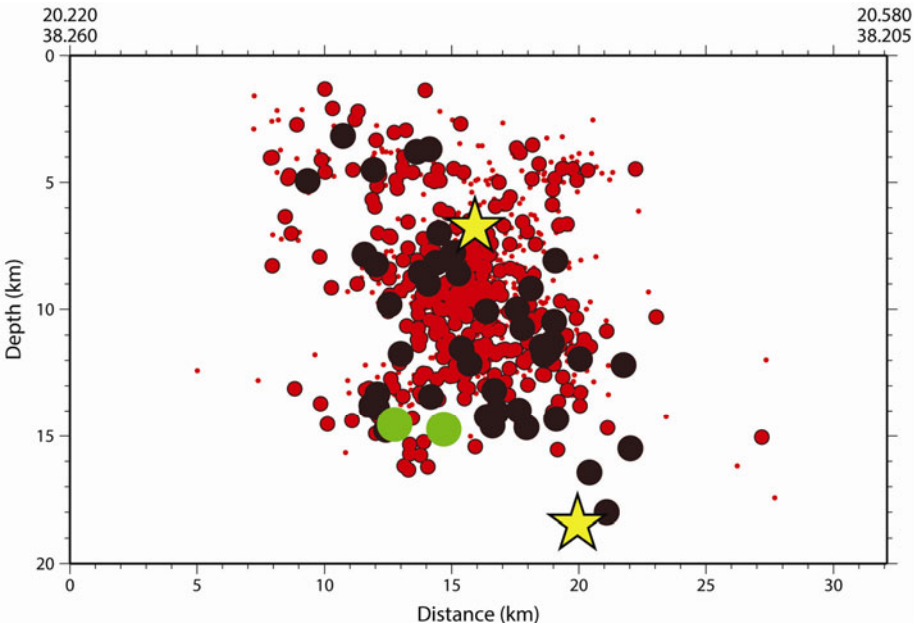


Fig. 6. Cross section normal to the orientation of the aftershock activity. The symbols size is proportional to the earthquake magnitudes. The three larger symbols correspond to earthquakes in the magnitude ranges of 4.0-4.9, 5.0-5.9, and 6.0-6.1.

Figure 6 shows a strike normal cross-section encompassing the aftershocks of both main events. Earthquakes that occurred offshore to the north are not included, since according to our interpretation they are associated with other smaller faults with orientations different than that of the main structure associated with the two main earthquakes. The aftershock zone extends in the range of 2-18 km, dips sharply to the ESE, with the first main shock focus at its lower part. The second main shock occurred at a depth of 7 km, which is probably the reason for the very high values of recorded accelerations. In addition to the main structure, which is delineated clearly from the space distribution of the aftershocks with $M \geq 4.0$, in the cross section the activation of other smaller faults is revealed by the space distribution of the lower magnitude earthquakes as a consequence of the stress transfer due to the coseismic slips of the main shocks. Although it is not completely clear, the strong $M5.5$ aftershock probably occurred on a patch of the main rupture being located between 13-16 km, perhaps in the transition ductile part of the seismogenic layer.

3. FAULT MODEL

Strike-slip faults are commonly segmented at all scales, typically in the form of *en échelon*, non-coplanar faults separated by offsets (or step overs). In our case the step-over zone accommodates continued strike slip displacement between the Kefalonia Fault to the south, and the Lefkada Fault to the north (Fig. 7). This zone comprises smaller parallel fault segments, WSW-ENE striking, almost perpendicular to the *T*-axis orientation at this site, thus controlling the extensional deformation. This set of the closely-spaced parallel seismic lineaments is considered to define strike-slip duplexes and forms a typical transfer zone, between the major Kefalonia and Lefkada Fault segments (shown in Fig. 2), with stepped strike-slip faults and bending in the orientation of the Kefalonia and Lefkada Faults.

In addition to the manifestation of the transfer zone which sheds light on the deformation pattern in the study area, one important component for seismic hazard assessment in particular, is that rupture terminates at this locus, and does not continue further northward. This is repeatedly documented by historical accounts (Papazachos and Papazachou 2003) evidencing that strong earthquakes are associated either with Lefkada or Kefalonia Fault branches separately. The activation of the two branches appeared to be synchronized several times in the past as a result of stress transfer between them (Papadimitriou 2002). One can then assume that the rupture extent has an upper limit that is controlled by the presence of this structure forming bends and stepped strike slip secondary faults, where the orientation of the main fault is deflected.

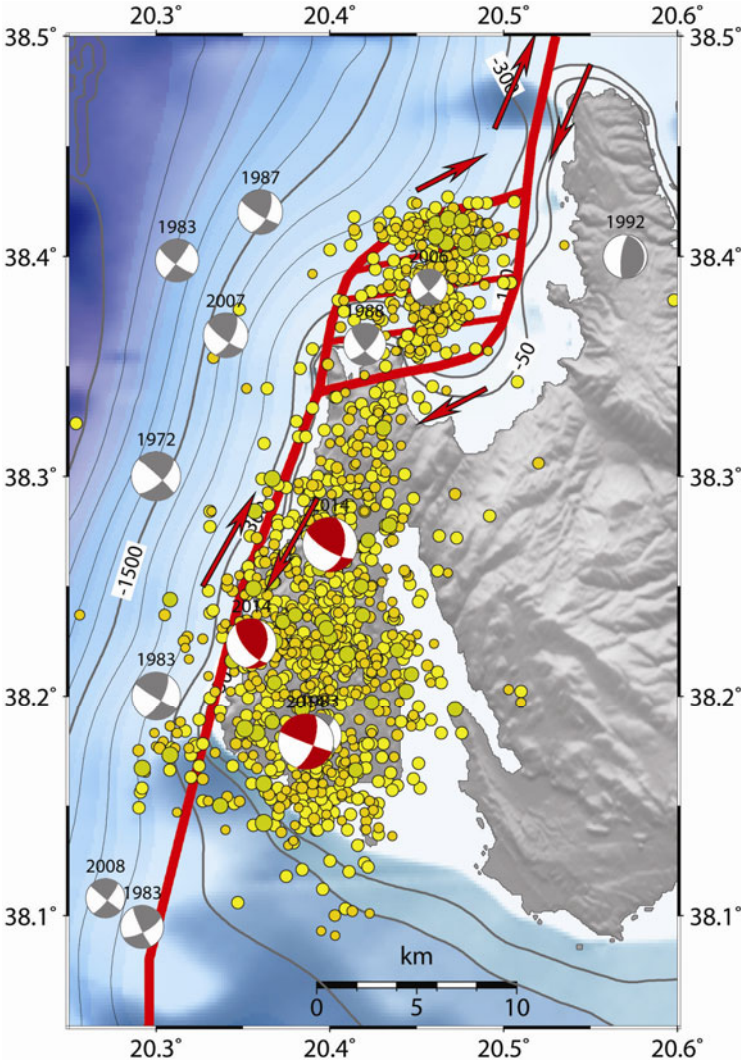


Fig. 7. Aftershock activity (circles) along with inferred fault traces and strong earthquakes fault plane solutions shown as equal area lower hemisphere projections. The compressions quadrants of the stronger earthquakes of the sequence are shown in red. The off fault aftershock activity forms a transfer zone of extensional step overs that connect the Kefalonia Transform Fault to the south with the Lefkada Transform Fault to the north.

4. COULOMB STRESS TRIGGERING

The Coulomb stress change patterns resulting from the two main shocks are shown in Fig. 8. The change in Coulomb failure function (ΔCFF) depends on

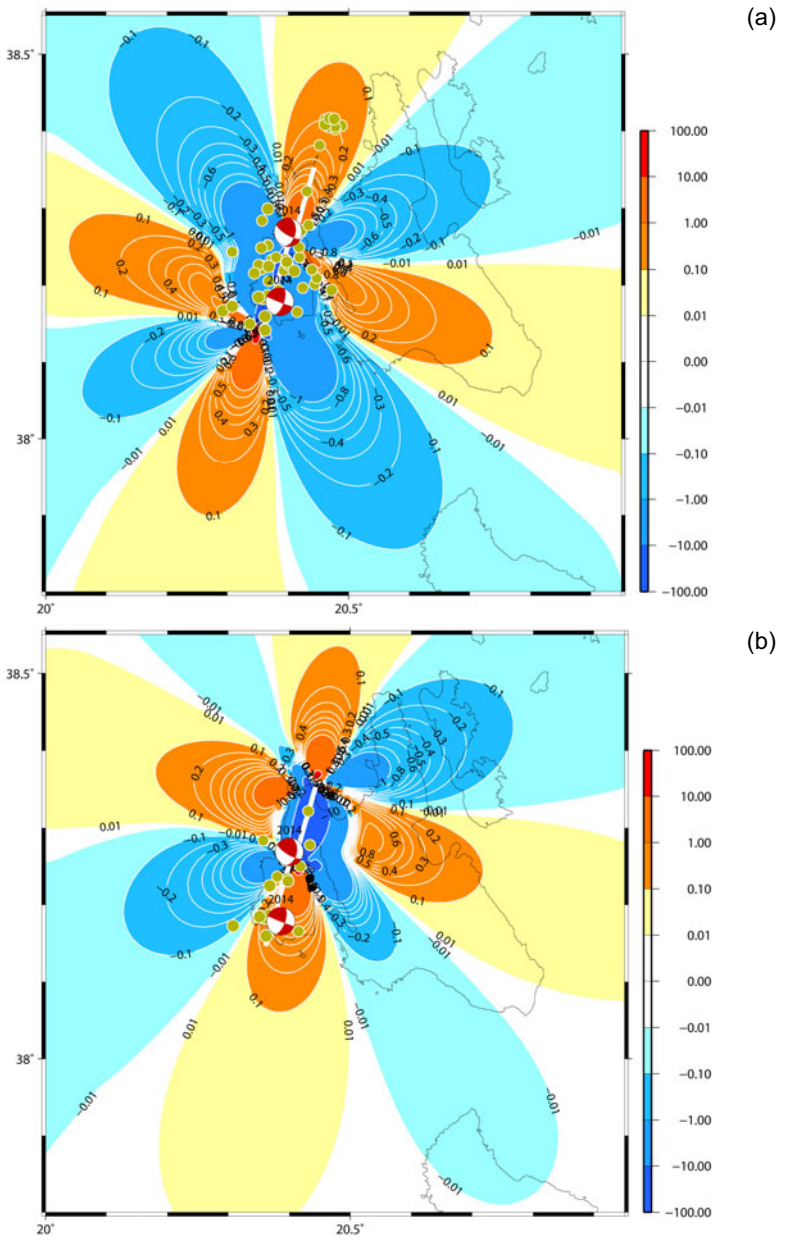


Fig. 8. Coulomb stress changes due to the coseismic slip of: (a) 26 January, and (b) 3 February main shocks, calculated at a depth of 9.0 km. Changes are denoted by the color scale to the right (in bars) and by numbers on the contour lines. The main shock epicenter is depicted by the star and aftershocks by circles, the color and size of which is scaled according to magnitude. The inferred fault trace is shown by the thick white line.

both changes in shear stress, $\Delta\tau$, and normal stress, $\Delta\sigma$, and takes the form $\Delta CFF = \Delta\tau + \mu(\Delta\sigma + \Delta p)$, where Δp is the pore pressure change within the fault, and μ is the friction coefficient which for dry model ranges between 0.6 and 0.8 (Harris 1998 and references therein). Throughout this study we ignore the time-dependent changes in pore fluid pressure and consider only the undrained case (Beeler *et al*. 2000), meaning that Δp depends on the fault-normal stress whereas the fluid mass content per unit volume remains constant. Induced changes in pore pressure resulting from a change in stress under undrained conditions, according to Rice and Cleary (1976) are calculated from $\Delta p = -B(\Delta\sigma_{kk}/3)$, where B is the Skempton's coefficient ($0 \le B \le 1$) and $\Delta\sigma_{kk}$ indicates the summation over the diagonal elements of the stress tensor. We consider a $\mu = 0.75$ and $B = 0.5$, which are close to the case of taking an apparent coefficient of friction $\mu' = 0.4$ suggested by Papadimitriou (2002) who tested different values of μ' for the study area. The shear modulus and Poisson's ratio are fixed as 3.3×10^5 bars and 0.25, respectively. The fault dimensions were defined according to the epicentral distribution and GCMT solutions, as mentioned in a previous section. The centroid moment tensor solutions were adopted for both the faulting type and the value of seismic moment (http://www.ldeo.columbia.edu/~gcmt/). The coseismic slip was then calculated from seismic moment and fault area for both events and the calculations were performed at a depth of 9 km.

Figure 8a evidences high positive stress changes at the location of the northern fault segment, which probably was enhanced and triggered to failure seven days later. The second main event epicenter in particular is located at the southernmost edge of the activated segment, where stress changes have attained the highest positive values (more than 10 bars). The north positive lobe also encompasses the cluster associated with the step over transfer zone. Figure 8b shows that ΔCFF continues to increase at the offshore area; nevertheless, paucity in $M4$ events is observed. Such aftershocks are enhanced in the southern zone instead.

5. DISCUSSION AND CONCLUSIONS

The 2014 Kefalonia doublet (M_w 6.1 and M_w 6.0) with the two main events being separated by seven days in time and about 10 km in space was accommodated by the area comprising Paliki Peninsula and continued offshore northwards, extended over 30-35 km. The rich aftershock production and the spatiotemporal occurrence pattern shed light on the kinematic and structural properties of the activated portions of the KTFZ. The two activated fault segments are compatible with dextral shearing along the Kefalonia Fault branch, the southern portion of which last failed in 1983. The 2014 seismic sequence may be then considered as the spatial continuation of the 1983 sequence with partial coincidence. The aftershock distribution, which ex-

panded up to the northernmost part of the activated area in one day, exhibits a comparatively less active patch, where the second main rupture was accommodated. The aftershock area can be divided into three rupture zones. The southern zone, extending from the first main shock epicenter, covers the southern half of the almost N-S elongated Paliki Peninsula. The second rupture zone starts just north to the first one, evidencing along strike stepping fault segments. To the north of the second zone limit at the northern coast of Paliki and offshore further northwards, a dense cluster appeared almost contemporaneously with the first main shock occurrence. The cluster comprises several $M \geq 4.0$ earthquakes with their number dramatically decreasing after the second main shock occurrence, and can be decomposed in a branch of WSW-ENE striking parallel lineaments.

Both main shocks express north-south dextral strike slip motion, whereas the northern cluster comprises WSW-ENE trending parallel lineaments also of oblique strike-slip motion, dextral with normal component. This constitutes one of the main findings of the current analysis, providing evidence of a transfer zone linking the Kefalonia Fault branch with the Lefkada dextral strike-slip fault branch, which constitutes the northward continuation of Kefalonia Transform Fault Zone.

We used the GCMT focal mechanism solutions and aftershock relocations to construct the source models for calculating the static stress changes due to the two main shocks' coseismic slips. Each earthquake alters the state of stress in its surroundings, and it is advantageous for seismotectonic and seismic hazard recognition purposes to discuss the issue in association with the seismicity manifestation. The offshore cluster triggered by the first main shock occurred in an area where the positive Coulomb stress changes were found to be greater than 0.2 bars, and presented an intense activity encompassing tens of the located aftershocks, several of them having $M \geq 4.0$. This intense activity lasted for about five days and decreased dramatically after the second main shock occurrence, although positive stress changes are accumulated to the ones created by the first main shock. This provides first evidence that the potential of this relay zone cannot exceed moderate magnitude earthquake occurrence. The short duration of the occurrence of stronger aftershocks, might also be the characteristics of the relay zone. The area of Myrtos gulf is characterized by a large number of small magnitude earthquakes in the aftershock period as well as in the last 40 years seismicity (Fig. 2). The area exhibits high activity after the earthquake of magnitude $M5.7$ in 2007. This behavior of seismicity and the absence of any known strong earthquake in the area is an evidence of aseismic movement in the area, which should be investigated using more data. The small magnitude earthquakes are a consequence of rupture of minor geometric anomalies on small faults having different orientations than the main tectonic line. Most

importantly, the later observations set the boundary of the northern extent for the ruptures being originated along the Kefalonia branch of the KTFZ. It also constitutes a major issue for the area's seismic hazard assessment, since it truncates the upper bound of maximum expected earthquake, which is not anticipated to be associated with a rupture exceeding this site to the north.

Acknowledgements. The constructive comments of Sebastiano d'Amico and a second anonymous reviewer, which improved the final version of the paper, are greatly appreciated. Gratitude is also extended to Prof. Telesca for his editorial assistance. Fault plane solutions data used in this paper came from http://www.ldeo.columbia. edu/~gcmt/ and published sources listed in the references. The stress tensors were calculated using a program written by J. Deng (Deng and Sykes 1997), based on the DIS3D code of S. Dunbar, which was later improved (Erikson 1986) and the expressions of G. Converse. The plots were made using the Generic Mapping Tools version 4.5.3 (www.soest.hawaii.edu/gmt, Wessel and Smith 1998). Geophysics Department Contribution 823.

References

Beeler, N.M., R.W. Simpson, S.H. Hickman, and D.A. Lockner (2000), Pore fluid pressure, apparent friction, and Coulomb failure, *J. Geophys. Res*. **105**, B11, 25533-25542, DOI: 10.1029/2000JB900119.

Deng, J., and L.R. Sykes (1997), Evolution of the stress field in southern California and triggering of moderate-size earthquakes: A 200-year perspective, *J. Geophys. Res*. **102**, B5, 9859-9886, DOI: 10.1029/96JB03897.

Erickson, L. (1986), User's manual for DIS3D: A three-dimensional dislocation program with applications to faulting in the Earth, M.Sc. Thesis, Geomechanics Applied Earth Science Dept., Stanford University, Stanford, USA, 167 pp.

Harris, R.A. (1998), Introduction to special section: Stress triggers, stress shadows, and implications for seismic hazard, *J. Geophys. Res*. **103**, B10, 24347-24358, DOI: 10.1029/98JB01576.

Haslinger, F., E. Kissling, J. Ansorge, D. Hatzfeld, E. Papadimitriou, V. Karakostas, K. Makropoulos, H.-G. Kahle, and Y. Peter (1999), 3D crustal structure from local earthquake tomography around the gulf of Arta (Ionian region, NW Greece), *Tectonophysics* **304**, 3, 201-218, DOI: 10.1016/S0040-1951(98)00298-4.

Karakostas, V. (2008), Relocation of aftershocks of the 2003 Lefkada sequence: Seismotectonic implications. **In:** *Proc. 3rd Hellenic Conf. Earthquake En-*

gineering and Engineering Seismology, 5-7 November 2008, *Athens, Greece*, CD ROM, 16 pp.

Karakostas, V.G., and E.E. Papadimitriou (2010), Fault complexity associated with the 14 August 2003 M_w6.2, Lefkada, Greece, aftershock sequence, *Acta Geophys.* **58**, 5, 838-854, DOI: 10.2478/s11600-010-0009-6.

Karakostas, V.G., E.E. Papadimitriou, and C.B. Papazachos (2004), Properties of the 2003 Lefkada, Ionian Islands, Greece, earthquake seismic sequence and seismicity triggering, *Bull. Seismol. Soc. Am.* **94**, 5, 1976-1981, DOI: 10.1785/012003254.

Karakostas, V.G., E.E. Papadimitriou, Ch.K. Karamanos, and D.A. Kementzetzidou (2010), Microseismicity and seismotectonic properties of the Lefkada–Kefalonia seismic zone, *Bull. Geol. Soc. Greece* **43**, 2053-2063.

Karakostas, V., E. Karagianni, and P. Paradisopoulou (2012), Space-time analysis, faulting and triggering of the 2010 earthquake doublet in western Corinth Gulf, *Nat. Hazards* **63**, 2, 1181-1202, DOI: 10.1007/s11069-012-0219-0.

Karakostas, V., E. Papadimitriou, and D. Gospodinov (2014), Modelling the 2013 North Aegean (Greece) seismic sequence: geometrical and frictional constraints, and aftershock probabilities, *Geophys. J. Int.* **197**, 1, 525-541, DOI: 10.1093/gji/ggt523.

Kiratzi, A.A., and C.A. Langston (1991), Moment tensor inversion of the 1983 January 17, Kefallinia event of Ionian islands (Greece), *Geophys. J. Int.* **105**, 2, 529-535, DOI: 10.1111/j.1365-246X.1991.tb06731.x.

Klein, F.W. (2000), User's guide to HYPOINVERSE-2000, a Fortran program to solve earthquake locations and magnitudes, Open File Rep. 02-171, Ver. 1.0, U.S. Geological Survey, Menlo Park, USA.

Kokinou, E., E. Papadimitriou, V. Karakostas, E. Kamberis, and F. Vallianatos (2006), The Kefalonia Transform Zone (offshore Western Greece) with special emphasis to its prolongation towards the Ionian Abyssal Plain, *Mar. Geophys. Res.* **27**, 4, 241-252, DOI: 10.1007/s11001-006-9005-2.

Louvari, E., A.A. Kiratzi, and B.C. Papazachos (1999), The Cephalonia Transform Fault and its extension to western Lefkada Island (Greece), *Tectonophysics* **308**, 1-2, 223-236, DOI: 10.1016/S0040-1951(99)00078-5.

Paige, C.C., and M.A. Saunders (1982), Algorithm 583: LSQR – sparse linear equations and least squares problems, *ACM Trans. Math. Software* **8**, 2, 195-209, DOI: 10.1145/355993.356000.

Papadimitriou, E.E. (1993), Focal mechanism along the convex side of the Hellenic Arc, *Boll. Geof. Teor. Appl.* **35**, 140, 401-426.

Papadimitriou, E.E. (2002), Mode of strong earthquake recurrence in the central Ionian Islands (Greece): Possible triggering due to Coulomb stress changes generated by the occurrence of previous strong shocks, *Bull. Seismol. Soc. Am.* **92**, 8, 3293-3308, DOI: 10.1785/0120000290.

Papadimitriou, E.E., and B.C. Papazachos (1985), Evidence for precursory seismicity patterns in the Ionian Islands (Greece), *Earthq. Predict. Res.* **3**, 95-103.

Papazachos, B.C., and C.C. Papazachou (2003), *The Earthquakes of Greece*, Ziti Publ., Thessaloniki, 304 pp.

Rice, J.R., and M.P. Cleary (1976), Some basic stress diffusion solutions for fluid-saturated elastic porous media with compressible constituents, *Rev. Geophys*. **14**, 2, 227-241, DOI: 10.1029/RG014i002p00227.

Scordilis, E.M., G.F. Karakaisis, B.G. Karakostas, D.G. Panagiotopoulos, P.E. Comninakis, and B.C. Papazachos (1985), Evidence for transform faulting in the Ionian Sea: The Cephalonia Island earthquake sequence of 1983, *Pure Appl. Geophys.* **123**, 3, 388-397, DOI: 10.1007/BF00880738.

Waldhauser, F. (2001), HypoDD – A program to compute double-difference hypocenter locations, Open File Rep. 01-113, U.S. Geological Survey, Menlo Park, USA.

Waldhauser, F., and W.L. Ellsworth (2000), A double-difference earthquake location algorithm: Method and application to the Northern Hayward Fault, California, *Bull. Seismol. Soc. Am*. **90**, 6, 1353-1368, DOI: 10.1785/0120000006.

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

Received 8 April 2014
Received in revised form 9 June 2014
Accepted 9 June 2014

DE GRUYTER OPEN

Acta Geophysica
vol. 63, no. 1, Feb. 2015, pp. 17-44
DOI: 10.2478/s11600-014-0255-0

Seismic Monitoring of Poland – Description and Results of Temporary Seismic Project with Mobile Seismic Network

Jacek TROJANOWSKI, Beata PLESIEWICZ, and Jan WISZNIOWSKI

Institute of Geophysics, Polish Academy of Sciences, Warszawa, Poland
e-mail: jtroj@igf.edu.pl

A b s t r a c t

The paper describes a temporary seismic project aimed at developing the national database of natural seismic activity for seismic hazard assessment, officially called "Monitoring of Seismic Hazard of Territory of Poland" (MSHTP). Due to low seismicity of Poland, the project was focused on events of magnitude range 1-3 in selected regions in order to maximize the chance of recording any natural event. The project used mobile seismic stations and was divided into two stages.

Five-year measurements brought over one hundred natural seismic events of magnitudes M_L range 0.5-3.8. Most of them were located in the Podhale region in the Carpathians. Together with previously recorded events this made it possible to conduct a preliminary study on ground motion prediction equation for this region. Only one natural event, of magnitude $M_L = 3.8$, was recorded outside the Carpathians in a surprising location in central-west Poland.

Key words: seismic monitoring, mobile network, seismicity of Poland, GMPE.

1. INTRODUCTION

Although Poland is known as a region of very low natural seismicity, some earthquakes occur there from time to time. The historical catalogue (Guterch B. 2009) consists of less than one hundred earthquakes in the time span of almost one thousand years (Fig. 1).

There are two main regions of natural seismicity − mountains in south Poland and Teisseyre–Tornquist Zone (TTZ), which passes through Poland as about 100 km wide band from NW to SE.

The TTZ is a passive contact zone between two stable platforms − the Precambrian East European Craton (EEC) and the Paleozoic West European Platform (WEP). This complicated and very interesting structure is very difficult to study because it is covered by thick sediments. Most of our knowledge about it comes from three wide-angle experiments covering Poland and vicinity with many 2D profiles: POLONAISE'97 (Guterch A. *et al.*

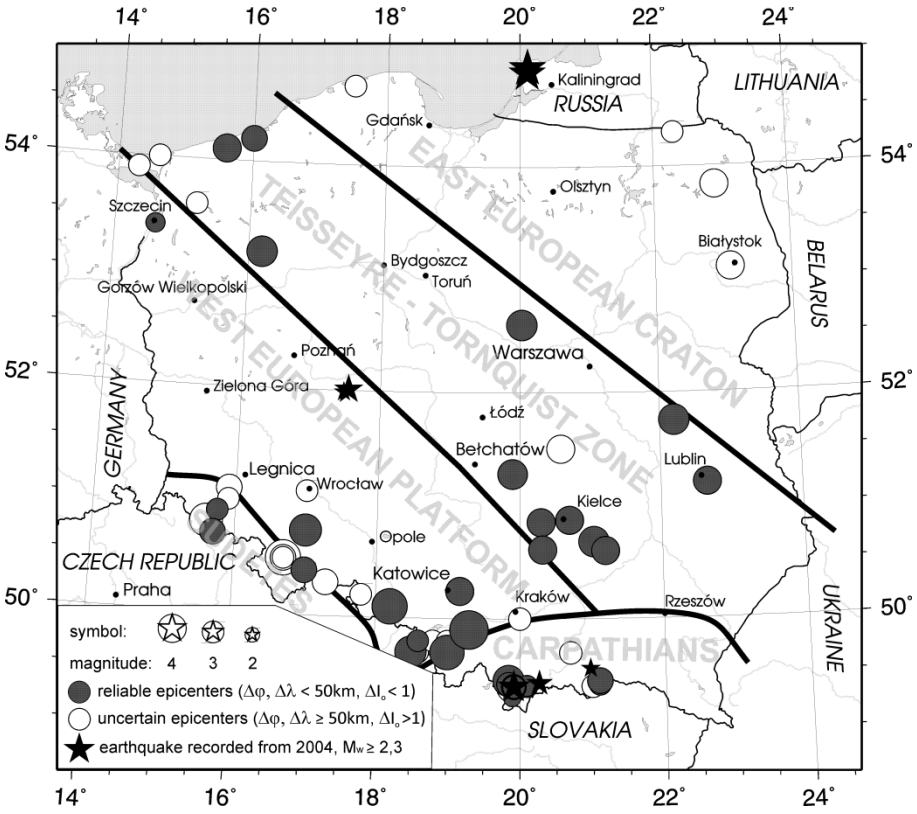


Fig. 1. Maps of the historical seismicity and strong recent earthquakes from 1400 to 2012 (data after Guterch B. (2009) and MSHTP).

1999), CELEBRATION 2000 (Guterch A. *et al.* 2003), and SUDETES 2003 (Grad *et al.* 2003).

A stable area of central and north Poland is pushed from the south by the Carpathians, which is manifested in mountains in south and south-east Poland. The Carpathians is young active orogen, contrary to older, Variscan Sudetes in south-west Poland.

Regardless of these significant genetic differences, south Poland in general is a site of the majority of earthquakes including the biggest ones, with maximum intensity of 7 in EMS-98 scale. The reason for using the EMS intensity scale, instead of magnitude, is that the biggest earthquakes occurred before any form of seismic measurement was set up in the area.

In other parts of Poland, there were much fewer earthquakes and they were smaller, hardly ever reaching intensity 6. But there is a spectacular exception of a recent well described earthquake – 5.2 magnitude earthquake on 21 September 2004 in Kaliningrad Oblast, about 50 km from the Polish-Russian border (Wiejacz 2006, Domański 2007). It was felt in the whole north Poland causing minor losses. It was also felt in Sweden and even Denmark. Together with another earthquake in southern Poland the same year, it was probably one of the important reasons for establishing the project focused on natural seismicity.

The project assumptions were based on historical seismicity (Guterch 2009) and seismic hazard study for Polish area (Schenk *et al.* 2000). Parameters of seismic hazard made it possible to assess that earthquake occurrence rate in the regions being considered is high enough to record at least a few events of magnitudes above $M_L = 1$. Because of too large area to monitor with 24 stations assigned to the project, it was decided to monitor only selected areas with the highest occurrence rate of earthquakes. Additionally, the project was divided into two 2.5-year stages. Some practical experience was gained during realization of PASSEQ experiment in the years 2006-2008 (Wilde-Piórko *et al.* 2008) which comprised almost 200 stations in temporary locations. On the contrary, though, in MSHTP the stress was put on mobility and immediate data transfer. In case of an exceptionally large event (above $M_L = 3$) it was assumed that some of the stations have to be able to be moved in two days from current positions to the epicentral area.

Although the project was focused on the natural seismicity only, there are regions of Poland with relatively high induced seismicity. The strongest earthquakes are induced by copper mines near Legnica and Głogów in south-west Poland and can exceed magnitude $M_L = 4$, which happens almost every year (*e.g.*, Lizurek and Wiejacz 2011, Orlecka-Sikora *et al.* 2012, Idziak and Dubiel 2011). Other regions highly influenced by induced seismicity is Upper Silesia in south Poland, where big coalfields have been intensively exploited for last centuries (Zuberek and Jochymczyk 2010) and Bełchatów

brown coal open-cast mine region (Wiejacz and Rudziński 2010). Events in these regions were not analyzed. There exist local networks maintained by mines and controlled by appropriate authorities, which monitor regions of induced seismicity.

2. INSTRUMENTATION

Mobile seismic network requires transportable equipment which can be easily deployed without much effort to prepare the site. The Institute of Geophysics, Polish Academy of Sciences (IGF PAS) decided to use its long experience in the field of developing data loggers (*e.g.*, Aleksandrowicz 1982, Hościłowicz *et al.* 1990, Olszewski and Wiszniowski 1993) and launched a new Net Data Logger (NDL) in 2008. The NDL served in the project with a sampling rate of 100 sps and dynamics of 132 dB. Continuously recorded data were stored on Compact Flash in the internal data format and immediately transferred through the Internet provided by GSM operators. The NDL together with external devices form a mobile station.

All stations were equipped with three-component short-period seismometers Lennartz LE-3DLite (1 Hz), which are appropriate to measure local and regional seismicity, as the main content of seismic waves comes in the range of a few Hz. Additionally, this type of seismometers is easy to handle and does not require time-consuming procedure as most long period seismometers (*e.g.*, STS-2), which is very important for projects requiring high mobility of stations.

Data downloaded from the stations are collected and archived by the SeisComp system (www.seiscomp3.org/wiki/doc). During the whole project, SeisComp has been extended and supplemented by our components which support handling a seismic network. It comprises a set of tools to control network status and check data quality.

For data processing there was chosen a Seismic Wave Interpretation Program (SWIP) developed by IGF PAS for the purpose of routine job in seismological observatories (private.igf.edu/~jwisz). It has a direct connection to MySQL data base of events and to all recorded data (via ArcLink protocol).

3. DETECTION METHODS

At first, only visual inspection of seismograms was done but it shortly appeared to be time consuming and not reliable. As the acquisition was based on SeisComp system, tools built in this system were tested − AutoPick and AutoLoc. Unfortunately, they are meant for other recording conditions. AutoPick is based on the ratio of Short Term Average to Long Term Average (STA/LTA), which makes it vulnerable to high amplitude disturbances

and generates many false detections which makes these tools inapplicable for our data.

The problem was that most stations were deployed in temporary locations close to human neighborhoods, which causes a high level of noise and disturbances in recorded seismic signal. Therefore, detection of small events is associated with unacceptable number of false detections.

It was decided to apply Real Time Recurrent Neural Network (RTRN) to detect small natural seismic events. It had already been studied on regional events by Wiszniowski (2000) but it got accommodated for local events (Wiszniowski *et al.* 2014). This method is able to assess relations of seismic signal in frequency domains as well as in time of seismic phases.

4. MEASUREMENTS

The seismic network used in MSHTP is registered in IRIS with a name PD – Polish Seismic Monitoring Network, but as a mobile network; individual station names have not been reserved. The full list of all stations which worked in the project is listed in the Appendix 1 and shown on maps in Figs. 2 and 3 for the first and the second stage of the project, respectively.

According to the contract with the project founder, in case of appearance of a natural earthquake of significant scale, some stations should be moved to the epicenter area in 48 hours to monitor potential aftershocks. For this reason, stations are designed and deployed so as to make mobilization and demobilization easy, without expensive and time-consuming site preparation.

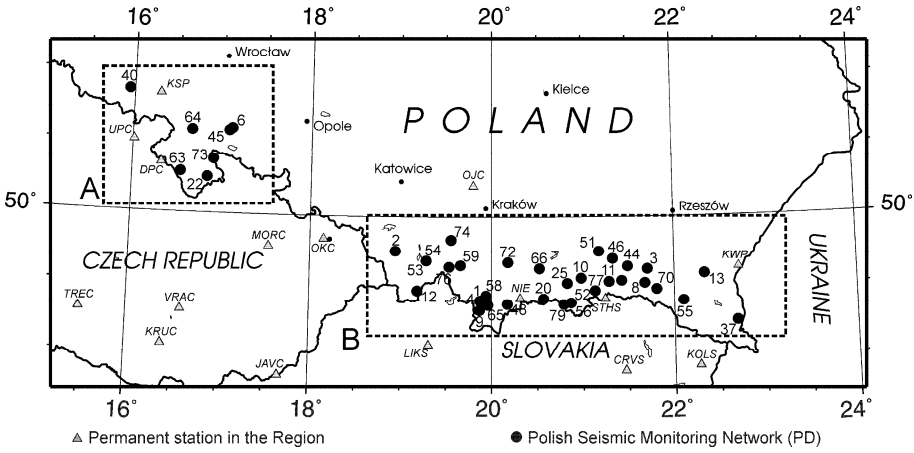


Fig. 2. Locations of stations in the Sudetes (region A) and Carpathians (region B) during the first stage of MSHTP in 2008-2010 (circles). Other stations in the region are marked by triangles.

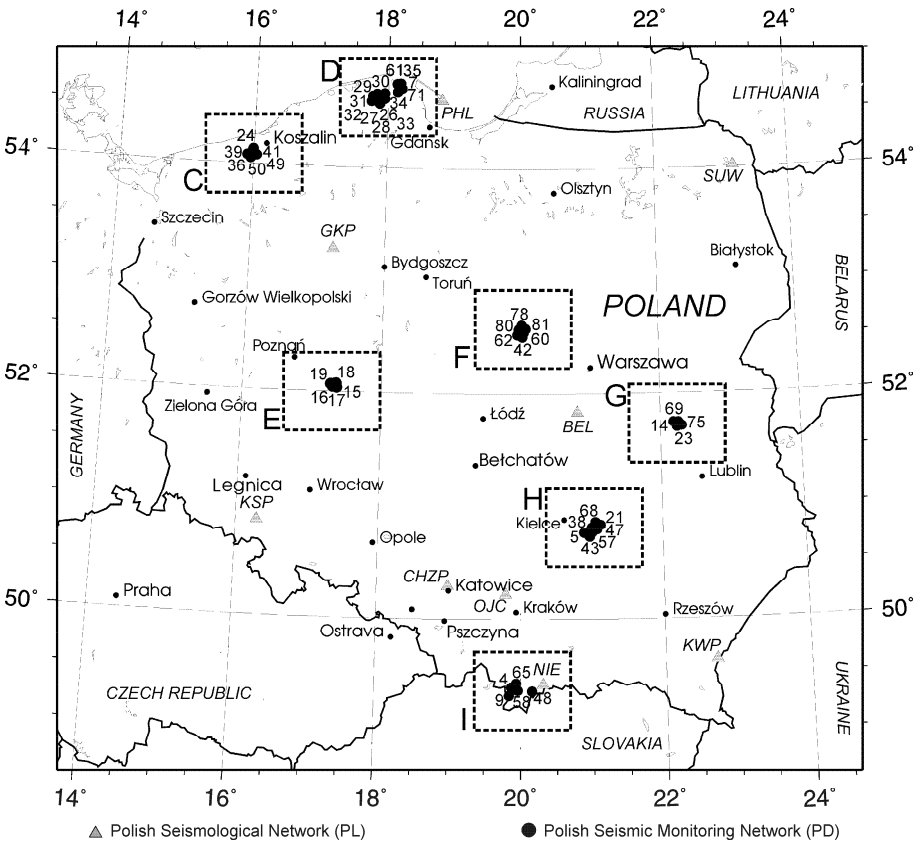


Fig. 3. All seismic stations maintained by IGF PAS during the second stage of MSHTP in the years 2010-2012 (circles). Other stations in the region are marked by triangles. Monitored regions: C – West Pomerania / near Koszalin, D – Pomerania / vicinity of Łebień and Żarnowiec, E – Wielkopolska / near Jarocin, F – Mazovia / near Płock, G – Mazovia / near Łuków, H – Holly Cross Mountains, I – Podhale / Tatra Mountains.

The project was focused on areas of known historical seismic activity. In the first stage, the monitoring covered mountain regions of the Sudetes and Carpathians in south Poland, where the majority of historical earthquakes occurred (Fig. 2). In the second stage, several regions of central and north Poland were selected (Fig. 3). It was assumed that in these places it is most probable to find seismic activity, although among these regions the region of Holy Cross Mountains has exceptionally high seismic hazard parameters (Schenk *et al.* 2000) which suggested very high chance of recording natural events. Additionally, three stations have remained during second stage in

Podhale/Carpathians, where continuous seismicity was discovered in the first stage.

Locations were selected in order to have good coverage of the monitored area and to avoid noisy areas. The second requirement means in general that stations should be put far from highly populated areas but it entails problems with infrastructure which is necessary to provide a station with power supply and internet access. Especially in mountains there was often a lack of GSM signal, which was necessary to transmit data. In most cases, stations were installed in private properties to assure protection and power supply. Although station locations were carefully selected, no tests were performed prior to station deployment. As a result, it was often necessary to move stations because of disturbances or high noise level which appeared after a station had been deployed. However, this was relatively easy due to simple installation procedure. Such a trial-error method led to 38 station locations during the first stage and 46 during the second stage.

Signal quality in different regions of Poland

Recording conditions in Poland vary because of differences in population density, industrialization, and geology. In general, low noise level is in southern Poland – in mountains. In central and northern Poland there are thick sediment layers and soils without any outcrops which are associated with higher noise level. Site selection and verification was carried out for every potential seismic station. A very useful criterion for noise assessment was to analyze power spectrum density of the recorded ground velocity. Sometimes high amplitude noise is concentrated around particular frequencies or frequency bands, which makes it possible to filter it out. This is especially easy for most signals generated by machines. The worst possible kind of noise is that related to the whole band of seismic waves (a few Hz), which makes it impossible to filter it out without significant loss of the seismic signal. Useful information about human generated noise comes from a comparison of day and night spectra, as human activity is usually higher during a day.

A comparison of power spectrum density for different regions of Poland is presented in Fig. 4. Day and night spectra were calculated for all stations and then averaged to represent respective regions. All spectra are plotted in the same scale to make it possible to easily compare noise levels between regions.

The best signal is in mountain regions, for which velocity power spectra density stays below 10^{-16} $(m/s)^2/Hz$ (regions A and B) or 10^{-15} $(m/s)^2/Hz$ (region H). All other regions with sedimentary background have values varying between 10^{-15} $(m/s)^2/Hz$ and 10^{-13} $(m/s)^2/Hz$.

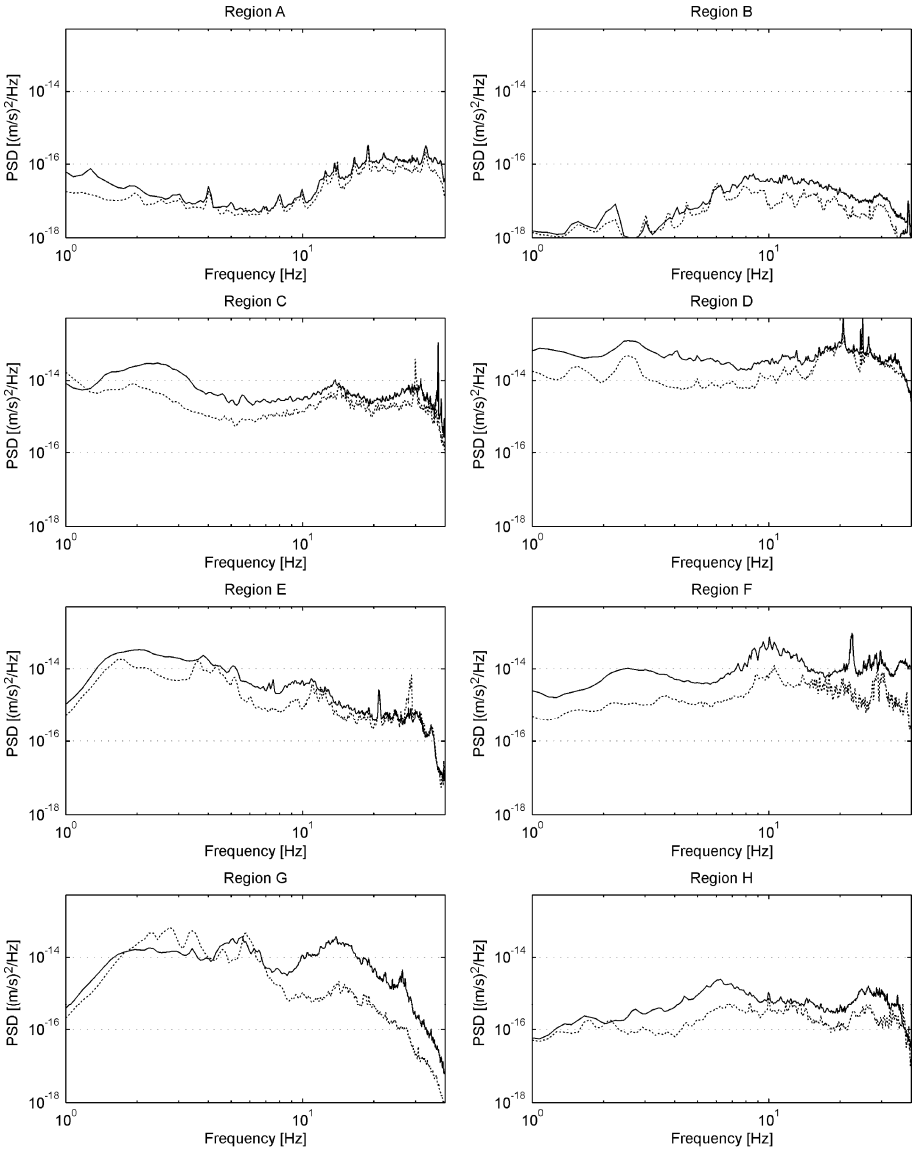


Fig. 4. Velocity power spectra density for stations in regions A-H shown in Figs. 2 and 3. Day spectra – solid line, night spectra – dashed line. Few days with small wind speed (average day wind speed below 3 m/s) were selected in each region. Day was defined as a time between 6 a.m. and 6 p.m. local time.

5. RESULTS

General description

The first stage of MSHTP covered southern Poland where the probability of earthquake occurrence is the highest. Indeed, it has been confirmed that Podhale is seismically active (over 100 events). There were also two micro-earthquakes recorded near Krynica in Beskid Sądecki. In other places monitored during the first stage of MSHTP, no natural event was recorded.

Places of previously recognized seismic activity in TTZ became the object of seismic monitoring in the second stage of MSHTP. Most of the stations were put in regions with thick sediment and soil layers which makes the noise level high. From this point of view, stations in Holy Cross Mountains had good recording conditions but a problem was with high activity of quarries in the region. To distinguish its records from natural events, source location and spectrograms of the recorded signal were analyzed. Finally, during the entire MSHTP no natural seismic event was found in the whole TTZ.

The last not described region of Poland is Wielkopolska in central-western Poland. It is a part of WEP and by January 2012 there was no single earthquake known there. Surprisingly, on 6 January 2012 at 15:38 UTC an earthquake of magnitude $M_L = 3.8$ frightened inhabitants of the area in a radius of 10 km from the epicenter. Later, macroseismic questionnaires were coming from distances over 60 km from the epicenter. Unfortunately, this region was not monitored and the nearest stations were about 100 km off. Immediately after the earthquake, five stations were moved there from other regions. The earthquake was described in detail by Lizurek *et al.* (2013).

Apart from seismic events, both natural and man-induced, there were also recorded non seismic events considered generally as disturbances but they are rarely recorded by more than one station. If so, it is necessary to verify a possible source to distinguish it from possibly natural seismic source. Events of this kind may be, for example, quarry blasts or jet sonic waves.

The next two sections describe in more details natural seismicity recorded in the project and an interesting example of non-seismic events recorded on an exceptionally large area. The last section concerns derivation of ground motion prediction equation (GMPE) for the Podhale region / Carpathians.

Seismicity of the Carpathians (Polish part)

Earthquakes felt in the Polish Carpathians were known for many years. Some of them are only mentioned in chronicles but more recent ones are better described (Guterch B. *et al.* 2005, Guterch B. 2006, 2007, 2009, Wiejacz and Dębski 2009). Seismicity concentrates in the west part of the region – in

the Podhale but some events occur also in the east part, mainly in Beskid Sądecki (Fig. 5). Before 2008, seismic events were rarely recorded but the MSHTP project has shown a continuous activity. Figure 6 shows an increasing capability to detect seismic events in this region with the majority of events recorded during the MSHTP project. In the Podhale region, where

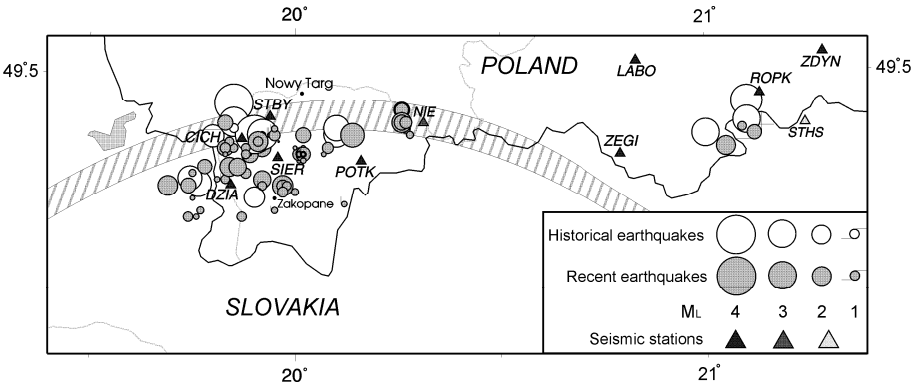


Fig. 5. Seismic events in Podhale. The most important geological structure is PKB – Pieniny Klippen Belt (hatched area). Seismic stations are marked with triangles.

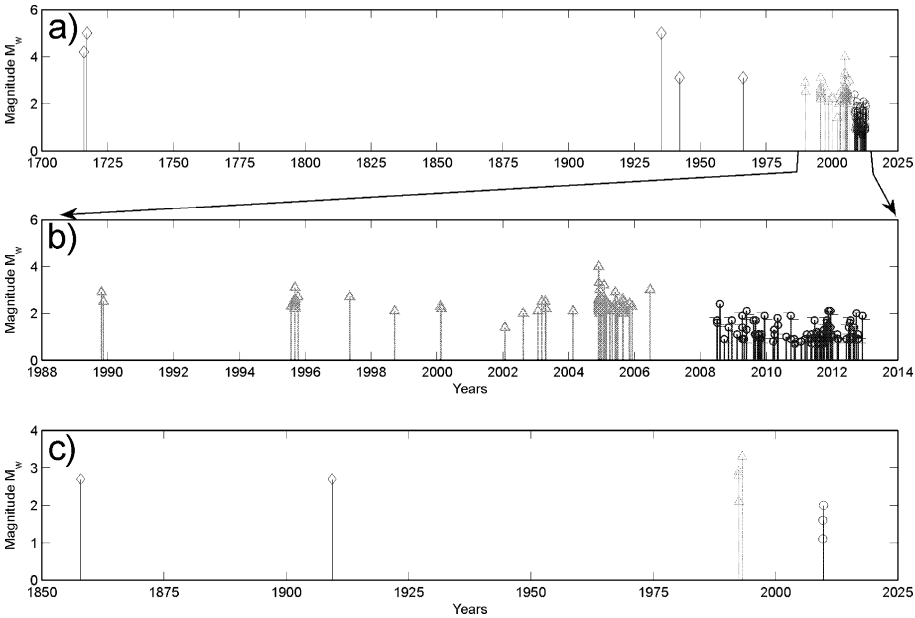


Fig. 6. Time sequence of all earthquakes in the Polish Carpathians separated into the Podhale region – panels (a) and (b), and Beskid Sądecki – panel (c). Diamonds – historical earthquakes, triangles – recorded instrumentally, and circles – recorded during MSHTP.

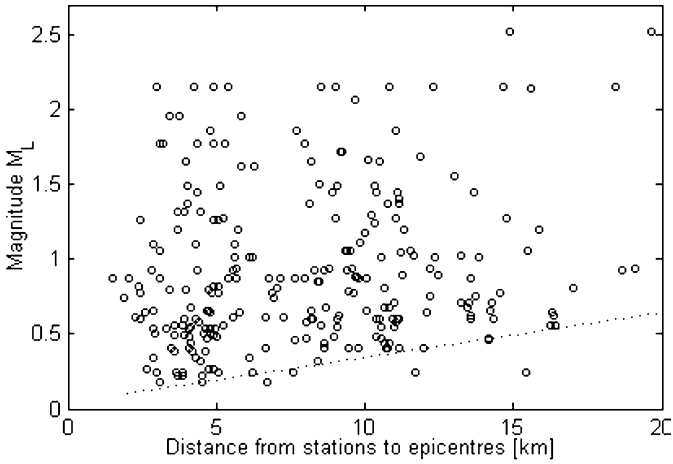


Fig. 7. The dashed line on the diagram shows a sensitivity of the network in the Podhale region. Each circle represents a magnitude calculated for an individual station.

several seismic events were recorded, a sensitivity of the seismic network, in terms of a minimum magnitude possible to be detected at a given distance, is shown on a diagram in Fig. 7. Generally, events below a dashed cut-off line are not detectable by this network, although in very good recording conditions it happens.

Over one hundred microearthquakes have been detected and 81 of them, of magnitude (M_L) range from 0.5 to 2.2, located (Appendix 2). Some of them are part of a swarm which took place in November–December 2011. All of the swarm events are located in the Pieniny Klippen Belt (PKB) formation and may be related to a nearby Czorsztyńskie Lake – an artificial reservoir on the Dunajec River, whereas the majority of other events are located to the south of PKB and are not related to the lake. In general, active regions fit well with regions where stronger events were recorded previously (Fig. 5).

Although the number of records was not sufficient to conduct full moment tensor analysis, it was possible to derive basic spectral parameters. The values of Ω_0 and f_0 of Brune's (1970) model were calculated separately for three components using the method of Andrews (1986). The spectrum of the signal was computed by the multitaper method of Park *et al*. (1987) with scaling of the spectra based on the Parseval equality. The seismic moment was calculated from three components (Wiejacz and Wiszniowski 2006).

There were selected three events of magnitude $M_L = 1$ and three of $M_L = 2$. A corner frequency (f_0) calculated for different stations was for $M_L \sim 1$, in the range of 5.9-11.9 Hz for *P* waves and 3.6-7.2 Hz for *S* waves. For

stronger events, of magnitude $M_L \sim 2$, the calculated f_0 was in the range of 4.1-11.7 Hz for *P* waves and 3.3-7.3 Hz for *S* waves. Spectral parameters for events of $M_L = 1$ are less reliable because, as a result of inelastic dumping, high frequency signal is below the noise level.

Ground Motion Prediction Equation (GMPE) for Podhale

Relatively high number of seismic events recorded in Podhale / West Polish Carpathians allowed us to derive GMPE for this region. The only equation used before was a very general equation used by Shenk *et al.* (2000) to jointly describe Czech, Polish, and Slovak region.

At first, a standard regression model of GMPE was used, which assumes geometric damping

$$\log Y = a_1 + a_2 m + a_3 \log\sqrt{d^2 + h^2} \quad , \tag{1}$$

where Y is the Peak Horizontal Acceleration (PHA) [m/s^2], m the magnitude, and d the distance [km]. Coefficients a_1, a_2, a_3, and h were estimated using three methods (Joyner and Boore 1993): single stage regression (SSR), two stage regression (TSR), and least squares. The results are presented in Table 1.

Table 1

Coefficients of Eq. 1 computed using different regression methods

Regression method	a_1	a_2	a_3	h	σ
SSR	–1.6957	1.1251	–2.7167	7.0646	0.255 (γ = 0.009)
TSR	–2.1658	1.1105	–2.4095	4.8663	0.25 (σ_e = 0.096)
Least squares	–2.0945	1.0985	–2.4438	5.2665	0.25

All three methods gave similar results within 30% confidence interval. Then, stability of the solution was tested for different distances to the epicenter (Table 2). Small variation of the parameters at every distance indicates good fit to the data, which can also be visually inspected in Fig. 8.

Apart from the standard form of the GMPE of Eq. 1, a GMPE with assumed anelastic damping was also tested. It is given by the general equation

$$\log Y = a_1 + a_2 m - \log\sqrt{d^2 + h^2} + a_3 d \tag{2}$$

and gave very similar results, in terms of σ value, to previous GMPE model for all source-receiver distances, although data extended by older events does not fit as well as previously, having higher σ-value (Table 3). This means that the previous GMPE model with elastic damping is better for the region of Podhale.

Table 2

Coefficients of the GMPE equation
calculated for events within different distances to the epicenter

Distance limit	Data points	a_1	a_2	a_3	h	σ
5	88	–2.4716	1.1782	–2.2901	5.2500	0.2055
10	165	–2.6556	1.165	–2.0001	6.9019	0.2262
15	230	–2.5984	1.1191	–2.0217	4.99	0.2262
20	244	–2.5411	1.1133	–2.084	5.6697	0.2239
30	256	–2.6027	1.1155	–2.0144	4.195	0.2208
50	269	–2.3774	1.0868	–2.2392	6.6336	0.2336
100	280	–2.3652	1.0818	–2.2476	5.7368	0.2399
100	288*	–2.3601	1.0741	–2.2457	5.2665	0.2529

*)including stronger events recorded prior to MSHTP project

Table 3

Coefficients of the GMPE Eq. 2
calculated for events within different distances to the epicenter

Distance limit	Data points	a_1	a_2	a_4	h	σ
5	88	–3.0214	1.1784	–0.0846	5.5441	0.2056
10	165	–2.8465	1.165	–0.0779	9.225	0.2256
15	230	–3.217	1.1192	–0.0429	3.7152	0.2262
20	244	–3.1949	1.1155	–0.0445	3.9287	0.2229
30	256	–3.3544	1.1161	–0.0301	2.2112	0.223
50	269	–3.3029	1.108	–0.034	2.7008	0.2254
100	280	–3.4162	1.0913	–0.0235	1.182	0.2456
100	288*	–3.4941	0.9547	–0.0062	0	0.3342

*)including stronger events recorded prior to MSHTP project

Finally, a model of the form of Eq. 1 obtained with TSR regression method ($a_1 = –2.17$, $a_2 = 1.1$, $a_3 = –2.4$, $h = 4.87$) was compared with other GMPE's. There was chosen an equation of Schenk *et al.* (2000) which was dedicated for the whole area of Czech, Poland, and Slovakia and five other equations used in the project SHARE (Delavaud *et al.* 2012) to describe central European region. An example plot for an event of magnitude $M_w = 4$ is shown in Fig. 9. For small distances, our solution gives higher values than other models except the one by Campbell (2003), but for distances over 10 km from the epicenter it gives the highest values. The most probable rea-

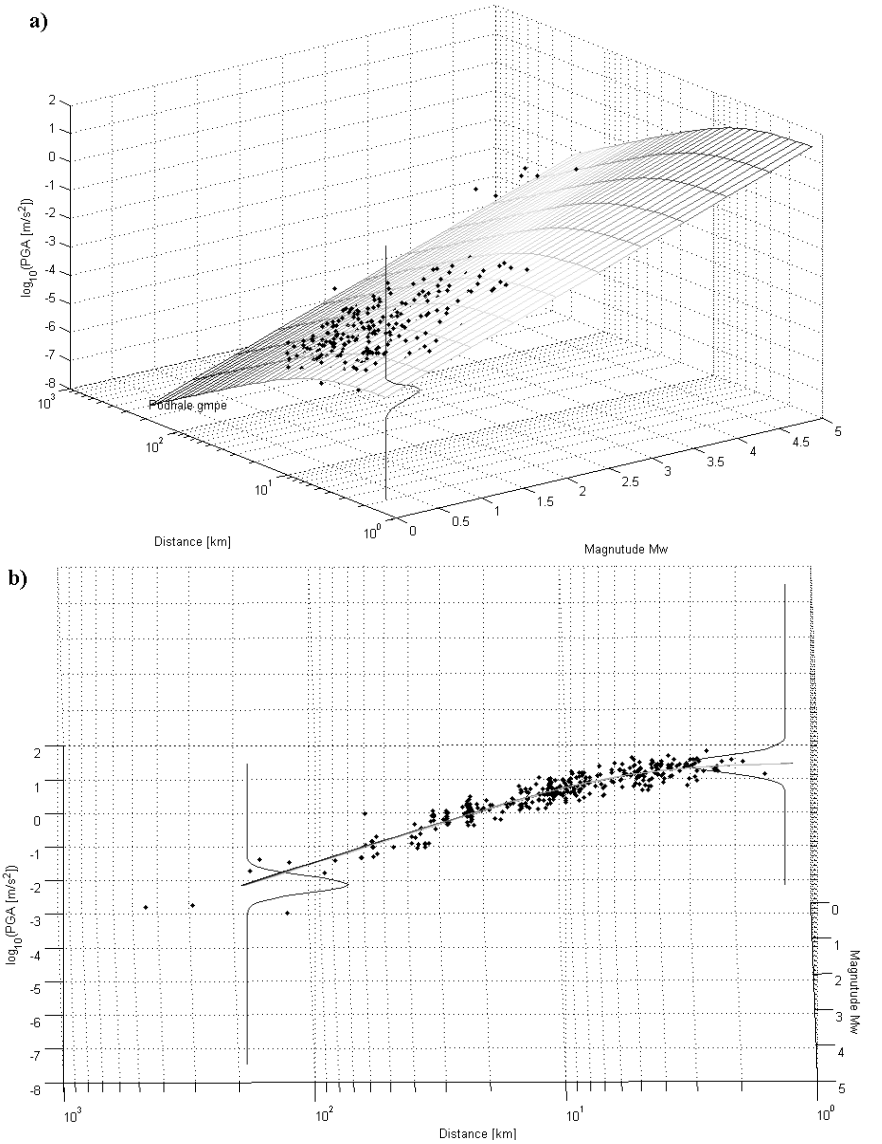


Fig. 8. TSR regression model as a surface in 3D plotted in two projections, (a) and (b), together with recorded peak horizontal acceleration – black dots. Uncertainty of the model is given by the probability distribution plotted in a vertical axis.

son for this difference from other models is that the majority of events in Podhale were of small magnitude and at small distance, which makes a weak fit for strong and distant events. To obtain a better GMPE for Podhale it is necessary to record stronger events with a wide spectrum of distances.

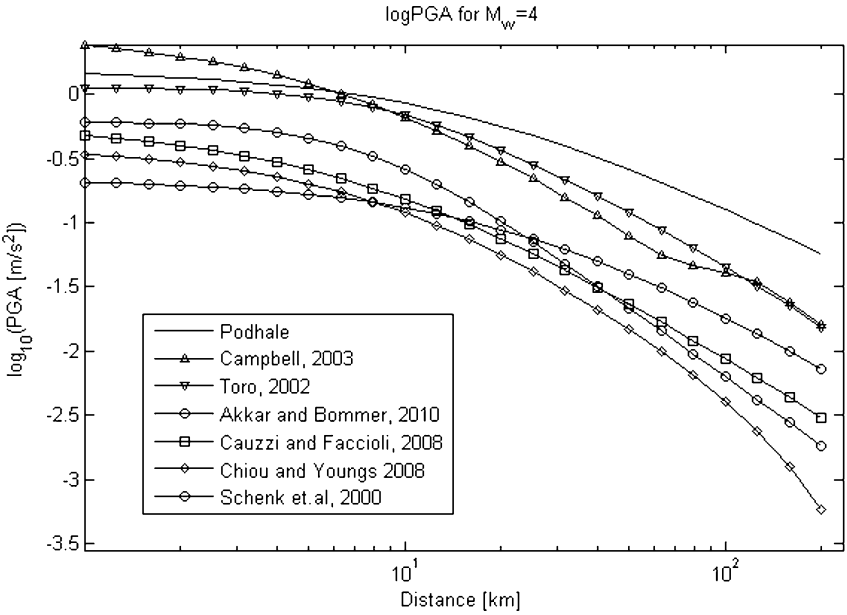


Fig. 9. Comparison of different GMPE's for an event with magnitude $M_w = 4$.

Non-seismic events

When continuous measurements are carried out, there are many non-seismic events recorded by seismic stations. Usually they are local and are recorded only by one station, so they are not even detected by algorithms. Sometimes, such signals are recorded by more stations but still in one region. It can be, for example, a sonic wave caused by an explosion or a shock wave of a supersonic jet. It is easy to distinguish such an event from seismic one because it has much smaller propagation velocity.

An interesting event happened on 25 February 2011, but there were more very similar ones. A set of regularly spaced (2 min interval) impulses was recorded by some stations in very distant regions (over 300 km). The best records for vertical components are presented in Fig. 10 and associated station locations in Fig. 11.

Joint epicenter location and velocity inversion derived a velocity of ~350 m/s which fits sonic wave speed very well. The locations obtained for the first impulse and the last impulse are almost the same and are very close to $\varphi = 54.50$N, $\lambda = 20.83$E (Fig. 11). This suggests that the source was immobile and was on the territory of Kaliningrad Oblast / Russia. It is still not clear what kind of source could generate such a signal.

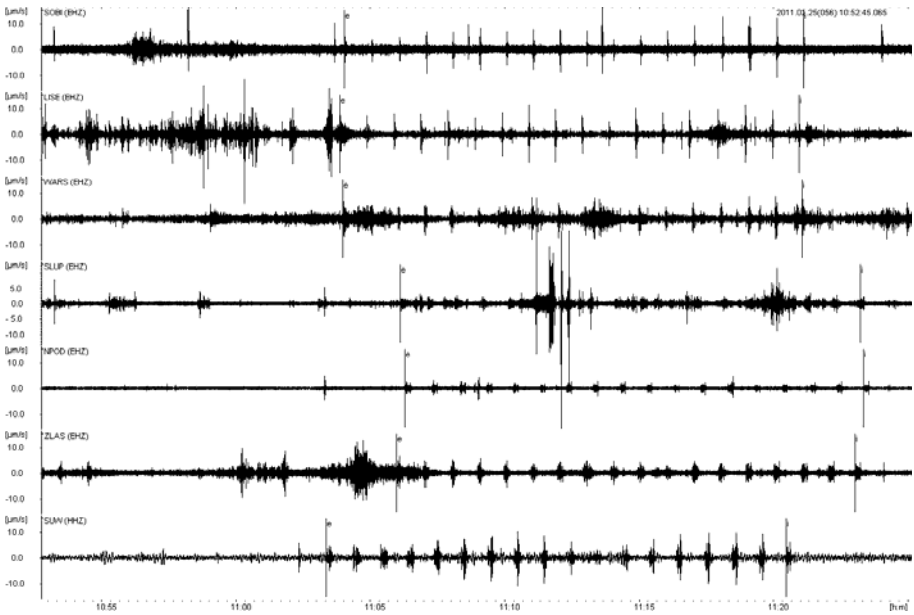


Fig. 10. Vertical components of stations: SOBI, LISE, WARS, SLUP, NPOD, ZLAS, SUW, which recorded the pulses in regular intervals in different regions of Poland. Delays between stations indicate much slower propagation than for seismic waves. First pulse on each seismogram is marked with "e" and the last one with "i".

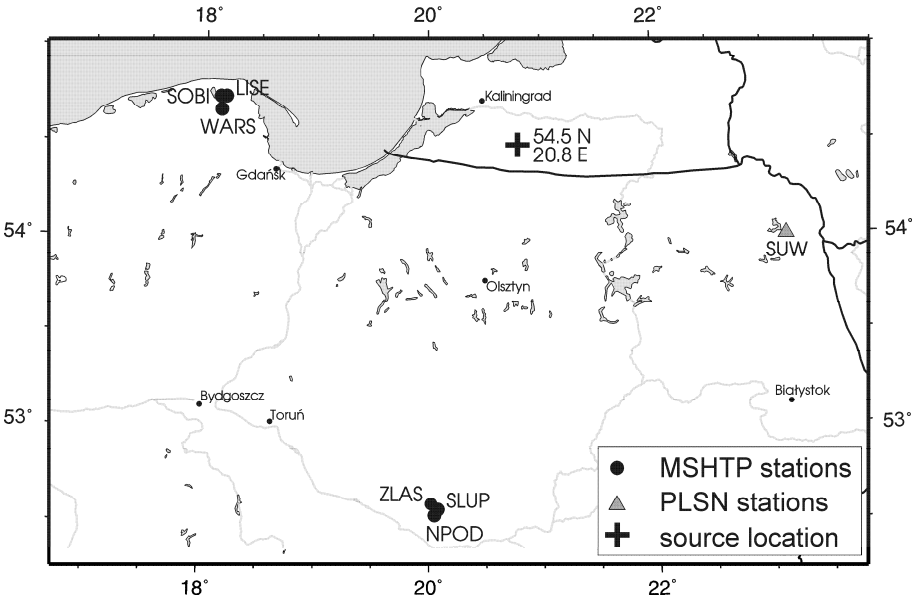


Fig. 11. Distribution of stations presented in Fig. 10 (circles) and source location obtained jointly with velocity inversion (cross).

6. CONCLUSIONS

"Monitoring of Seismic Hazard of Territory of Poland" was the first project of this scale in Poland that focused on local natural seismicity. The existing seismological network (9 stations) is meant for global and regional scale monitoring measurements. It is not capable to detect and locate events smaller than magnitude 3. A new seismic network composed of 24 mobile seismic stations significantly improved this sensitivity. Together with new seismic stations a new acquisition system was set up and new tools for maintaining a network and for data processing were developed, which makes a room for a further network growth.

Monitored regions were selected on the basis of analysis of historical seismicity and were scattered over the area of Poland. A five-year project confirmed seismic activity of the Carpathians (Stage 1, Region A, and Stage 2, Region I, Figs. 2 and 3), mainly in the Podhale region, where over 100 events have been recorded, and near Krynica / Beskid Sądecki / Carpathians (three events). At the beginning of the project both places have been already known for historical earthquakes and recent seismic activity. Second region with a surprisingly large earthquake is near Jarocin / Wielkopolska / western Poland (Stage 2, Region E, Fig. 3). This region was not considered as a potential source of such earthquake and was not covered by monitoring before the earthquake occurrence.

In other mountain regions, Sudetes (Stage 1, Region A, Fig. 2) and Holy Cross Mountains (Stage 2, Region H, Fig. 3), there has not been recorded any natural seismic event, which suggests that return period of detectable events is too long comparing to the monitoring period. This argument is valid for other regions as well, but bad recording conditions additionally reduce the detection possibility. The best recording conditions, in terms of noise level, are in mountains in the south, where sediments are very thin and it is often possible to put a seismometer directly on a rock. Much greater noise, by about two orders of the PSD magnitude, is in regions of sedimentary background (Stage 2, Regions C-G, Fig. 3).

The number of events in the Podhale region made it possible to conduct a preliminary study on ground motion prediction equation for this region, which is a key for future seismic hazard assessments. The obtained equation fits the data well but there is a lack of strong motion records, which makes the solution less adequate for higher magnitudes.

The project confirmed that some regions of Poland are seismically active, which makes them interesting objects of future studies. It was also proved that applied mobile monitoring network is a reliable and adequate tool for measurements of this type.

Acknowledgments. The project "Monitoring of Seismic Hazard of Territory of Poland" (MSHTP) was carried out by the Institute of Geophysics, Polish Academy of Sciences, in the period 2007-2012 on the order of the Ministry of the Environment and financed from funds of the National Fund for Environmental Protection and Water Management (contract No. 445/2007/Wn-07/F6-bp-tx/D).

Appendix 1

List of stations in the project

Station name	Stage	Site name	Recording period	Lat. [°]N	Long. [°]E	Elev. [m]	Region
AWRA	1	Ciche	17 Jun 2008 – 10 Aug 2008	49.40	19.87	670	B
BLAT	1	Błatnia	6 May 2009 – 2 Oct 2009	49.75	18.94	887	B
BOBR	1	Bóbrka	9 Sep 2008 – 19 May 2010	49.62	21.71	334	B
CICH	1+2	Ciche	11 Aug 2008 – 31 Dec 2012	49.39	19.87	670	B, I
CISO	2	Cisów	13 Jul 2010 – 31 Dec 2012	50.77	20.90	331	H
DEBO	1	Dębowiec	23 Sep 2008 – 1 Dec 2008	50.59	17.11	278	A
DOMA	2	Domatowo	28 Jul 2010 – 31 Dec 2012	54.71	18.21	103	D
DUKL	1	Dukla	19 Jun 2008 – 6 Aug 2008	49.52	21.68	502	B
DZIA	2	Dzianisz	27 Oct 2010 – 31 Dec 2012	49.32	19.84	921	I
FLOR	1	Florynka	18 Jun 2008 – 4 Sep 2008	49.55	20.98	403	B
GLAD	1	Gładyszów	26 Oct 2009 – 9 Jan 2010	49.53	21.29	521	B
GLIN	1	Glinka	17 Sep 2009 – 14 Apr 2010	49.46	19.18	673	B
HOLU	1	Hołuczków	17 Jul 2008 – 8 Jun 2010	49.58	22.33	381	B

to be continued

continuation

Station name	Stage	Site name	Recording period	Lat. [°]N	Long. [°]E	Elev. [m]	Region
HORD	2	Hordzieszka	7 Jul 2010 – 21 Oct 2010	51.74	22.18	175	G
JAR1	2	Mniszew	8 Jan 2010 – 31 Dec 2012	52.04	17.69	81	E
JAR2	2	Stęgosz	8 Jan 2010 – 31 Dec 2012	52.04	17.52	97	E
JAR3	2	Ludwinów	8 Jan 2010 – 31 Dec 2012	52.05	17.62	82	E
JAR4	2	Pogorzelica	9 Jan 2010 – 31 Dec 2012	52.14	17.59	71	E
JAR6	2	Lgów	9 Jan 2010 – 31 Dec 2012	52.10	17.53	72	E
JAWO	1	Jaworki – Biała Woda	18 Jun 2008 – 20 Jul 2008	49.41	20.57	616	B
JELE	2	Jeleniów	27 Oct 2010 – 31 Dec 2012	50.83	21.13	347	H
KLET	1	Kletno	2 Dec 2008 – 15 Dec 2008	50.25	16.84	878	A
KONO	2	Konorzatka	7 Jul 2010 – 31 Dec 2012	51.71	22.24	158	G
KRAS	2	Kraśnik Koszaliński	23 Aug 2010 – 31 Dec 2012	54.13	15.99	30	C
LABO	1	Łabowa	10 Sep 2008 – 19 May 2010	49.52	20.83	509	B
LE01	2	Karlikowo Leborskie	12 Jul 2011 – 4 Oct 2011	54.64	17.78	73	D
LE02	2	Rekowo Leborskie	13 Jul 2011 – 4 Oct 2012	54.63	17.79	66	D
LE03	2	Obliwice	14 Jul 2011 – 7 Oct 2013	54.62	17.74	56	D
LE04	2	Lebien	14 Jul 2011 – 4 Oct 2014	54.64	17.72	50	D
LE05	2	Basewice	20 Jul 2011 – 25 Sep 2015	54.67	17.79	69	D
LE06	2	Krepa Kaszubska	21 Jul 2011 – 7 Oct 2016	54.63	17.66	73	D

to be continued

continuation

Station name	Stage	Site name	Recording period	Lat. [°]N	Long. [°]E	Elev. [m]	Region
LE07	2	Niebedzino	21 Jul 2011 – 7 Oct 2017	54.57	17.62	55	D
LE08	2	Lubowidz	21 Jul 2011 – 7 Oct 2017	54.54	17.81	29	D
LE09	2	Leczyn Dolny	22 Jul 2011 – 6 Oct 2019	54.64	17.95	125	D
LISE	2	Lisewo	28 Jul 2010 – 31 Dec 2012	54.75	18.18	25	D
LULE	2	Lulewice	14 Jul 2010 – 31 Dec 2012	54.05	15.95	24	C
LUTW	1	Lutowiska	18 Jul 2008 – 31 May 2010	49.24	22.69	650	B
MAKO	2	Lechów	13 Jul 2010 – 31 Dec 2012	50.81	21.00	334	H
MERZ	2	Mierzynek	13 Jul 2010 – 31 Dec 2012	54.10	15.90	32	C
MNIS	1	Mniszków	9 Aug 2008 – 31 May 2010	50.86	15.94	626	A
NOSO	2	Nosówko	13 Jul 2010 – 23 Aug 2010	54.10	16.00	29	C
NPOD	2	Nowy Podleck	23 Aug 2010 – 6 Jan 2012	52.55	20.05	127	F
OCIE	2	Ociesęki	13 Jul 2010 – 31 Dec 2012	50.73	20.97	312	H
OSIE	1	Osiek Jasielski	19 Jun 2008 – 3 Sep 2008	49.64	21.49	226	B
OSIN	1	Osina Wielka	8 Aug 2008 – 4 Sep 2008	50.58	17.07	229	B
PAGO	1	Pagorzyna	9 Sep 2008 – 27 May 2010	49.69	21.33	360	B
PLUC	2	Płucki	25 Oct 2010 – 31 Dec 2012	50.80	21.07	315	H
POTK	2	Potok	26 Oct 2010 – 31 Dec 2012	49.36	20.16	886	I
PUS2	2	Pustkowo	23 Aug 2010 – 31 Dec 2012	54.08	16.04	20	C

to be continued

continuation

Station name	Stage	Site name	Recording period	Lat. [°]N	Long. [°]E	Elev. [m]	Region
PUST	2	Pustkowo	27 Jul 2010 – 23 Aug 2010	54.06	16.03	19	C
RACL	1	Racławice	7 Aug 2008 – 30 Aug 2008	49.75	21.18	342	B
ROPK	1	Ropki	9 Aug 2008 – 1 Jun 2010	49.46	21.13	563	B
RYC2	1	Rychwałdek	5 May 2009 – 17 May 2010	49.68	19.28	464	B
RYCH	1	Rychwałdek	1 Nov 2008 – 5 May 2009	49.68	19.28	493	B
RZEP	1	Rzepedź	19 Jul 2008 – 31 May 2010	49.39	22.10	458	B
SCWK	1	Szczawnik	9 Aug 2008 – 29 Oct 2008	49.38	20.87	523	B
SEDK	2	Sędek	13 Jul 2010 – 31 Dec 2012	50.77	21.01	396	H
SIER	1+2	Sierockie	11 Aug 2008 – 31 Dec 2012	49.36	19.96	1007	B, I
SKAW	1	Skawica	20 Jul 2008 – 12 May 2010	49.65	19.66	686	B
SLUP	2	Słupca	1 Jul 2010 – 6 Jan 2012	52.58	20.08	139	F
SOBI	2	Sobieńczyce	28 Jul 2010 – 31 Dec 2012	54.75	18.13	102	D
SOCH	2	Sochocino-Praga	1 Jul 2010 – 6 Jan 2012	52.57	19.98	129	F
SPAL	1	Spalona	17 Dec 2008 – 1 Jun 2010	50.28	16.54	760	A
SRGO	1	Srebrna Góra	8 Aug 2008 – 1 Jun 2010	50.58	16.66	483	A
STBY	1+2	Stare Bystre	20 Jul 2008 – 31 Dec 2012	49.43	19.94	682	B, I
STRO	1	Stronie	13 Sep 2008 – 1 Jun 2010	49.62	20.53	578	B
SWIA	1	Świątkowa Wielka	8 Aug 2008 – 4 Sep 2008	49.54	21.42	444	B

to be continued

continuation

Station name	Stage	Site name	Recording period	Lat. [°]N	Long. [°]E	Elev. [m]	Region
SWKR	2	Święty Krzyż	25 Oct 2010 – 31 Dec 2012	50.86	21.05	556	H
SZCZ	2	Szczałb	7 Jul 2010 – 21 Oct 2010	51.77	22.24	159	G
SZKL	1	Szklary	7 Aug 2008 – 31 May 2010	49.47	21.81	559	A
WARS	2	Warszkowo	28 Jul 2010 – 31 Dec 2012	54.68	18.14	43	D
WILC	1	Wilczyce	11 Aug 2008 – 31 May 2010	49.67	20.18	615	B
WRZO	1	Wrzosówka	7 Aug 2008 – 1 Jun 2010	50.38	16.90	722	A
ZAGR	1	Zagórze	21 Jul 2008 – 16 Jun 2010	49.83	19.55	337	B
ZAKE	2	Zakępie	7 Jul 2010 – 21 Oct 2010	51.73	22.30	158	G
ZAWO	1	Zawoja	17 Jun 2008 – 20 Jul 2008	49.64	19.53	613	B
ZDYN	1	Zdynia	30 Sep 2009 – 26 Oct 2009	49.49	21.27	521	B
ZDZW	2	Zdziar Wielki	2 Jul 2010 – 23 Aug 2010	52.63	20.05	134	F
ZEGI	1	Żegiestów	29 Oct 2008 – 31 May 2010	49.37	20.79	466	B
ZLAS	2	Zdziar Las	23 Aug 2010 – 6 Jan 2012	52.61	20.02	146	F
ZOCH	2	Żochocino	1 Jul 2010 – 23 Aug 2010	52.63	20.12	127	F

Appendix 2

List of events localized in Podhale/Carpathians during MSHTP project

No.	Date	Time	Latitude [°N]	Longitude [°E]	Depth [km]	M_L
1	5 Jul 2008	20:11:10.7	49.320	19.742	5	1.6
2	6 Jul 2008	11:58:11.5	49.354	19.780	7	1.5
3	5 Aug 2008	12:28:54.5	49.405	20.142	4	2.5
4	21 Sep 2008	16:05:23.8	49.356	20.013	3	0.4
5	10 Nov 2008	07:55:45.1	49.385	20.084	3	1.1
6	16 Dec 2008	23:31:54.9	49.373	20.012	4	1.6
7	12 Feb 2009	15:20:25.5	49.321	19.916	5	0.9
8	10 Apr 2009	17:45:39.1	49.379	19.917	3	2.0
9	10 Apr 2009	19:00:51.6	49.391	19.904	4	1.3
10	12 Apr 2009	00:27:31.2	49.391	19.910	6	0.5
11	30 Apr 2009	05:44:31.2	49.388	19.916	4	0.5
12	29 May 2009	14:57:24.4	49.317	19.690	3	2.1
13	29 May 2009	20:33:28.3	49.344	19.878	1	1.0
14	13 Aug 2009	19:47:23.9	49.384	19.834	3	1.7
15	29 Aug 2009	23:20:19.9	49.384	19.846	3	0.9
16	3 Sep 2009	20:40:52.7	49.392	19.835	3	0.9
17	7 Sep 2009	16:52:39.7	49.401	20.015	3	1.6
18	1 Oct 2009	07:06:41.7	49.393	19.903	3	0.8
19	4 Oct 2009	00:32:59.4	49.393	19.904	4	0.8
20	4 Oct 2009	23:50:21.9	49.308	19.975	3	0.8
21	16 Oct 2009	04:22:33.4	49.368	19.884	3	0.9
22	28 Oct 2009	22:00:28.2	49.331	19.924	8	0.6
23	12 Nov 2009	14:38:18.3	49.320	19.976	3	0.9
24	15 Dec 2009	11:46:51.3	49.332	19.916	7	1.9
25	19 Mar 2010	07:18:54.9	49.380	19.904	4	0.3
26	31 Mar 2010	04:09:01.9	49.384	19.907	4	1.1
27	3 Apr 2010	20:28:21.1	49.312	19.970	5	1.0
28	4 Apr 2010	07:18:33.6	49.406	19.851	4	0.7
29	7 May 2010	20:50:00.4	49.367	19.885	4	1.8
30	13 May 2010	19:02:02.7	49.396	19.921	5	1.3
31	11 Aug 2010	05:05:49.6	49.392	19.910	4	0.7

to be continued

continuation

No.	Date	Time	Latitude [°N]	Longitude [°E]	Depth [km]	M_L
32	1 Oct 2010	06:31:53.3	49.389	19.912	3	1.8
33	2 Oct 2010	23:17:54.9	49.391	19.910	4	0.5
34	7 Nov 2010	09:52:42.9	49.387	19.923	4	0.5
35	8 Nov 2010	23:16:41.2	49.376	19.884	5	0.6
36	10 Nov 2010	00:32:52.9	49.392	19.963	4	0.2
37	19 Nov 2010	02:56:19.3	49.398	19.924	5	0.2
38	25 Jan 2011	02:26:28.5	49.381	20.000	3	0.4
39	8 Mar 2011	01:50:32.1	49.331	19.813	4	0.6
40	28 Mar 2011	01:53:40.3	49.332	19.828	5	0.9
41	12 May 2011	01:38:58.2	49.291	20.119	3	0.7
42	17 May 2011	20:21:50.8	49.306	19.993	3	0.7
43	21 May 2011	01:33:01.2	49.397	19.956	3	0.2
44	21 Jun 2011	14:54:25.9	49.418	19.829	4	1.4
45	11 Jul 2011	21:14:07.8	49.398	19.943	4	0.6
46	11 Jul 2011	22:06:24.2	49.280	19.774	5	0.8
47	12 Jul 2011	02:41:53.8	49.272	19.735	5	0.9
48	12 Jul 2011	02:54:13.3	49.267	19.762	5	0.6
49	24 Jul 2011	03:28:57.3	49.270	19.869	5	1.0
50	26 Jul 2011	00:53:33.8	49.296	19.746	5	0.5
51	29 Jul 2011	20:28:58.8	49.391	19.907	5	0.3
52	22 Aug 2011	22:20:52.9	49.374	20.072	2	0.4
53	3 Sep 2011	19:20:09.7	49.370	19.889	2	0.2
54	26 Sep 2011	03:48:16.9	49.374	20.018	4	0.3
55	26 Sep 2011	16:30:54.8	49.401	19.925	5	0.8
56	26 Sep 2011	22:23:07.9	49.392	19.934	4	0.4
57	6 Oct 2011	23:05:21.8	49.385	19.831	6	1.1
58	6 Nov 2011	11:35:07.1	49.366	19.834	5	0.5
59	6 Nov 2011	13:19:37.6	49.368	19.833	5	1.5
60	26 Nov 2011	16:31:42.6	49.386	19.910	3	2.1
61	26 Nov 2011	18:13:37.1	49.385	19.907	3	0.9
62	15 Dec 2011	03:05:42.9	49.320	19.968	2	2.1
63	15 Dec 2011	04:56:33.7	49.318	19.968	2	1.3
64	23 Dec 2011	17:07:03.8	49.339	19.750	3	0.8
65	8 Jan 2012	21:45:55.5	49.283	19.946	3	0.6
66	1 Mar 2012	19:34:38.3	49.377	19.882	3	0.9

to be continued

continuation

No.	Date	Time	Latitude [°N]	Longitude [°E]	Depth [km]	M_L
67	5 Mar 2012	14:24:45.6	49.372	20.011	3	0.6
68	8 Mar 2012	21:33:30.3	49.379	20.019	3	0.5
69	9 Jun 2012	03:11:50.2	49.366	20.018	4	0.6
70	6 Jul 2012	06:18:25.3	49.397	19.950	4	1.2
71	16 Jul 2012	09:15:28.6	49.368	20.018	3	1.4
72	17 Jul 2012	00:47:16.8	49.369	20.015	4	0.7
73	17 Jul 2012	05:27:13.0	49.365	20.016	3	0.6
74	17 Jul 2012	23:15:12.7	49.369	20.014	4	0.4
75	21 Jul 2012	22:36:52.9	49.370	20.015	4	0.4
76	25 Aug 2012	23:27:40.4	49.407	19.948	4	0.6
77	28 Sep 2012	17:09:34.2	49.351	19.845	3	1.9
78	16 Oct 2012	14:35:31.5	49.365	20.015	3	0.8
79	16 Oct 2012	14:48:30.0	49.374	20.020	4	0.4
80	4 Dec 2012	03:57:55.2	49.392	19.924	4	0.6
81	4 Dec 2012	13:09:30.9	49.351	19.855	3	1.9

References

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.

Aleksandrowicz, D. (1982), Automatic seismic station ASS PCM 6/10, *Acta Geophys. Pol.* **30**, 4, 381-392.

Andrews, D.J. (1986), Objective determination of source parameters and similarity of earthquakes of different size. **In:** S. Das, J. Boatwright, and C.H. Sholtz (eds.), *Earthquake Source Mechanics*, Geophysical Monograph, Vol. 37, American Geophysical Union, Washington, D.C., 259-267, DOI: 10.1029/ GM037p0259.

Brune, J.N. (1970), Tectonic stress and the spectra of seismic shear waves from earthquakes, *J. Geophys. Res.* **75**, 26, 4997-5009, DOI: 10.1029/JB075i026 p04997.

Campbell, K.W. (2003), Prediction of strong ground motion using the hybrid empirical metod and its use in the development of ground-motion (attenuation) relations in eastern North America, *Bull. Seismol. Soc. Am.* **93**, 3, 1012-1033, DOI: 10.1785/0120020002.

Cauzzi, C., and E. Faccioli (2008), Broadband (0.05 to 20 s) prediction of displacement response spectra based on worldwide digital records, *J. Seismol*. **12**, 4, 453-475, DOI: 10.1007/s10950-008-9098-y.

Chiou, B.S.-J., and R.R. Youngs (2008), An NGA model for the average horizontal component of peak ground motion and response spectra, *Earthq. Spectra* **24**, 1, 173-215, DOI: 10.1193/1.2894832.

Delavaud, E., F. Cotton, S. Akkar, F. Scherbaum, L. Danciu, C. Beauval, S. Drouet, J. Douglas, R. Basili, M.A. Sandikkaya, M. Segou, E. Faccioli, and N. Theodoulidis (2012), Toward a ground-motion logic tree for probabilistic seismic hazard assessment in Europe, *J. Seismol*. **16**, 3, 451-473, DOI: 10.1007/s10950-012-9281-z.

Domański, B.M. (2007), Source parameters of the 2004 Kaliningrad earthquakes, *Acta Geophys.* **55**, 3, 267-287, DOI: 10.2478/s11600-007-0021-7.

Grad, M., A. Špičák, G.R. Keller, A. Guterch, M. Brož, E. Hegedüs, and Working Group (2003), SUDETES 2003 seismic experiment, *Stud. Geophys. Geod.* **47**, 3, 681-689, DOI: 10.1023/A:1024732206210.

Guterch, A., M. Grad, H. Thybo, and G.R. Keller (1999), POLONAISE'97 – an international seismic experiment between Precambrian and Variscan Europe in Poland, *Tectonophysics* **314**, 1-3, 101-121, DOI: 10.1016/S0040-1951 (99)00239-5.

Guterch, A., M. Grad, G.R. Keller, K. Posgay, J. Vozár, A. Špičák, E. Brückl, Z. Hajnal, H. Thybo, O. Selvi, and CELEBRATION 2000 Experiment Team (2003), CELEBRATION 2000 seismic experiment, *Stud. Geophys. Geod.* **47**, 3, 659-669, DOI: 10.1023/A:1024728005301.

Guterch, B. (2006), Seismic events in the Orawa – Nowy Targ Basin, Western Carpathians, November 30, 2004 – December 2005, *Acta Geodyn. Geomater*. **3**, 3, 85-95.

Guterch, B. (2007), Seismological Bulletin 2004, Local Earthquakes Recorded by Polish Seismological Stations, *Publs. Inst. Geophys. Pol. Acad. Sci*. **B-40**, 397.

Guterch, B. (2009), Seismicity in Poland in the light of historical records, *Prz. Geol.* **57**, 6, 513-520 (in Polish).

Guterch, B., H. Lewandowska-Marciniak, and J. Niewiadomski (2005), Earthquakes recorded in Poland along the Pieniny Klippen Belt, Western Carpathians, *Acta Geophys. Pol.* **53**, 1, 27-45.

Hościłowicz, M., J. Olszewski, and J. Wiszniowski (1990), Microcomputer seismic station MK-1, *Acta Geophys. Pol.* **38**, 2, 141-147.

Idziak, A.F., and R. Dubiel (eds.) (2011), *Geophysics in Mining and Enviromental Protection*, Geoplanet: Earth and Planetary Sciences, Springer, Berlin Heidelberg, DOI: 10.1007/978-3-642-19097-1.

Joyner, W.B., and D.M. Boore (1993), Methods for regression analysis of strong-motion data, *Bull. Seismol. Soc. Am.* **83**, 2, 469-487.

Lizurek, G., and P. Wiejacz (2011), Moment tensor solution and physical parameters of selected recent seismic events at Rudna Copper Mine. **In:** A.F. Idziak and R. Dubiel (eds.), *Geophysics in Mining and Environmental Protection*, Geoplanet: Earth and Planetary Sciences, Springer, Berlin Heidelberg, 11-19, DOI: 10.1007/978-3-642-19097-1_2.

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.

Olszewski, J., and J. Wiszniowski (1993), A microcomputer-based seismic station, *Pol. Tech. Rev.* **1**, 18-21.

Orlecka-Sikora, B., S. Lasocki, G. Lizurek, and Ł. Rudziński (2012), Response of seismic activity in mines to the stress changes due to mining induced strong seismic events, *Int. J. Rock Mech. Min. Sci.* **53**, 151-158, DOI: 10.1016/j.ijrmms.2012.05.010.

Park, J., C.R. Lindberg, and F.L. Vernon III (1987), Multitaper spectral analysis of high-frequency seismograms, *J. Geophys. Res.* **92**, B12, 12675-12684, DOI: 10.1029/JB092iB12p12675.

Schenk, V., Z. Schenková, P. Kottnauer, B. Guterch, and P. Labák (2000), Earthquake hazard for the Czech Republic, Poland and Slovakia – Contribution to the ILC/IASPEI Global Seismic Hazard Assessment Program. **In:** G.A. Papadopoulos, T, Murty, S. Venkatesh, and R. Blong (eds.), *Natural Hazards. State-of-the-Art at the End of the Second Millennium*, Kluwer Academic Publ., Dordrecht, 331-345, DOI: 10.1007/978-94-017-2386-2_14.

Toro, G.R. (2002), Modification of the Toro et al. (1997) attenuation equations for large magnitudes and short distances, Tech. Rep., Risk Engineering, Boulder, Colorado, USA.

Wiejacz, P. (2006), The Kaliningrad earthquakes of September 21, 2004, *Acta Geodyn. Geomater*. **3**, 2, 7-16.

Wiejacz, P., and W. Dębski (2009), Podhale, Poland, earthquake of November 30, 2004, *Acta Geophys*. **57**, 2, 346-366, DOI: 10.2478/s11600-009-0007-8.

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.

Wiejacz, P., and J. Wiszniowski (2006), Moment magnitude determination of local seismic events recorded at selected Polish seismic stations, *Acta Geophys.* **54**, 1, 15-32, DOI: 10.2478/s11600-006-0003-1.

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.R. 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, T. Plenefisch, B. Růžek, S. Sliaupa, P. Środa, M. Świeczak, T. Tiira, P. Voss, and P. Wiejacz (2008), PASSEQ 2006-

2008: Passive seismic experiment in Trans-European Suture Zone, *Stud. Geophys. Geod.* **52**, 3, 439-448, DOI: 10.1007/s11200-008-0030-2.

Wiszniowski, J. (2000), Application of real time recurrent neural network for seismic event detection, *Acta Geophys. Pol.* **48**, 1, 1-26.

Wiszniowski, J., B.M. 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.

Zuberek, W.M., and K. Jochymczyk (eds.) (2010), *Origin and Seismic Hazard Assessment in the Upper Silesian Coal Basin*, Wyd. Uniwersytetu Śląskiego, Katowice, 95 pp. (in Polish).

Received 16 September 2013
Received in revised form 16 June 2014
Accepted 18 June 2014

DE GRUYTER OPEN

Acta Geophysica
vol. 63, no. 1, Feb. 2015, pp. 45-61
DOI: 10.2478/s11600-013-0197-y

On Choosing Effective Elasticity Tensors Using a Monte-Carlo Method

Tomasz DANEK[1,2] and Michael A. SLAWINSKI[1]

[1]Department of Earth Sciences, Memorial University of Newfoundland, St. John's, Canada; e-mail: mslawins@mun.ca

[2]Department of Geoinformatics and Applied Computer Science, AGH – University of Science and Technology, Kraków, Poland
e-mail: tdanek@agh.edu.pl (corresponding author)

A b s t r a c t

A generally anisotropic elasticity tensor can be related to its closest counterparts in various symmetry classes. We refer to these counterparts as effective tensors in these classes. In finding effective tensors, we do not assume *a priori* orientations of their symmetry planes and axes. Knowledge of orientations of Hookean solids allows us to infer properties of materials represented by these solids. Obtaining orientations and parameter values of effective tensors is a highly nonlinear process involving finding absolute minima for orthogonal projections under all three-dimensional rotations. Given the standard deviations of the components of a generally anisotropic tensor, we examine the influence of measurement errors on the properties of effective tensors. We use a global optimization method to generate thousands of realizations of a generally anisotropic tensor, subject to errors. Using this optimization, we perform a Monte Carlo analysis of distances between that tensor and its counterparts in different symmetry classes, as well as of their orientations and elasticity parameters.

Key words: anisotropy, vertical seismic profile (VSP), inversion.

1. INTRODUCTION

The purpose of this paper is to examine a generally anisotropic elasticity tensor, expressed in terms of twenty-one elasticity parameters and obtained from vertical seismic profiling (VSP) measurements, to infer properties of materials represented by this tensor. Beginning with a generally anisotropic tensor, we are able to infer these properties by relating this tensor to its closest counterparts in the sense of Frobenius norm, as defined by Gazis *et al.* (1963), in all material symmetries of Hookean solids, as shown by Danek *et al*. (2013). Herein, we focus our attention on examining symmetries used in seismology: monoclinic, orthotropic, transversely isotropic, and isotropic tensors. Following the definition and nomenclature of Kochetov and Slawinski (2009a, b), we refer to these counterparts as effective tensors of these symmetry classes. Note that we use generally anisotropic and orthotropic, not triclinic, and orthorhombic crystals, respectively. The latter terms are associated with lattice symmetries of crystals, while in seismology we deal with continua and their symmetries, which are symmetries of the elasticity tensor. In such a case, orthotropic refers to three mutually orthogonal symmetry planes.

Consideration of a generally anisotropic tensor allows us to examine effective tensors belonging to distinct symmetry classes without a bias of prior assumptions. In other words, the choice of the symmetry class model is guided by the data from which the generally anisotropic tensor is derived.

Obtaining these orientations and parameter values is a mathematically involved process. Explicit underpinnings of the methodology used in this paper are presented by Danek *et al*. (2013), and the reader is referred to that publication and references therein. Herein, we provide an overview to render the present paper self-contained.

To infer information about materials examined through VSP measurements, we consider relationships between the obtained tensor and its symmetric counterparts. Such a tensor was obtained by Dewangan and Grechka (2003) from multi-component and multi-azimuth walkaway VSP data, and such relationships are considered in terms of distance between tensors, as proposed by Gazis *et al*. (1963). The concept of such a distance is discussed by several researchers, including Norris (2006), Bóna (2009), and Kochetov and Slawinski (2009a, b). The present work, which is formulated in the context of a computationally efficient global optimization, allows us to obtain thousands of solutions within a few hours on a multi-core CPU computer. Hence, we can infer properties of materials, together with reliability of such inferences, by examining distributions, illustrated by histograms, of the elasticity parameters of effective tensors and distributions of orientations of these tensors. A computationally efficient scheme is crucial for generating such distributions.

As discussed by Kochetov and Slawinski (2009a), the distance between a generally anisotropic tensor and its counterpart belonging to a given symmetry class is obtained by finding the orientation that minimizes the distance. Performing a search under all orientations leads to a highly nonlinear optimization problem, which commonly exhibits many local minima. In the past, local-optimization methods have been used, which must be restricted to the vicinity of the global minimum to avoid convergence to a local one. A restriction based on visual examination of distance plots was proposed by Kochetov and Slawinski (2009a, b). Such an examination, however, is practically impossible if we wish to perturb the generally anisotropic tensor thousands of times to consider the effect of errors on the distribution of values that describe properties of effective tensors. Herein, we address this problem by applying a global optimization. Using this method, we can find effective tensors by a Monte Carlo (MC) method (see, *e.g.*, Tarantola 2005), and determine distributions of their orientations and of their elasticity parameters. These distributions arise from errors in which the original tensor is given. That is, our inversion consists of distributions of values that describe properties and orientations of effective tensors. These distributions, which are akin to error bars, allow us to gain an insight into the reliability of a given effective tensor in representing the generally anisotropic one.

This paper has a following layout. First, we review the concept of the effective elasticity tensor and describe the global optimization used in its search. Then, using this optimization, we examine the generally anisotropic tensor obtained by Dewangan and Grechka (2003) and discussed also in Chapter 9 of Tsvankin and Grechka (2011). Next we analyze the sensitivity of the solution to perturbation of elasticity tensor C using MC technique. This way we can evaluate the reliability of the solutions.

2. EFFECTIVE ELASTICITY TENSOR

For a fixed coordinate system, we can relate a general elasticity tensor, c, to its counterpart, c^{sym}, which belongs to a particular symmetry class. Tensor c^{sym} is the orthogonal projection of c, in the sense of the Frobenius inner product (which is the sum of products of the corresponding components, $a_{\mathrm{ijkl}}\, b_{\mathrm{ijkl}}$), onto the linear space containing all tensors of that symmetry class, as described by Gazis *et al*. (1963).

The distance-squared between c and c^{sym} is:

$$d^2_{\mathrm{sym}} = \left\| c - c^{\mathrm{sym}} \right\|^2 = \left\| c \right\|^2 - \left\| c^{\mathrm{sym}} \right\|^2 . \tag{1}$$

The second equality is a consequence of the orthogonality of c and c^{sym}. Components of tensor c^{sym}, and, hence, the value of distance obtained from

Eq. 1, depend on the orientation of the coordinate system. To determine the effective tensor without assuming an *a priori* orientation, we must minimize d_{sym} in Eq. 1 with respect to orientations; in other words, we have to perform the minimization of d_{sym} under all rotations. To find the solution of this high dimensional minimization problem, taking into account an existence of multiple local minima, we formulate a metaheuristic global approach. In this approach, no prior knowledge about a solution is required. The only requirement is that – given two arbitrary points within that space – a candidate for a solution can be chosen based on the difference of value of target function. The manner in which candidates are selected depends on the choice of algorithm. As a search strategy, we choose particle swarm optimization (PSO) because of its simplicity and speed of computation. This search strategy was formulated by Kennedy and Eberhart (1995) and used by Danek *et al*. (2013) to find the closest tensor of a given symmetry class. Furthermore – unlike other metaheuristics, say, genetic algorithms or simulated annealing – PSO does not require algorithm-parameter tuning (see Donelli *et al*. 2006). PSO is the stochastic technique that simulates social behavior of animals searching for food, as exemplified by a swarm of fish, insects, *etc*. In the present case, each particle represents a set of quaternion parameters in a four-dimensional solution space. We choose this representation because quaternions are particularly convenient for describing three-dimensional rotations (see, *e.g*., Stillwell 2008). In particular, they are computationally more convenient than the Euler angles. During the optimization process, each particle is "aware" of three positions: its current position, x_i, its best individual position, p_i, and the best position of the entire swarm, p_g. Best positions are points in the solution space for which a target function exhibits the lowest value obtained in all previous iterations.

The amplitude of a jump from the previous to the current position is defined by parameter v_i, called velocity; its value depends on the difference between the best position of an individual particle and the best position found so far by all particles. The canonical PSO formula is (Clerc and Kennedy 2002)

$$\begin{aligned} v_i &= \chi\left(v_i + U\left(0,\Phi_1\right)\otimes\left(p_i - x_i\right) + U\left(0,\Phi_2\right)\otimes\left(p_g - x_i\right)\right), \\ x_i &= x_i + v_i \ , \\ \chi &= \frac{2}{\Phi - 2 + \sqrt{\Phi^2 - 4\Phi}} \ , \\ \Phi &= \Phi_1 + \Phi_2 > 4 \ . \end{aligned} \tag{2}$$

where U represents uniform distribution and $\otimes$ is a component wise multiplication. Commonly, Φ, which is the sum of weights of a personal and

swarm information, Φ_1 and Φ_2, respectively, is usually set to 4.1, which means that a constant velocity multiplier, χ, is approximately 0.73 and U is a random number between 0 and approximately 1.5, if both weights are equal. This scheme guarantees convergences without particle velocity limitations. The flowchart of the algorithm is presented in Fig. 1. Since the elasticity tensor possesses index symmetries, $c_{ijkl} = c_{jikl} = c_{klij}$ (see, *e.g.*, Slawinski 2010), we can write its components as entries of a symmetric 6×6 matrix. Hence, Hooke's law,

$$\sigma_{ij} = \sum_{kl=1}^{3} c_{ijkl} \varepsilon_{kl} \ , \tag{3}$$

can be written in a manner that allows us for a convenient display of elasticity parameters, namely,

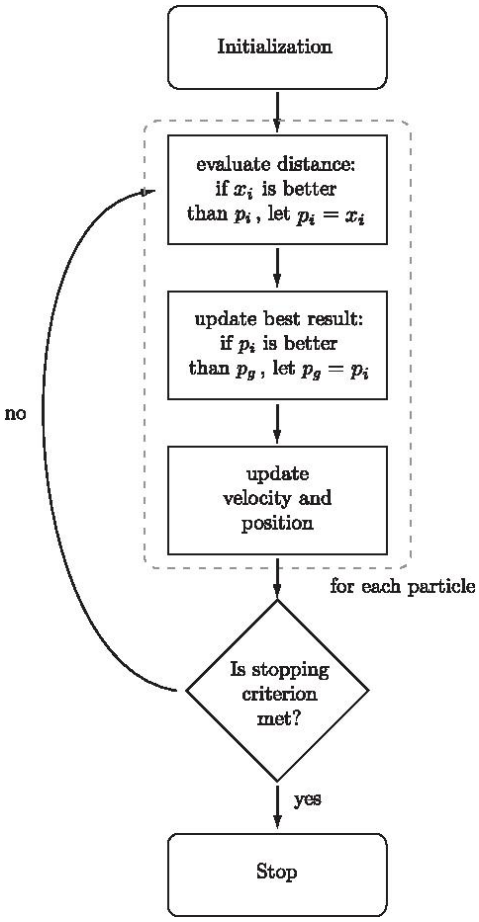


Fig. 1. Flowchart describing the particle-swarm-optimization algorithm: x_i is the current particle position, p_i is its best position, and p_g is the best position for the entire swarm; see text for details about the algorithm.

$$\begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sqrt{2}\sigma_{23} \\ \sqrt{2}\sigma_{13} \\ \sqrt{2}\sigma_{12} \end{bmatrix} = \begin{bmatrix} c_{1111} & c_{1122} & c_{1133} & \sqrt{2}c_{1123} & \sqrt{2}c_{1113} & \sqrt{2}c_{1112} \\ c_{1122} & c_{2222} & c_{2233} & \sqrt{2}c_{2223} & \sqrt{2}c_{2213} & \sqrt{2}c_{2212} \\ c_{1133} & c_{2233} & c_{3333} & \sqrt{2}c_{3323} & \sqrt{2}c_{3313} & \sqrt{2}c_{3312} \\ \sqrt{2}c_{1123} & \sqrt{2}c_{2223} & \sqrt{2}c_{3323} & 2c_{2323} & 2c_{2313} & 2c_{2312} \\ \sqrt{2}c_{1113} & \sqrt{2}c_{2213} & \sqrt{2}c_{3313} & 2c_{2313} & 2c_{1313} & 2c_{1311} \\ \sqrt{2}c_{1112} & \sqrt{2}c_{2212} & \sqrt{2}c_{3312} & 2c_{2312} & 2c_{1312} & 2c_{1212} \end{bmatrix} \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ \sqrt{2}\varepsilon_{23} \\ \sqrt{2}\varepsilon_{13} \\ \sqrt{2}\varepsilon_{12} \end{bmatrix}, \tag{4}$$

which we refer to as the Kelvin notation, we shall denote the elasticity tensor by C. Tensor C, which includes factors of $\sqrt{2}$ or 2 in its entries, allows us to keep the same norm for both the strain and stress tensors, and as a consequence allows us to conveniently examine rotations associated with symmetry classes (see, *e.g.*, Chapman 2004). Also, unlike the so-called Voigt notation, Eq. 4 is a vector equation. Using the Kelvin notation, we can write the squared distance between c and c^{sym} as

$$d^2_{\mathrm{sym}} = \left\|C\right\|^2 - \left\|C^{\mathrm{sym}}\right\|^2, \tag{5}$$

which is equivalent to Eq. 1.

3. NUMERICAL STUDY

The crux for obtaining effective tensors by realizations of a generally anisotropic tensor perturbed by errors relies on the aforementioned global optimization. We apply this method to the tensor obtained from VSP measurements by Dewangan and Grechka (2003):

$$C = \begin{bmatrix} 7.8195 & 3.4495 & 2.5667 & \sqrt{2}(0.1374) & \sqrt{2}(0.0558) & \sqrt{2}(0.1239) \\ 3.4495 & 8.1284 & 2.3589 & \sqrt{2}(0.0812) & \sqrt{2}(0.0735) & \sqrt{2}(0.1692) \\ 2.5667 & 2.3589 & 7.0908 & \sqrt{2}(-0.0092) & \sqrt{2}(-0.0286) & \sqrt{2}(0.1655) \\ \sqrt{2}(0.1374) & \sqrt{2}(0.0812) & \sqrt{2}(-0.0092) & 2(1.6636) & 2(-0.0787) & 2(0.1053) \\ \sqrt{2}(0.0558) & \sqrt{2}(0.0735) & \sqrt{2}(-0.0286) & 2(-0.0787) & 2(2.0660) & 2(-0.1517) \\ \sqrt{2}(0.1239) & \sqrt{2}(0.1692) & \sqrt{2}(0.1655) & 2(0.1053) & 2(-0.1517) & 2(2.4270) \end{bmatrix} \tag{6}$$

The components of this generally anisotropic tensor are the density-scaled elasticity parameters; their units are $\mathrm{km}^2/\mathrm{s}^2$. Entries of matrix 6 were obtained with the following standard deviations:

$$S = \pm \begin{bmatrix} 0.1656 & 0.1122 & 0.1216 & \sqrt{2}(0.1176) & \sqrt{2}(0.0774) & \sqrt{2}(0.0741) \\ 0.1122 & 0.1862 & 0.1551 & \sqrt{2}(0.0797) & \sqrt{2}(0.1137) & \sqrt{2}(0.0832) \\ 0.1216 & 0.1551 & 0.1439 & \sqrt{2}(0.0856) & \sqrt{2}(0.0662) & \sqrt{2}(0.1010) \\ \sqrt{2}(0.1176) & \sqrt{2}(0.0797) & \sqrt{2}(0.0856) & 2(0.0714) & 2(0.0496) & 2(0.0542) \\ \sqrt{2}(0.0774) & \sqrt{2}(0.1137) & \sqrt{2}(0.0662) & 2(0.0496) & 2(0.0626) & 2(0.0621) \\ \sqrt{2}(0.0741) & \sqrt{2}(0.0832) & \sqrt{2}(0.1010) & 2(0.0542) & 2(0.0621) & 2(0.0802) \end{bmatrix} \tag{7}$$

Unlike matrixes 6 and 7 does not consist of components of a tensor; it does not satisfy the conditions of tensorial transformations. Thus, S is fixed in the coordinate system in which components 6 are expressed; it can be used as a measure of errors in that system only.

In writing Eqs. 6 and 7, we do not imply that the number of decimal points corresponds to the number of significant digits. We use more decimals to examine numerical stability and to compare accurately our results with those of Kochetov and Slawinski (2009a, b).

To examine the influence of errors, we generate thousands of realizations of tensor 6 with random perturbations whose standard deviations are given in matrix 7. For each realization, using the PSO method, we obtain the effective orthotropic tensor, whose natural orientation is illustrated in Fig. 2, and

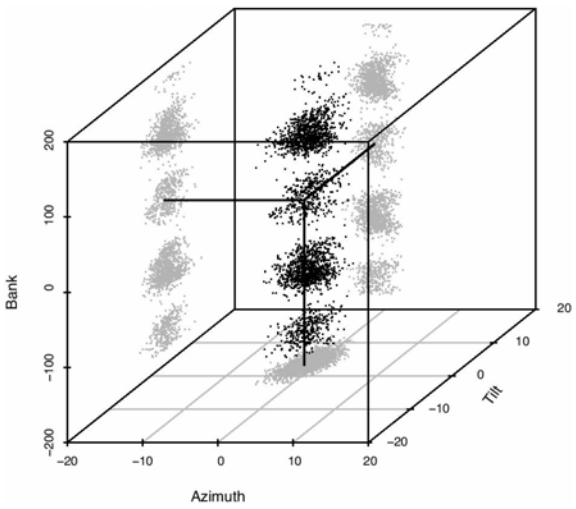


Fig. 2. Four clusters of the effective orthotropic tensors: each black dot is the orientation of the effective tensor corresponding to a realization of tensor 6 subject to errors 7; gray points are projections of black points. The vertex of black lines is the orientation of tensor 8, which results from tensor 6, without errors, and is the closest to the original coordinate system. The axes denote the three Euler angles. Each cluster corresponds to an equally valid natural coordinate system; their slightly different appearances are a result of random perturbation.

whose axes are the three Euler angles: the azimuth, defined by rotation about the x_3-axis, the tilt, by rotation about the x_1-axis, and the bank, by rotation about the new x_3-axis. Values of these angles are obtained from the quaternions used in the PSO. This figure contains four clusters, which correspond to orientations of natural coordinate systems of orthotropic tensors. Orientations of these systems differ in bank by $\pi/2$ or its multiple. None of these systems has a privileged status; our choice is a matter of convenience.

4. EFFECTIVE TENSOR

Choosing effective tensor

Since a symmetry class has more than one natural coordinate system – each associated with an aforementioned cluster – for our examinations, we choose systems that are closest to the one in which components (Eq. 6) are expressed. Following the global optimization, components of the effective orthotropic tensor – derived from tensor 6 without errors 7 – in the closest natural coordinate system are

$$C^{eff} = \begin{bmatrix} 7.7740 & 3.3634 & 2.4276 & 0 & 0 & 0 \\ 3.3634 & 8.3762 & 2.4879 & 0 & 0 & 0 \\ 2.4276 & 2.4879 & 7.0810 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2(1.6497) & 0 & 0 \\ 0 & 0 & 0 & 0 & 2(2.0784) & 0 \\ 0 & 0 & 0 & 0 & 0 & 2(2.3323) \end{bmatrix} \tag{8}$$

and the azimuth, tilt, and bank are −2.4°, 2.6°, and 19.3°, respectively. Note the similarity (expected) between tensors 8 and 6. Also to ensure consistency, note that expression 25 in Kochetov and Slawinski (2009a) and expression 25 in Danek *et al*. (2013) describe the same effective tensor but stated in a natural coordinate system that, relative to expression 8, is rotated by $\pi/2$ about the new x_3-axis.

According to the work of Dewangan and Grechka (2003) and Kochetov and Slawinski (2009a), tensor 6 can be represented by its counterpart exhibiting orthotropic symmetry. Let us use our method to examine whether or not a lesser or greater symmetry is a good representation of tensor 6 subject to errors 7.

Using the aforementioned global optimization, we obtain the distribution of shortest distances of tensor 6 to effective tensors of monoclinic, orthotropic, transversely isotropic, and isotropic symmetries. We compare these distributions to the distribution of the Frobenius norm of errors 7, which we obtain by generating random realizations of a zero tensor with these errors. This operation results in a distribution that is a square-root of the sum of thirty-six squares of independent random variables, M_{ij}, having a normal dis-

tribution with the zero mean and standard deviations given in matrix 7. Even though this distribution could be obtained analytically (see, *e.g.*, Mathai and Provost 1992), we use the same numerical method that we use to obtain the distributions of distances of tensor 6 to effective tensors.

To clarify that the expected value of the norm of

$$\left[M_{ij}\right] = \begin{bmatrix} N(0,S_{11}^2) & \cdots & N(0,S_{16}^2) \\ \vdots & \ddots & \vdots \\ N(0,S_{61}^2) & \cdots & N(0,S_{66}^2) \end{bmatrix}, \tag{9}$$

where $N(0,S_{ij}^2)$ are random variables and S_{ij} are entries of matrix 7, is not equal to zero, let us consider the variance

$$\mathrm{Var}(X) := E\left[X^2\right] - \left(E\left[X\right]\right)^2 , \tag{10}$$

where X is a random variable and E denotes the expected value. In our case, $E[X] = 0$, so the expected value of the square of random variable is equal to its variance. Hence, the expected value of the norm of matrix 9 is

$$E\left(|M|\right) = \sqrt{\sum_{i,j=1}^{6} S_{ij}^2} , \tag{11}$$

which is the norm of matrix 7, namely, 0.7844. This value is in agreement with the value obtained numerically: 0.7747, which corresponds to the location of the apex of the black line in Fig. 3.

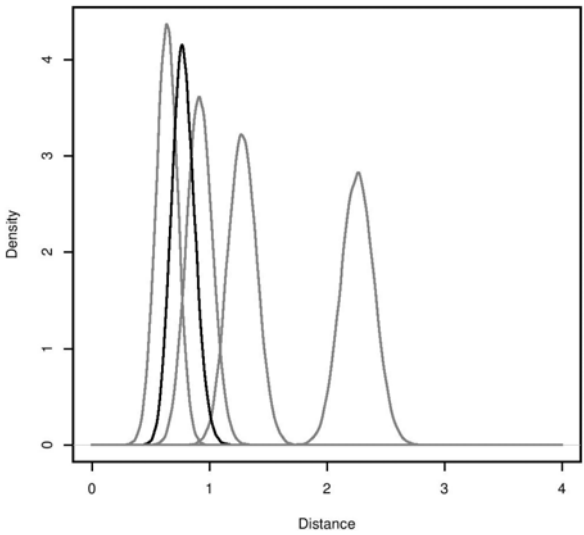


Fig. 3. Density distributions: the black line represents density of the norm of matrix 9, whose mean value – by expression 11 – is equal to the norm of matrix 7. Proceeding from left to right, gray lines represent densities of distance distributions for monoclinic, orthotropic, transversely isotropic, and isotropic symmetries, respectively.

In Figure 3, the distribution of errors 7 overlaps not only with the distribution of distances to the orthotropic symmetry but also to the monoclinic and transversely isotropic symmetries. As suggested by Kochetov and Slawinski (2009b), we view the overlap between the distribution of the distance from tensor 6 with its symmetric counterpart and with the distribution of the Frobenius norm of errors 7 as an indication that a symmetric tensor might represent tensor 6. Hence, the following question arises: could we choose a tensor of monoclinic or transversely isotropic symmetry to represent tensor 6?

To address this question, we note that the natural-coordinate expressions of the orthotropic, and higher, symmetries require $c_{1112} = c_{2212} = c_{3313} = 0$ (see, *e.g.*, Slawinski 2010). For examination of this issue, we generate thousands of realizations of tensor 6 subject to errors 7, and express them in the orientations of their closest monoclinic counterparts. Since $(0, 0, 0)$ is in the center of the obtained cluster shown in Fig. 4, we conclude that tensor 6 with errors 7 appears to be more symmetric than monoclinic.

Also, the natural-coordinate expressions of transverse isotropy require $c_{1111} = c_{2222}$, $c_{1133} = c_{2233}$, and $c_{2323} = c_{1313}$. To examine this issue, we generate thousands of realizations and express them in the orientations of their closest transversely isotropic counterparts. Examining the left panels in

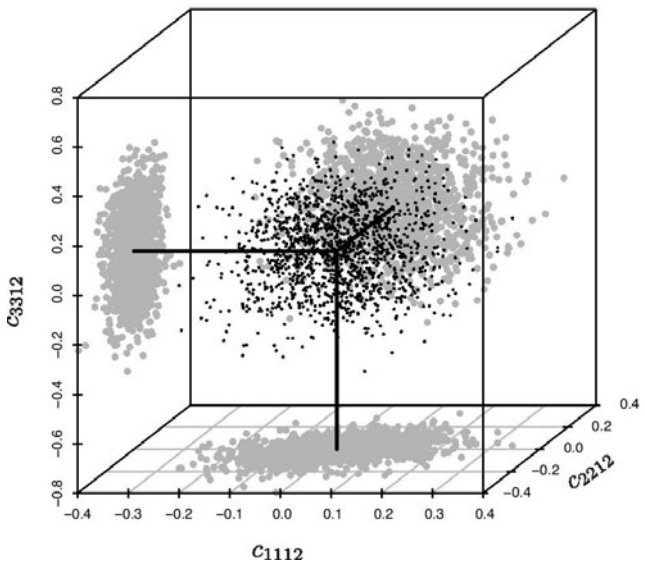


Fig. 4. Selected entries of realizations of tensor 6 subject to errors 7 in coordinate systems whose orientations correspond to the closest monoclinic tensors. Note that – with no constraints applied – the values of c_{1112}, c_{2212}, and c_{3313} are scattered around zeros, which are highlighted by solid lines. From this pattern we infer that tensor 6 can be represented by an effective tensor whose symmetry is higher than monoclinic.

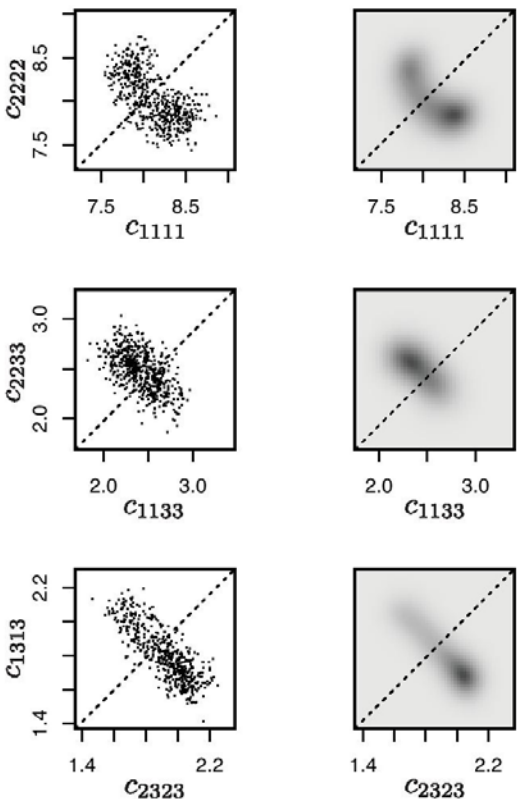


Fig. 5. Selected entries of realizations of tensor 6 subject to errors 7 expressed in coordinate systems whose orientations correspond to the closest transversely isotropic tensors, and their marginal distributions. For picture clarity, only 1000 points are shown on the left panel.

Fig. 5, we observe that obtained clusters are crossed by dashed lines showing these equalities; however, from examination of the right panel, we see that two-dimensional marginal distributions show that areas of the highest density are away from these equalities. We conclude that tensor 6 with errors 7 exhibits a lesser symmetry than transverse isotropy.

Thus, we choose the orthotropic symmetry to represent tensor 6.

Properties of chosen effective tensor

Having accepted that tensor 6 subject to errors 7 can be represented by an effective orthotropic tensor, let us examine its properties in the context of errors. These properties are the elasticity parameters, whose distributions are illustrated in Fig. 6, and orientation, whose distribution is illustrated in Fig. 7.

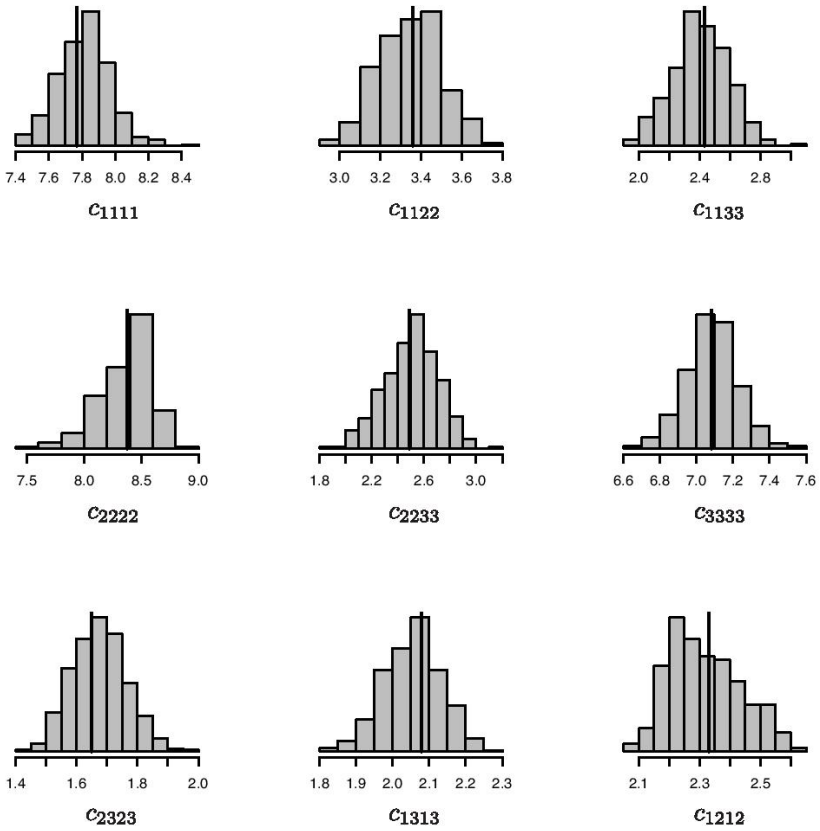


Fig. 6. Histograms of density-scaled elasticity parameters of the effective orthotropic tensors obtained from realizations of tensor 6, subject to errors 7, in natural coordinates whose orientations are illustrated in Fig. 7. Solid lines correspond to the values for the error-free case.

Furthermore, from the properties of this tensor we can infer properties of materials examined by VSP measurements from which tensor 6 is obtained.

The behavior of the histograms in Fig. 6, including their unimodality and relatively confined widths, suggests that the orthotropic symmetry contains much information about tensor 6 subject to errors 7, which is consistent with the symmetry choice discussed above. Examining the azimuth and tilt displayed in Fig. 7, we see that one of the symmetry planes of the effective orthotropic tensor is close to horizontal. The value of the bank indicates that the axes of the system in which tensor 6 is expressed are oblique to natural coordinates of the effective tensor. This information can be used to infer orientations of layers and fractures in a manner akin to those examined by Grechka and Kachanov (2006).

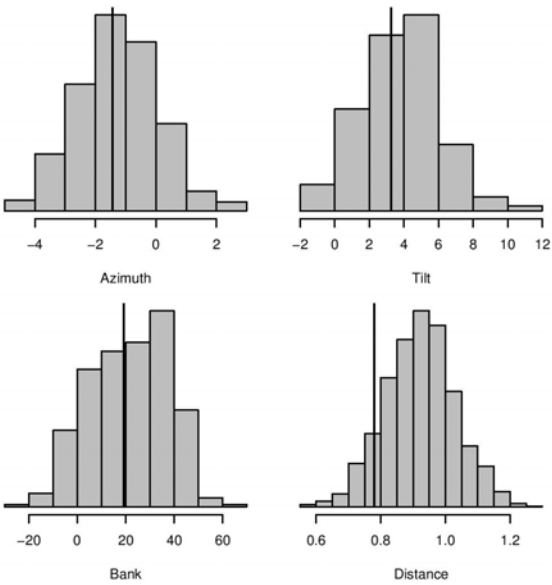


Fig. 7. Histograms of the azimuth, tilt, bank (in degrees), and distance (in km^2/s^2), from tensor 6 to effective orthotropic tensors obtained from realizations of tensor 6 subject to errors 7. The first three histograms are projections of the cluster in Fig. 2 – whose bank is close to zero – to the three axes therein. Solid lines correspond to the values for the error-free case.

To gain an insight into the strength of anisotropy under consideration, following the formulation proposed by Tsvankin (1997) let us express the pertinent components of tensor 8 in terms of the seven parameters that are zero for the case of isotropy. Our results shown in Fig. 8 are consistent with those presented by Dewangan and Grechka (2003) and further elaborated by Tsvankin and Grechka (2011), Section 7.13.

As shown in Fig. 8, distributions of several among the aforementioned parameters contain zero. Nevertheless, absolute average values of $\delta^{(1)}$ and $\gamma^{(2)}$ are close to 0.2, which suggests that anisotropy is not weak. Moreover, the shear-wave-splitting coefficient, $(c_{1313} - c_{2323})/(2c_{2323})$, which is important in fracture detection, is about 0.12; again, it is similar to 0.1 obtained by Tsvankin and Grechka (2011). This value is relatively large since, typically, the observed splitting coefficients are less than 0.05 (Tsvankin 2013, pers. comm.).

Note that values of i in $\delta^{(i)}$, $\varepsilon^{(i)}$, and $\gamma^{(i)}$, in this figure, are interchanged with respect to values in Kochetov and Slawinski (2009b) because the coordinate systems differ by $\pi/2$ about the new x_3-axis. Similarly, the interchange in Dewangan and Grechka (2003) and Tsvankin and Grechka (2011) is a consequence of coordinate systems belonging to different clusters shown in

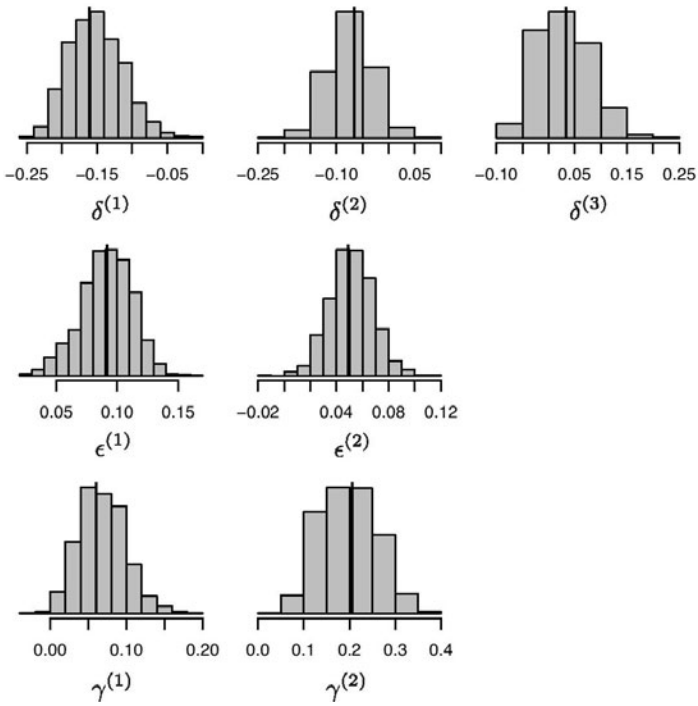


Fig. 8. Histograms of the seven elasticity parameters of orthotropic symmetry, whose values are zero in the case of isotropy. Solid lines correspond to the values for the error-free case.

Fig. 2. None of the clusters is privileged, as long as all results are expressed with respect to the same system.

5. CONCLUSIONS

The presented method allows us to infer from seismic measurements information about materials represented by a generally anisotropic tensor. This method extends the approach introduced by Kochetov and Slawinski (2009a) in two important ways.

First, as discussed by Danek *et al*. (2013), it invokes a global optimization method, which allows us to directly consider tensors of all symmetry classes, regardless of their orientation being described by two or three Euler angles. We note that, as discussed by Kochetov and Slawinski (2009a, b), constraining the local search to obtain absolute minima requires an examination that is possible only for tensors whose orientations are described by only two Euler angles; such tensors are either monoclinic or transversely isotropic.

Second, this direct approach allows us to perturb the values of the original tensor thousands of times to obtain estimates of its effective tensors in the presence of errors. Hence, by using the Monte-Carlo approach, we can estimate ranges of values for elasticity parameters and orientations of effective tensors for cases such as that represented by tensor 6 with errors 7.

Within the assumption of a normal distribution of errors, one could also examine the best fit in terms of likelihood by including errors in the distance function, as considered by Bóna (2009). We use the same errors to perturb the original tensor and – while remaining within coordinate-invariant definition of distance – obtain distributions of solutions, within whose range we would find the effective tensor obtained by the approach of Bóna (2009).

In this study, we confirm and further quantify conclusions obtained originally by Dewangan and Grechka (2003) and examined also by Kochetov and Slawinski (2009a) about the symmetry, orientation, and component values of tensor 8. In particular, we conclude that tensor 6 with errors 7 is consistent – in the Monte-Carlo sense – with the orthotropic symmetry class. Also the results of this paper are consistent with comments of Grechka and Kachanov (2006), according to whom orthotropy might suffice for many scenarios encountered in exploration seismology.

Acknowledgments. This work was inspired by collaboration with the late Albert Tarantola. Also, the authors acknowledge discussions with, and fruitful suggestions of Misha Kochetov, Ken Larner, Daniel Peter, Michael Rochester and Ilya Tsvankin, editorial help of David Dalton and graphic support of Elena Patarini. TD received funding from the Atlantic Innovation Fund and the Research and Development Corporation of Newfoundland and Labrador through the High Performance Computing for Geophysical Applications Project and from Polish National Science Center through grant number 2011/01/B/ST10/07305. MS's research was partially supported by the Discovery Grant of The Natural Sciences and Engineering Research Council of Canada. This research was performed in the context of The Geomechanics Project supported by Husky Energy.

References

Bóna, A. (2009), Symmetry characterization and measurement errors of elasticity tensors, *Geophysics* **74**, 5, 75-78, DOI: 10.1190/1.3184013.

Chapman, C. (2004), *Fundamentals of Seismic Wave Propagation*, Cambridge University Press, Cambridge.

Clerc, M., and J. Kennedy (2002), The particle swarm – explosion, stability, and convergence in a multidimensional complex space, *IEEE Trans. Evolut. Comp.* **6**, 1, 58-73, DOI: 10.1109/4235.985692.

Danek, T., M. Kochetov, and M.A. Slawinski (2013), Uncertainty analysis of effective elasticity tensors using quaternion-based global optimization and Monte-Carlo method, *Q. J. Mech. Appl. Math*. **66**, 2, 253-272, DOI: 10.1093/qjmam/hbt004.

Dewangan, P., and V. Grechka (2003), Inversion of multicomponent, multiazimuth, walkawayVSP data for the stiffness tensor, *Geophysics* **68**, 3, 1022-1031, DOI: 10.1190/1.1581073.

Donelli, M., G. Franceschini, A. Martini, and A. Massa (2006), An integrated multiscaling strategy based on a particle swarm algorithm for inverse scattering problems, *IEEE Trans. Geosci. Remote Sens.* **44**, 2, 298-312, DOI: 10.1109/TGRS. 2005.861412.

Gazis, D.C., I. Tadjbakhsh, and R.A. Toupin (1963), The elastic tensor of given symmetry nearest to an anisotropic elastic tensor, *Acta Crystallogr.* **16**, 9, 917-922, DOI: 10.1107/S0365110X63002449.

Grechka, V., and M. Kachanov (2006), Seismic characterization of multiple fracture sets: Does orthotropy suffice? *Geophysics* **71**, 3, D93-D105, DOI: 10.1190/1.2196872.

Kennedy, J., and R. Eberhart (1995), Particle swarm optimization. **In:** *Proc. IEEE Int. Conf. Neural Networks, 27 November – 1 December 1995, Perth, Australia*, 1942-1948.

Kochetov, M., and M.A. Slawinski (2009a), Estimating effective elasticity tensors from Christoffel equations, *Geophysics* **74**, 5, 67-73, DOI: 10.1190/1.3155163.

Kochetov, M., and M.A. Slawinski (2009b), On obtaining effective orthotropic elasticity tensors, *Q. J. Mech. Appl. Math*. **62**, 2, 149-166, DOI: 10.1093/qjmam/hbp001.

Mathai, A.M., and S.B. Provost (1992), *Quadratic Forms in Random Variables: Theory and Applications*, Statistics: Textbooks and Monographs, Vol. 126, Dekker, New York.

Norris, A.N. (2006), The isotropic material closest to a given anisotropic material, *J. Mech. Mater. Struct*. **1**, 2, 223-238, DOI: 10.2140/jomms.2006.1.223.

Slawinski, M.A. (2010), *Waves and Rays in Elastic Continua*, World Scientific Publ., Singapore.

Stillwell, J. (2008), *Naive Lie Theory. Undergraduate Texts in Mathematics*, Springer, New York, DOI: 10.1007/978-0-387-78214-0.

Tarantola, A. (2005), *Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM, Philadelphia.

Tsvankin, I. (1997), Anisotropic parameters and P-wave velocity for orthorhombic media, *Geophysics* **62**, 4, 1292-1309, DOI: 10.1190/1.1444231.

Tsvankin, I., and V. Grechka (2011), *Seismology of Azimuthally Anisotropic Media and Seismic Fracture Characterization*, Geophysical References Series, Society of Exploration Geophysicists, DOI: 10.1190/1.9781560802839.

Received 30 July 2013
Received in revised form 7 October 2013
Accepted 14 October 2013

DE GRUYTER OPEN

Acta Geophysica
vol. 63, no. 1, Feb. 2015, pp. 62-102
DOI: 10.2478/s11600-014-0206-9

Effect of Stress-Strain Conditions on Physical Precursors and Failure Stages Development in Rock Samples

Kamel BADDARI[1,2,4], Anatoly D. FROLOV[3], Victor TOURTCHINE[4], Fayçal RAHMOUNE[4], and Said MAKDECHE[4]

[1]Laboratory of Physics of the Earth UMBB, Boumerdes, Algeria
e-mail: badari@umbb.dz (corresponding author)
[2]University of Bouira, Bouira, Algeria
[3]Geophysical Division NCG, Russian Academy of Sciences, Moscow, Russia
[4]Laboratory LIMOSE UMBB, Boumerdes, Algeria

A b s t r a c t

Precursory stages of failure development in large rock samples were studied and simultaneous observations of the space-time variation of several physical fields were carried out under different stress-strain states. The failure process was studied in detail. A hierarchical structure of discreet rock medium was obtained after loading. It was found that the moisture reduced the rock strength, increased the microcrack distribution and influenced the shape of the failure physical precursors. The rise in temperature up to 400 °C affected the physical precursors at the intermediate and final stages of the failure. Significant variations were detected in the acoustic and electromagnetic emissions. The coalescence criterion was slightly depending on the rock moisture and temperature effect. The possibility of identifying the precursory stage of failure at different strain conditions by means of a complex parameter derived from the convolution of physical recorded data is shown. The obtained results point out the efficiency of the laboratory modelling of seismic processes.

Key words: stress-strain conditions, failure, physical precursor, rock sample.

1. INTRODUCTION

Numerous experiments have been executed in order to understand the process of deformation and the failure of rock samples in relation with the evolution of mechanical instability within the rock masses during geodynamic processes in the earth crust (Benson *et al.* 2007, Heap *et al.* 2011, Bizzari and Cocco 2006, Kuksenko 2005, Lockner *et al*. 1986, Schubnel and Guéguen 2003, Zang *et al.* 2000). Several physical precursors were recorded as the time of rock failure approaches (Adushkin and Turuntaev 2005, Baddari *et al.* 2011, 2012, Darot and Reuschlé 2000, Heap *et al.* 2009, Jouniaux *et al*. 2001, Kadomtsev *et al*. 2011, Cai and Liu 2009, Sobolev 1995). The aim of seismic process experiments modeling in laboratory conditions was to ensure some sort of similarity to natural process (Baddari *et al.* 1996, Brace and Kohlstedt 1980, Glover *et al.* 1996, Jouniaux *et al*. 2006). In most experiments, it has been concluded that variations of electrical resistivity recorded during the process of deformation and fracture of rock samples could be caused by changes of stress, porosity, saturation, and global permeability of the medium (Chen *et al.* 1993, Lockner and Byerlee 1986, Lockner *et al.* 1986, Ponomarev 1987). A number of laboratory observations show that acoustic emission by rock samples in deformation reflects the process of crack genesis and development. The sources and mechanisms of acoustic emission in rock specimens are the growth and closure of microcracks and macrocracks, collapse of pores, twinning processes, motion of dislocations, *etc*. (Lavrov and Shkuratnik 2005). It is known that the process of rock failure is responsible for the reduction in *P* wave velocity that precedes failure in rock. The dilatancy-diffusion earthquake model predicted a decrease and then recovery of V_P/V_S ratio velocity ratio in the epicentral area prior to large earthquakes (Sobolev 1995, Stanchits *et al.* 2003). Indeed, the variations in electrical resistivity and elastic wave velocities could be, to a certain extent, a mirror of the stress-strain state of a rock and the development of heterogeneity in it. Some interesting laboratory results showed the existence of electromagnetic radiation, emitted from loaded rocks (Baddari *et al.* 2011, Bahat *et al.* 2005). The analysis of the obtained results warrants some assumptions on the origin of the registered electromagnetic radiation and the electrical potential difference during the rock failure. The electromagnetic pulses emitted during the nucleation and the expansion of local mechanical instability in the deformed rock volume can be explained by the formation of charged dislocations, electrokinetics phenomena, dislocation and discharge processes, and other physical mechanisms (Baddari *et al*. 2011). It has been suggested that electromagnetic emission can be generated by oscillating dipoles created by ions moving collectively as a surface wave on both faces of the crack (Frid *et al.* 2003, Lacidogna *et al*. 2011), and the electrochemical effects on the boundaries of minerals grains, *etc.* (Soloviev and Spivak

2009). Changes in physico-mechanical state of a rock caused by the rise of cracks concentration areas disrupt the stationary regime. The registered electrical disturbances can be qualitatively explained within the framework of the ions transfer mechanism (Baddari *et al*. 1999, Sobolev 1995), and the electro kinetic effect when water is present (Pozzi and Jouniaux 1994), *etc*. The analyses of some deformation features of rock samples testify that these mechanisms exist, generally, both in dry and moistened conditions.

Besides, the physical precursors of fracture are subject to the impact of numerous factors such as confining loading, pore water pressure and temperature. Fracture under AE feedback and constant strain rate loading were studied in detail by Baddari and Frolov (2010), Lockner and Stanchits (2002), Smirnov and Ponomarev (2004), and Thompson *et al*. (2006). Feedback loading allowed to study in detail the mechanisms of formation and evolution of fracture process. The role of liquid penetration into a stressed rock was investigated by Evans (2005), Kuksenko e*t al.* (2011), Schubnel *et al.* (2007), Shearer (1999), Sobolev and Ponomarev (2011), and Stanchits *et al.* (2003). It is shown that the water initiation into a rock material causes changes in its various physical features (Gupta 2005, Simpson *et al.* 1988, Sobolev and Ponomarev 2011). It should be noted that, as water penetrates into stressed material, the pore pressure increases causing the dilatancy of the rock. It is known that the acoustic emission (AE) due to water infusion into loaded samples is a possible model of swarm seismicity (Smirnov *et al.* 2010, Sobolev *et al*. 2010). The effects of temperature on the physico-mechanical properties of rock samples and the thermal cracking effects on physical rock properties have been studied by Baddari *et al.* (2012), David *et al.* (1999), Fortin *et al.* (2011), Vinciguerra *et al*. (2005), and Wan *et al.* (2009). The thermal cracking effect was also studied by Reuschlé *et al*. (2006). It has been shown that the behavior of rocks under high temperature was different from their behavior at room temperature. Changes in physical and structural rock properties have been obtained under an increased temperature. At room temperature, rock is generally in brittle failure mode. Changes in elastic modulus, Poissons's ratio, thermal expansion coefficient, and structural chemical reactions have been obtained with the temperature increase. The influence of temperature on various rocks parameters differs from one rock to another. Young's modulus, for example, for rocks like andesite, granite and quartz trachyte, gradually decreased about 20-30% along with a temperature increasing from 20 to 600 °C. Very weak changes have been recorded for the elastic modulus as the temperature increased over 300 °C for rocks like tuft and pottery stone (Xu *et al.* 2008, Zhang *et al.* 2009).

Currently, the problematic of extrapolating laboratory results to nature meets a number of difficulties associated both with technology and methods.

The matter concerns the reliability of the obtained precursors, their universality and the precision of their skill determination**.**

The present work describes the effect of strain-stress conditions on rock macrofailure formation. The use of different large rock samples in our tests permits to explorer the universality of the physical precursors of failure whatever the type of rock, the monitoring of several parameters simultaneously, the differential properties of rocks, the distribution of different physical parameters over the rock volume with using various pick-ups attached to the sample and the failure due to accumulation in the rock volume of elastic deformation. The laboratory experiments carried out in the present work address the study of the effect of stress-strain conditions on the process of rock failure and some spatial resolution of the precursory behavior under various deformation conditions. For this purpose, we analyze the deformation to failure observations on large rock samples in five types of experiments:

I The deformation of dry samples under action of controlled biaxial stress.

II The deformation of moistened samples under controlled biaxal loading.

III The deformation of samples during the variation of strain regime, when the vertical load P was stabilized constant at the stage (0.8 to 0.9) t/t_f and the horizontal load H decreased up to zero.

IV The deformation of samples under conditions of uniaxial load and relatively high strain rate (10^{-6} s^{-1}).

V The deformation of samples under controlled uniaxial load and various temperatures.

Interest of these experiments concerns the fact that a large number of factors affects rock massif in the earth crust conditions, such as mechanical instability, pressure impact, moisture, temperature, *etc*. In order to reinforce our approach, a multi statistical analysis has been applied to summarize information on series of physical failure precursors of a strained rock, which allows to identify a macrofailure precursor by means of a complex parameter.

2. EXPERIMENTAL METHODOLOGY

2.1 Rock samples and deformation apparatus

Large rock samples have been submitted to a series of independent laboratory experiments. Samples of granite, marble, diabase, pyrophyllite and concrete were shaped as a rectangular parallelepiped $450 \times 450 \times 350$ mm in size. Two pairs of glass plates, acting as stress concentrators, of dimensions of $150 \times 150 \times 7$ mm, have been inserted within the sample under an angle of 40° in relation to the vertical axis of the pattern, in order to favor the shear-

ing rupture (Fig. 1). This laboratory simulation can be regarded as a simplified model of the seismic gap process development.

A petrographic analysis gave compositions as follows. The granite contained 42% of albite, 42% of quartz, 39% of biotite, and admixture of muscovite, pyrite and other components, the grain size was 1-2 mm, the density 2.7 g/cm^3, the bulk porosity 2.3%. The marble contained 96% of calcite, 2.5% of dolomite, and 1.5% of other minerals, the grain size was 0.07-1.5 mm, the density 2.7 g/cm^3, the bulk porosity 0.34%. The diabase contained 60% of plagioclase, 20% of olivine, 10% of monocline pyroxene, and 10% of augite, diopside, ilmenite, and apatite, the grain size was less than 1.5 mm, and the bulk porosity was 6%. The sample of pyrophyllite contained more of 45% of silicates of aluminum, the density was 2.7 g/cm^3, the bulk porosity 5%. The concrete sample was prepared out of cement 25%, quartz sand 20%, fine-grained granite 48%, and water 10%.

A servo controlled and AE rate feedback press was used to load rock samples. The apparatus was equipped with a load and a displacement meas-

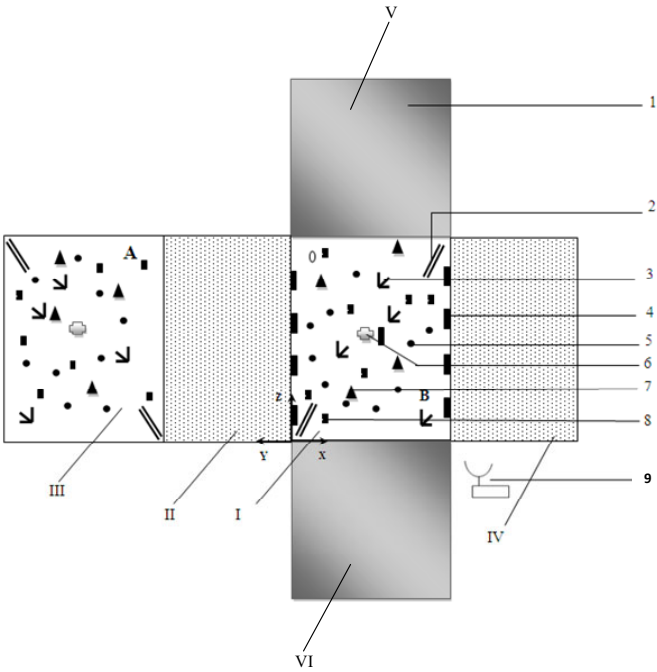


Fig. 1. Schematic diagram of the experiment arrangement. 1 – rock sample; 2 – artificial stress concentrator; 3 – strain rosette; 4 – AE transducer; 5 – elastic wave sound; 6 – thermal sensor; 7 – pick-up electric resistivity electrode; A, B – feeding electrodes; 0 – zero electrode; 8 – self potential electrode; 9 – antenna. I – VI: faces of the sample.

uring systems. The deformation apparatus was linked to a computer-controlled loading in order to measure and maintain the loading force, the preset load rate, the mobile clamp displacement and to record the experimental results. The program system was developed as a programmable feedback controller**.** A transducer was fixed in the middle part of the sample in order to provide feedback for controlling the loading rate as a function of the acoustic activity. The particular load regime has an inverse coupling with AE: with the increase of this latter, the speed of loading is diminished. A feedback between an AE pulse counter and the axial stress ensured a constant AE velocity throughout the trial. The AE activity used in the chain of the inverse command corresponded to the first dozens of events per second. AE rate growth conducted immediately to a decrease in the axial load by the reduction of the motion speed of the piston pressure, making it possible to monitor the failure process under quasi-static conditions. The considerate strain control regime was selected in accordance with the notion relative to the process occurring in the earthquake focal zones. The axial load was measured at the rate of 1s intervals on a PC. Experimental equipment and techniques were described in details in our previous work (Baddari *et al.* 2011, 2012). Laboratory experiments using the inverse loading technique by acoustic emission enabled us to extend the final phase of source development and analyze in details the acoustic regime (Baddari and Frolov 2010, Dresen *et al.* 2010, Lockner *et al*. 1991, Smirnov *et al*. 2010).

2.2 The experiments

Experiments were executed under five strain conditions. The deformation of dry samples under action of controlled biaxial stress has been executed in conditions of a variable vertical load P, and a constant (75 MPa) horizontal load H. Three cycles of stress variation are identified in mean load P as follows: 1 – growth to close to maximum values P_{max} in the interval $(0\text{-}0.45)t/t_f$, where t is the experimental current time, named with respect to a total lifetime t_f of experiment; 2 – constant stress up to $(0.75\text{-}0.8)t/t_f$; 3 – drop of stress due to macrofailure initiation in the interval $[(0.75\text{-}0.8) - 1]t/t_f$. The loading has been realized in the condition of invariability of the AE intensity.

In order to study the role of pore water in the development of failure, similar samples, but sutured with water were deformed under action of controlled biaxial load used in the previous experiment. After drying to constant weight by heating the rock samples at $T°$= 110-130 °C and the air was removed from pores, the samples were placed in water basin in a vacuum chamber during 60 days. The maximum moisture absorption was characterized by constant sample weight. The total saturation corresponded to 0.3-5% moisture content, depending on porosity of each rock sample.

To study the effect of strain regime, the vertical load P was stabilized constant at stage (0.8 to 0.9)t/t_f and the horizontal load H decreased up to zero. Samples similar to that used in previous tests were loaded under conditions of uniaxial load and relatively high strain rate (10^{-6} s^{-1}).

Finally, the same samples were deformed under controlled uniaxial load and various temperatures. In these last experiments, the samples were heated and kept at a given temperature, until the rupture was carried out, by means of a high temperature furnace formed by bifacial heat transmitter. Each sample was heated by an imposed temperature mechanism on two opposite sides at a temperature rate of 2 °C. The temperature was maintained constant during 2.5-3.5 hours, and then the sample was heated from outside, until reaching the selected temperature on the two free sides, where thermal sensors were cemented in order to measure the temperature in the sample under study. Measurements of mechanical, acoustic, and electrical parameters were carried out in the intervals of selected temperatures: 100, 200, 300, and 400°C. Experimental equipment and techniques of the influence of temperature on the kinetics of the deformation, and the rock rupture process were described and discussed in details by Baddari *et al.* (2012).

2.3 Data acquisition and data processing

We have studied the local strain field, acoustic and electromagnetic emissions, and characteristics of elastic waves by means of ultrasonic monitoring, the own electrical polarization and the electrical resistivity (Fig. 1).

A system of three rosettes was fixed right on the line connecting the stress concentrators, *i.e.*, in the route of the main failure and the fourth rosette in the stable area unaffected by the macrofailure. The relative strain readings were measured from vertical ε_z and horizontal ε_x strain components. The first tensor invariant of the plane strain $\Im = \varepsilon_z + \varepsilon_x$ was computed with an absolute error of 5×10^{-5}.

Acoustic signals were received by eight of 10 mm diameter, 1 mm-wide piezoelectric transducers (PZT) of eigenfrequency of 0.2 MHz. The sensors were linked to input amplifiers, oscillographs GRS-6052, and computer. The acoustic events recording was carried out in numerical form of the time of the first arrivals and the amplitudes of the first maxima of signals. The numerical channel resolution for the determination of the arrival time was of 0.05 μs. To pinpoint AE sources, a seismological algorithm (Baddari *et al.* 2011, Smirnov *et al.* 2010) has been used to locate hypocenter coordinates by the difference in AE signals registration time. The acoustic events having a travel time higher than 5 μs were eliminated from the AE catalogue. The uncertainty in determining the hypocenters coordinates was of 1-1.5 mm, according to the acoustic event value. The AE catalogue was composed by the Cartesian coordinates, the time, and the energy class C ($C = 2 \lg A$, where A

is the amplitude of the pulse in mV brought back to a distance of 10 mm from the hypocenter; A^2 is proportional to the energy of the event). Parameter C in this case is similar to the energy class used in seismology. The calculations have been realised in windows of 1000 events and varying by a step of 500.

The electromagnetic emission (EMR) monitoring was performed using a unit of three magnetic antennas of different magnetic permeability of low, mid, and high frequencies, located near the lateral surface of the sample at 0.5 m. The experimental technique and methodology were described in our previous work (Baddari *et al.* 2011). The received signals that occurred in the rock sample as it was deformed were amplified and processed in a formation block. Regular frequency characteristics of the signal reception chain in a tuning fork varied from 400 Hz to 3 MHz. These signals were passed through a quadratic amplitude detector, from which the processed signal arrived to the block of analogue memory where its maximum amplitude was stored in an interval of 10 μs. At the end of this period, a rectangular pulse of positive polarisation of duration of 2-3 μs was formed at the output of an electronic switch. Then, this pulse was directed to the analyser *H*/024-07 in order to obtain the energetic spectra of EMR. To eliminate disturbances on the energetic spectrum, an adjusted compensatory antenna was installed 2 m away from the rock sample in its controlled area. The received signals by this antenna were amplified and compared to the threshold of the triggering off the installation reserved to produce an electric potential for the command block. If the signal from this antenna exceeds a certain level, the command block was blocked and did not transmit signals to the electronic switch, so the undesired signals did not arrive at the input of the pulse analyser. Finally, the investigated frequency range was from first kHz up to 2 MHz.

The ultrasonic wave velocity was determined by compressional transducers with a free resonance frequency from 80 to 100 kHz and $10 \times 10 \times 10$ mm size. The records equipment included a GS13 pulse generator, SP5022 seismoscope, 7402FM frequency meter, GRS6052 oscillograph, and PC. The transducers were used as radiators and detectors of ultrasonic signals. Mainly *P*-wave sensors were used, but these can easily be replaced by *S*-wave ones. Ultrasonic monitoring started under no-load conditions, but as cycled loading progressed, series of measurements along the routes were taken. The arrival time of the first break was used to calculate the ultrasonic velocities. The sensors were fixed on the faces of samples and the number of monitored ray paths was 30; enabling to study velocity variations in measuring *P*- and *S*-wave arrival times with an accuracy of about 0.7%. The elastic wave velocities have been measured following the ultrasound traces under angles of 45° with respect to the side *I* of rock sample allowing to record *P*- and *S*-waves.

The electrical field was measured by nonpolarizable silver chloride electrodes enclosed in special holders, which were mounted on the free faces of the specimen. Their potential stability did not exceed 0.1-0.2 mV. Silver chloride electrodes were used as supply electrodes, both tapping off the 2 V power source. Rock samples were polarized during 30 s. After the current interruption by the disconnection of the power source, the electric potential was registered. Different combinations of pick-offs served as receiving pairs of electrodes MN and enabled to study the structure of the electrical potential in detail. Measurements were taken regularly every 1-2 min. The background values had been estimated prior to the experiment by carrying out an aerial survey of surface potentials on a 2×2 cm grid on all free faces of the specimen. As a result, it was possible to detect changes of the order of 1 mV and less in the self potentials.

The apparent electrical resistivity monitoring was studied using direct current and electrode devices in different orientations. Various combinations of the nonpolarizable sensors acted as dipoles. The feeding graphite electrodes, diameter in 2 and 1 mm thick, were attached on free surfaces of the rock sample. A stabilized current to 0.01% has been injected through these electrodes. While waiting 30 s, the measurements started when the polarization effect dropped to 0.5% of a measured value. Measurements in continuous and automatic regime of the apparent electrical resistivity were realized at an electrode pairs. It should be noted that the measurements realized during ten days in unloaded samples showed constancy in electric parameter measurements down to 0.5-2% precision. The methodology enabled us to evaluate relative changes of resistivity both in the entire sample and in separate parts of it. Dozens of electric surveys were conducted during loading cycle. The apparent resistivity was obtained using the formula $\Delta\rho_a = (\rho_t - \rho_0)/\rho_0$, where ρ_0 is the value of ρ_a before containment, and ρ_t is the value of ρ_a at the time t. In all cases, the used technical methods and continuous measurements insured high sensitivity and reliability of the apparent electrical resistivity parameter. Similar technical methodology of electrical parameters measurement was used by Baddari *et al.* (1999, 2012), Ponomarev (1987), and Sobolev (1995).

3. RESULTS

3.1 Defect accumulation and hierarchical failure process in dry samples

Structural discontinuity of rock massif in the earth crust is due to seismogenic and volcanic activities (Ammon *et al.* 2008, Kuksenko *et al.* 2011, Shearer 1999). Microscopic examination of sections of samples and thin plates, after experiments using dry samples, was used to study the specificity of the failure in each deformed rock. For samples of granite, diabase, and pyrophyllite,

Fig. 2. Schema of the location of cracks in rock samples after the experiments in dry strain conditions: granite (a), marble (b), diabase (c), pyrophyllite (d), and concrete (e).

we noticed cracks systems of h 0.1-0.2 mm length with a dominant orientation angles in the range of 30-45° and 55-60° with respect to the vertical stress P. Short subparallel cracks oriented under angles of 20-35° in relation to the main crack were observed in the diabase and pyrophyllite. Marble sample showed shear failures under an angle of 35-60°. The main macrofailure was a result of formation of subparallel system of microcracks in the granite and diabase samples. Microcracks were oriented under angles of

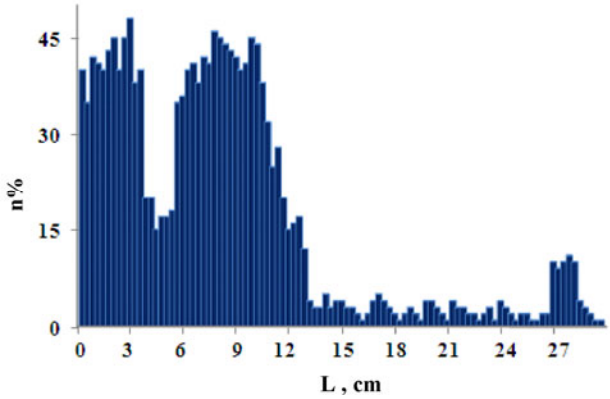


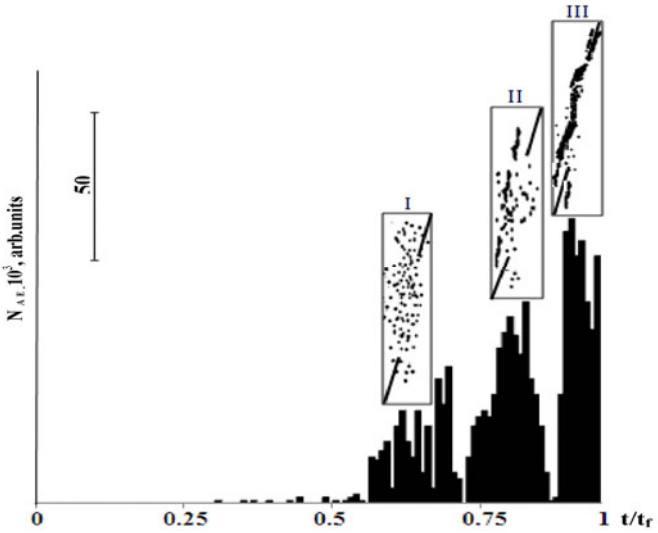
Fig. 3. Histogram of mean geometric size distribution of adjacent failures in diabase sample.

30 to 50° in the concrete sample. Figure 2 gives schemes of observed, according to a visual examination, of macrocracks in all samples after the experiments. It shows a blocky hierarchical formation. The ratio of two basic adjacent sizes, for all samples, during and after experiments, gave $L_i/L_{i-1} = \beta = 10^{\alpha}$, where $\beta = 2.5\text{-}3.5$ and $\alpha = 0.01\text{-}0.08$. Figure 3 shows, as an example, a diagram of mean geometric size distribution of adjacent failures in diabase sample. The histogram shows three maxima and the ratio of adjacent blocks mean sizes is in the order of 3. Similar results have been obtained by Asatryan *et al.* (1993). This structural discreteness reminds that after a seismic shock or anthropogenic activity, the continuity of rock massif is an object of considerable heterogeneity of faults, inter-block contacts and various fractures, which will act as stress concentrators under the action of natural stresses. This testifies to the self-similarity of failure process at different scale levels in agreement with Baddari and Frolov (2010), Sadovsky *et al.* (1991), Sobolev (1995) that binds the rupture process structure in the geophysical field to some characteristic sizes.

Acoustic and electromagnetic emissions produced in the process of deformation and failure were used in studying the redistribution of cracks in the fracture process accumulation. As an example, Fig. 4a shows changes in the number of acoustic events for the macrofailure zone in the granite obtained between the stress concentrators at different strain-stress stages. We note a gradual rise in the number of acoustic signals reaching its peak before the failure. The projection of the AE coordinates allows to reveal three characteristic stages of the focus genesis. Stage I corresponds to the formation of independent dispersed microcracks in the volume rock sample. The size of cracks at this stage is of the order of those grains (Sobolev and Ponomarev 2003). Characteristic microcracks dimensions corresponding to the mini-

mum acoustic signals amplitudes of 8-10 mV was about 70-90 μ in concordance with the results obtained by Baddari *et al*. (2011) and Ponomarev *et al.* (1997). Microscopic measurements gave approximately the same lengths.

(a)



(b)

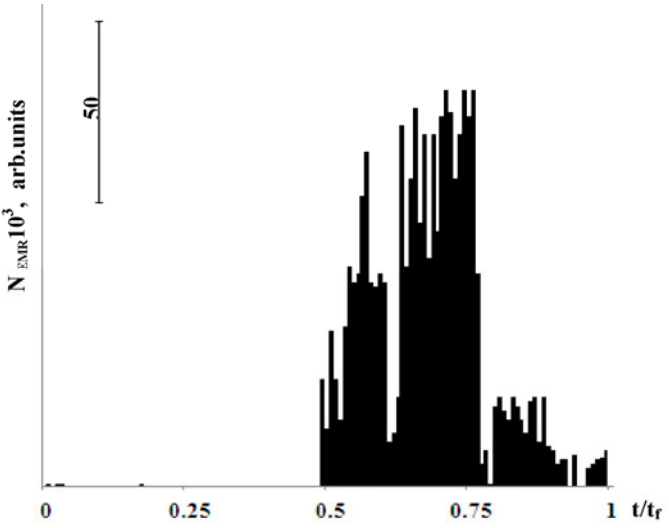


Fig. 4: (a) Energy released spectra of AE during the deformation of granite sample. Schemes of cracks hypocenters are shown above the plot. I – III – three stress-strain stages process; (b) Typical record of the total added released EMR energy in diabase sample.

It should be noted that the size of cracks emitting the lowest acoustic signals is calculated by the formula (Baddari *et al*. 1996, Ponomarev *et al.* 1997): $l_i = \beta A^{2/3}$, where A is the signal amplitude and β is a constant. The minimum amplitudes of the AE recorded during our tests were 8-10 mV. Stage II corresponds to an increasing accumulation of microcracks leading to their concentration and the formation of close microcracks sets able to interact and unify. The latter gave rise to defects with dimensions by an order of magnitude larger. Stage III is characterized by a localization of the main failure focus, which will develop the macrofailure. Figure 4a shows two acoustic quiescence periods at 0.71 and 0.9 [t/t_f], corresponding to the scenario preceding the respective transition from lower (i – 1) to higher (i) and from (i) to (i + 1) levels of cracking hierarchy. Using the formula

$$\log l_i = a \log E + b\,, \tag{1}$$

with $a \approx 0.25$, we can deduce that the variation of the energy E of an order of magnitude induced a twofold increase in the length l_i of the crack (Baddari *et al*. 2011, Sobolev and Ponomarev 2003). The decay in acoustic activity at each loading step obeys generally the Omori law (Lavrov and Shkuratnik 2005, Lockner 1993, Rudajev *et al.* 1996, Smirnov and Ponomarev 2004, Utsu 2002)

$$\Delta n / \Delta t = A_0 (t + c)^{-p}\,, \tag{2}$$

where A_0 is the initial number of acoustic signals in a time unit, c and p are constant values, and t denotes the time. An estimate of $0.55 \le p \le 2.0$ for values of p has been obtained. Utsu *et al.* (1995) obtained $0.6 \le p \le 2.5$ in relation with aftershock activity. Note that A_0 increases and the parameter p decreases prior to the progressive macrofailure. The decrease in p can be understood as a prediction criterion of the expected rock instability.

The experiments show that the EMR amplitude A changed from 10 to 120 mV. Figure 4b gives a typical record of the total added released energy summed up in 100 s interval of EMR in the stress concentrators zone of the diabase sample. The energy of an electromagnetic event was estimated from the amplitude of the first maximum of a signal. The first phase, corresponding to a gradual rise in differential loading, is characterized by a low intensity of EMR events, which begin to be recorded at the end of this phase at t/t_f = 0.48-0.5, and reflects the beginning of weak microcracks process. The stage from 0.49 to 0.78 [t/t_f] is marked by numerous minor cracks accumulation and is characterized by rising EMR intensity. The non-uniformity of this intensity reflects the relaxation character in the variation of the stress concentration in the sample which has been shared in robust and fragile zones. At the time of loading t/t_f = 0.78-0.9, the EMR variations decreased reflecting a gradual concentration of cracks around an incipient macrofailure. The

last phase 0.9-1 [t/t_f] unfolds against a falling EMR and the main fault taking shape. The dynamic of increase and relaxation of EMR, in our opinion, reflects the growth and closure of microcracks, and dislocations motion at grains boundaries and grains contacts in the rock body. The generation of minor cracks gives rise to high frequency EMR signals of no less 10-15 μs duration and more than 300 kHz base frequency; however their fusion upon reaching a critical concentration leads to low frequency EMR pulses (10-20 kHz) and 150-350 μs duration. We noticed a gradual increase in the amount of the last signals about 40-60 min prior the main shock triggering.

The analysis of registered EMR pulses allowed to classify arbitrarily the EMR energy into three classes: $C_1 = 1.2\text{-}2.2$, $C_2 = 2.2\text{-}2.8$, and $C_3 \geq 2.8$, which correspond to three sizes of cracks l_1, l_2, and l_3 (Eq. 1). Three classes were selected according to the recorded number of electromagnetic signals N_1, N_2, and N_3, respectively. Figure 5 shows an example of the variation of EMR for the three classes $C_1 - C_3$ in the experiment with the pyrophyllite. It gives the curve of loading and the number of EMR pulses per unit of time N, normalized to their ultimate values. We noted that phases of EMR evolution have been maintained. The EMR showed an electromagnetic quiescence more visible and longer for C_2 and C_3 at $t/t_f = 0.94\text{-}1$ and at $t/t_f = 0.96\text{-}1$, respectively (Fig. 5b, c). An electromagnetic quiescence is likely to occur only under a definite relationship between energies of registered minor shocks viewed as a background for the impending macroshock.

The possible physical explanation of the delocalized failure ultimate concentration can be based on the equation of the solid state strength kinetic theory (Kuksenko 2005, Sobolev 1995, Zhurkov 1984)

$$t = t_0 \, e^{(E_0 - \gamma\sigma)/kT} , \tag{3}$$

where E_0 is the failure activation energy, close in value to interatomic bond energy, k is the Boltzmann constant, t_0 is the pre-exponential factor with value equaling that of solid state atoms heat fluctuation period, and γ is a structurally sensitive parameter defining local overstresses. Equation 3 expresses the relationship revealed between the lifetime t, the tensile stress σ, and absolute temperature T. For the case of a constant temperature, Eq. 3 may be transformed to a simpler relation for the lifetime

$$t = B \, e^{-\alpha\sigma} , \tag{4}$$

where $B = t_0 \, e^{U_0/kT}$ and $\alpha = \gamma/kT$. By analogy to the kinetic equation, it is possible to assume that the rate of crack accumulation dN/dt can be defined as follows:

$$dN / dt = (dN / dt)_0 e^{(E_0 - \gamma\sigma)/kT} , \tag{5}$$

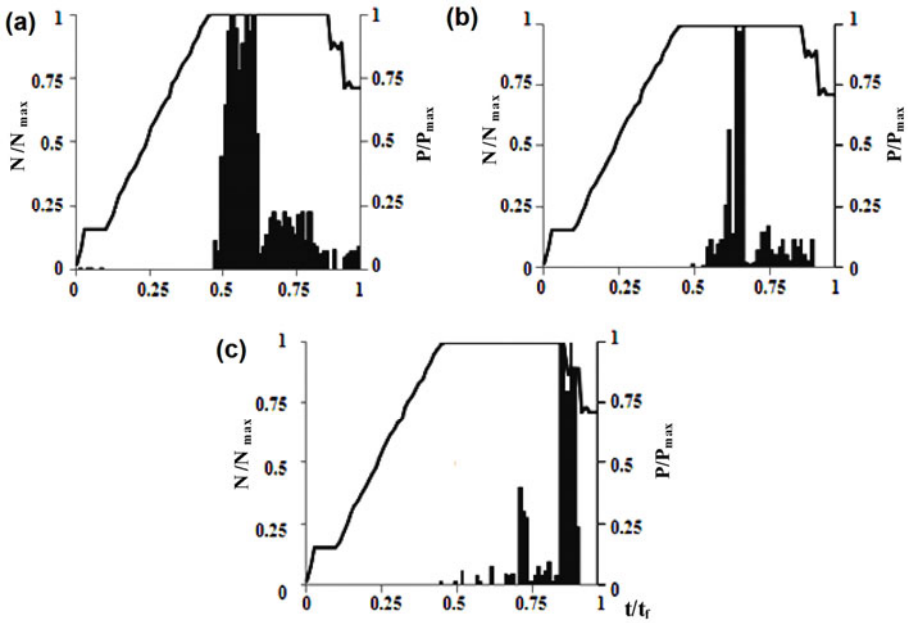


Fig. 5. Variations of EMR for energetic classes: (a) C_1, (b) C_2, and (c) C_3 in pyrophyllite sample. N/N_{max} and P/P_{max} represent the number of EMR pulses per unit of time N and the load P normalized to their ultimate values respectively.

which testifies that the process of kinetics of crack accumulation determines the lifetime t of a stressed rock which is made up of the duration stage deformation t_i and that t_s of a failure source rise $t = t_i + t_s$, with $t_s < t_i$. The failure focus rise is a result of the microcracking transition from delocalisation to localisation, and consequently changes in the physical subsystems structure of rock material. The associated physical subsystems, conditioning the response to external mechanical influences, are ionic, dipolar, and electronic holes. The changes in the spatial structure are accompanied by some changes in mechanical, electrical, and other physical properties of the medium, most of which can be recorded by remote sensing methods.

Results of experimental investigations of failure show that structures cracks sets obey the fractal statistics (Baddari *et al.* 2011, 2012; Sobolev 1995, Zavyalov 2006, Zhurkov *et al.* 1980). In a chaotic distribution of microcracks, the probability of the formation of a set of x adjacent cracks, *i.e.*, the transition from delocalization to localization, is

$$p_n = \frac{\langle x \rangle^x}{x!} \tag{6}$$

with the mean number $\langle x \rangle$ of cracks by cluster equal to $\langle x \rangle = x/K$ and the mean inter-crack distance K during their volumetric concentration per unit of size $N = xV^{-1}$ equal to

$$K = N^{-1/3} / \overline{l} \,, \tag{7}$$

where V is the volume of the confined rock, and $\overline{l}$ is the mean size of the accumulated cracks given by

$$\overline{l} = \frac{1}{x} \sum_{i=1}^{x} l_i \tag{8}$$

with l_i the size of the crack. Equation 6 reminds the Poisson formula without the multiple $e^{-\langle x \rangle}$. The use of the formula of Sterling for the factorial of a large number $x >> 1$ in Eq. 6 leads to

$$(2\pi x)^{1/2} x^x e^{-x} \le x! \le (2\pi x)^{1/2} x^x e^{-x} \left(1 + (12x)^{-1}\right). \tag{9}$$

As $\sum_{x=1}^{\infty} p_x = 1$, in the case of a dense set fluctuation, we have

$$(2\pi)^{-1/2} \le \sum_{x=1}^{\infty} \left(\frac{e}{K}\right)^x \tag{10}$$

and the criterion of clustering cracks to generate a quick failure of the rock is

$$K \le e\left(1 + (2\pi)^{-1/2}\right). \tag{11}$$

The absolute error ΔK is $\Delta K \le 0.5 \ 10^{i-j+1}$ with $i - j + 1 = -8$ and consequently the exact value of the calescence criterion is $K^* = 3.8$, which constitutes the threshold of the transition from the stage of crack accumulation to their start to coalesce and the formation ensembles of clusters in the cracked body.

Since we did not measure the crack size l_i (Eq. 8) during the experiments, the methodology explained in Baddari *et al*. (2012), Smirnov *et al.* (1995) has been used, and therefore

$$K = x^{2/3} V^{1/3} \Big/ \beta \sum_{i=1}^{x} A_i^{2/3} \,, \tag{12}$$

where A_i is the amplitude of the acoustic signal and β is a constant, evaluated empirically equal to 0.2 mm/mV $^{2/3}$.

The location of the failure was controlled by the Pearson criterion

$$\chi = (N_i - x_i)^2 / (x_i)^2 \,, \tag{13}$$

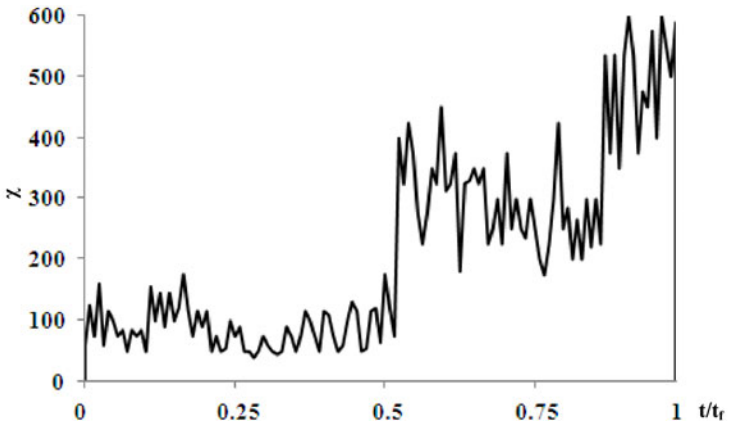


Fig. 6. Variation of the Pearson criterion during the deformation of concrete sample.

where x_i is the number of cracks in the sample unit volumes for an uniform distribution of acoustic signals, *i.e.*, received signals in the specified amplitudes range $A + dA$, and N_i the real number of signals in these volumes, *i.e.*, all emitted signals. The log$A$ values were divided into ten categories and the number of signals was calculated for each category. As a typical example Fig. 6 shows the variation of χ for the deformed sample in experiment with the concrete. The value of χ remained almost constant until $t = 0.5$ $[t/t_f]$ and then increased sharply up to $t = 0.8$ $[t/t_f]$, which indicated the location of AE. A second growth is recorded from $t = 0.9$ $[t/t_f]$ until the end of the test, which confirms the hierarchical progression of the failure. The stochastic nature of the EA has been preserved.

3.2 Parameters of physical fields in deformed dry samples under a strain control regime

Experimental data have yielded a series of stable variations in physical parameters depending on the strain-stress state of the rock. Figure 7 shows the experimental results with dry samples. In this case, the samples have been heated in kilns at temperatures of 110-130 °C until they reached constant weight. Figure 7 Ia shows the variation of the first flat strain tensor invariant $\mathcal{J}$ for rosette situated in the affected area by the macrocrack. We can identify three stages of deformation illustrating the different stages of the deformation process evolution. The first stage, A, from 0 to 0.5 $[t/t_f]$ corresponded to the stage of separate stable microscopic cracks accumulation in the entire solid body. The second one, B, was characterized by a relative slowdown in growth of $\mathcal{J}$ and its transition by an extremum. It has been observed in the area of maximal stresses corresponded to the irreversible rock deformation.

It has been characterized by microcracks growth process and the beginning of their mutual interaction triggering finally a macrofailure nucleus in the area of an incipient macrofailure. The final stage, *C*, characterized by negative deviations from the X-axis, unfolded during the step of failing stress while the macrofailure nucleus increased and the main fault was shaped.

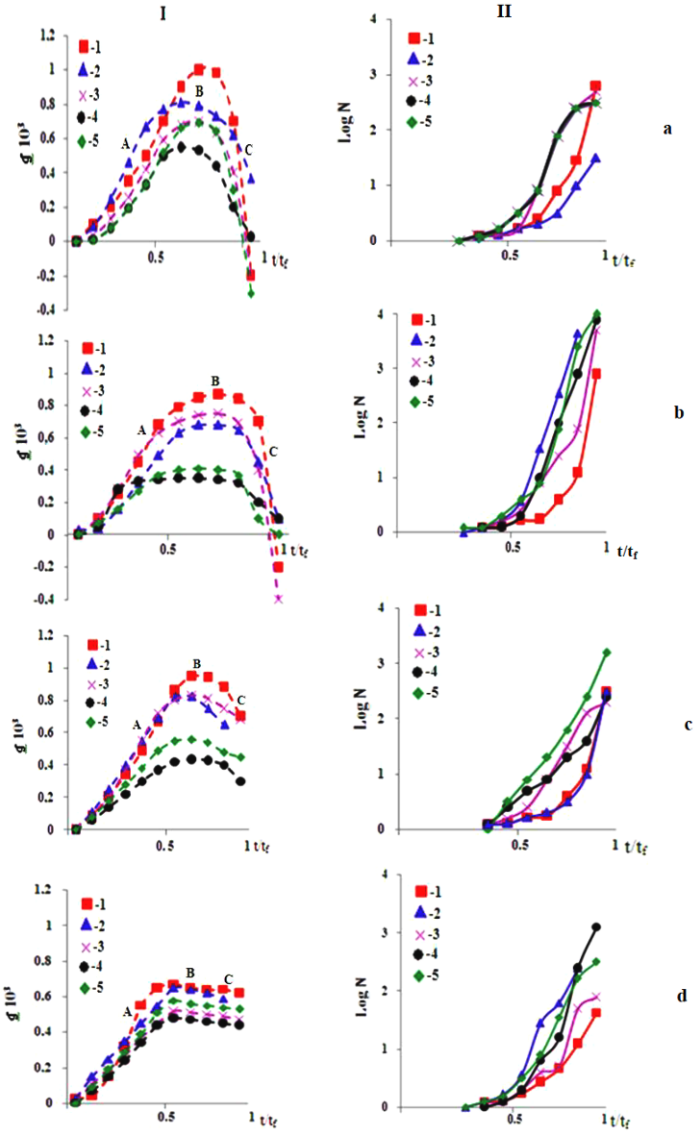


Fig. 7. Continued on next page.

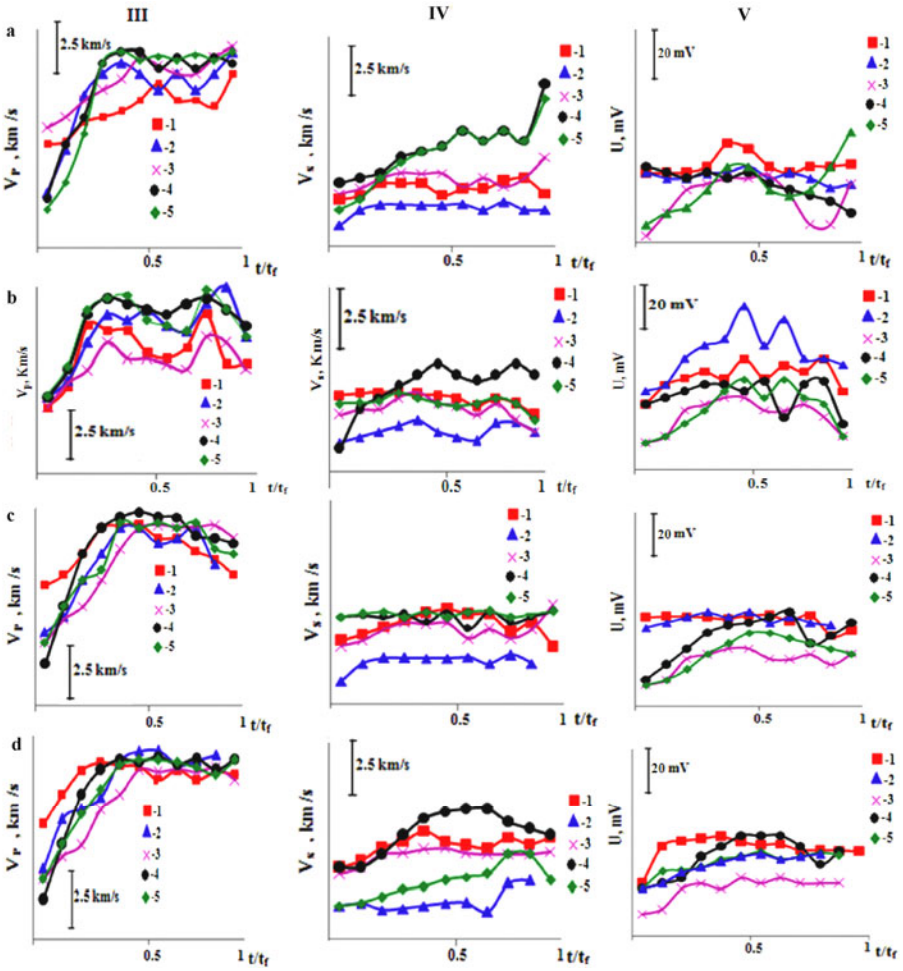


Fig. 7. Variations of physical parameters during the deformation and the failure rock samples under various strain conditions: dry samples (a), moistened samples (b), variation of the strain regime (c), and deformation at high strain rate (d). 1 – granite, 2 – marble, 3 – diabase, 4 – pyrophyllite,5 – concrete. I – first tensor invariant of flat deformation, II – AE intensity N, III – V_P, IV – V_S, V – electrical potential ΔU.

It should be noted that the positive deviations of $\mathcal{J}$ from the X-axis point to the element square contraction, while negative – dilatancy (Baddari *et al.* 2012, Sobolev 1995).

The initial stage (up to 0.4-0.5 t/t_f) is characterized, generally, by low activity of the AE received by all the transducers (Fig. 7 IIa), strong increase of V_P and low variation of V_S for routes situated along the macrofailure path

(Fig. 7 III-IVa), and a constant or negligible variation of the electrical potential ΔU for the pyrophyllite and concrete samples, and growth of this latter for diabase and marble samples. A slight increase of ΔU was noted for the granite at the end of this stage (Fig. 7 Va). These changes are uniform across the sample. This step corresponds to a material sample consolidation and the beginning of microcracking. This stage is characterized by the extensively nonuniform and isolated spatial distribution of microcracks. The second stage of the sample deformation (0.5-0.8 t/t_f) is characterized by a growth of the AE for all samples, bay – shaped oscillations, and drops in variations of V_P and V_S, which accompanied the dilatancy, the decrease of ΔU for granite, diabase, pyrophyllite, concrete, and a decrease followed by growth of this latter for the marble. An increase by intervals in the activity of the AE was observed. This step corresponded to an intense sample micofissuration. The third stage (over 0.8-0.85 t/t_f), characterized by a slow deformation of the sample, is accompanied by sudden increases of the AE for all samples, increases for granite, marble, diabase, and concrete samples, during the last phase of this stage (0.9-1 t/t_f) in V_P, a decrease of V_P for the pyrophyllite, increases of V_S for the diabase, pyrophyllite, and concrete samples, a decrease of V_S for the granite and its negligible variations for the marble. This deformation stage has increased ΔU in the marble, diabase, concrete and granite during the last step. Drop in ΔU was recorded for the pyrophyllite just before the end of the test. The analysis of the distribution of acoustic signals showed a cluster of AE hypocenters, which testifies to a step of the progressive macrofailure and the beginning disintegration phase in the rock material. We noted a dispersion of V_P and V_S velocities for the second and third stages caused probably by the increase volume of inhomogeneity and anisotropy due to the fracturing propagation, tensile, shears and various structural damages in stressed sample. An important elastic anisotropy was developed prior to failure, which was due to oriented microcracks, their concentration and their distribution pattern. The mean square velocity δ deviation determined for 15 routes shows that the increase in stress is accompanied by increased velocities scatter (Fig. 8). The failure occurred differently in the different zones of intersection of theses paths. Figure 8 demonstrates, as an indication, the behavior of the ratio V_P/V_S velocities variations for sets of elements 1-3 and 2-4 situated along and across the stress-concentrators, respectively in a granite sample. The behavior of the V_P/V_S ratio through the trace 2-4 was antagonistic that 1-3, which reflected the anisotropy of strength and the distribution of mechanical stresses in deformed rock sample. Roughly similar graphs were obtained for the other rock sample. It has been noted that the dynamics of cracking structure formation is more active in the vertical direction than in the horizontal one for all rocks. The ratio V_P/V_S

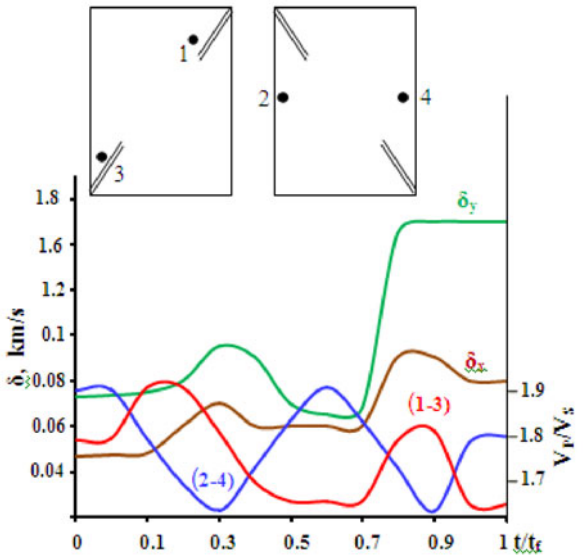


Fig. 8. V_P/V_S ratio observed in macrofailure area for routes 1-3 and 2-4. σ_y and σ_x are the mean square velocities computed according to the vertical Y and horizontal X-axis of the granite sample.

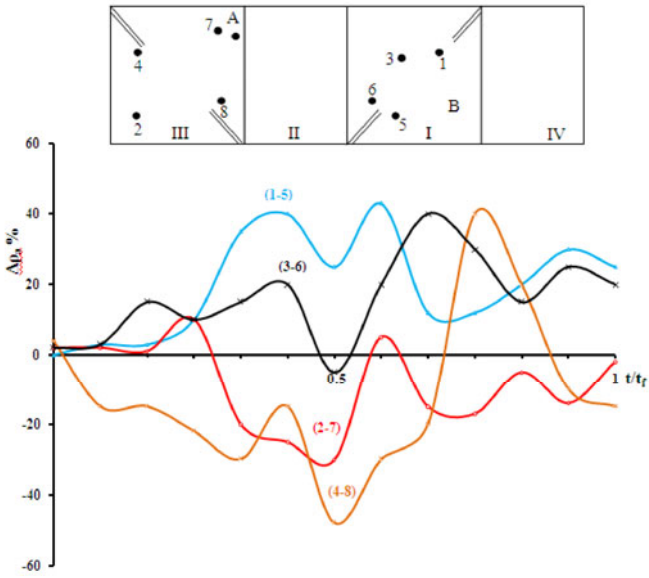


Fig. 9. Typical anisotropy registered in electrical resistivity measured for different positions of pick-up electrode pairs in granite sample. Scheme of four (I-IV) sample facets with the introduced stress concentrators, pick-up electrodes: 1, 2, 3, 4, 5, 6, 7, and feeding electrodes A, B are shown above the plot.

showed a negligible change in the stable zones for all samples. Similar, on the whole, results have been obtained in our previous work (Baddari and Frolov 2010). Anisotropy in $\Delta\rho_a$ was recorded during the half of the first stage of deformation between 0.3 and 0.5 [t/t_f], the second anisotropy at (0.6-0.8)t/t_f and last one at (0.8-1)t/t_f of rock failure (Fig. 9). A bay variation of $\Delta\rho_a$ has been obtained in the routes passing through the jumper center between the stress concentrators in the experiment with the granite sample. The bay's reversal was recorded in perpendicular routes. The increase of $\Delta\rho_a$ in perpendicular direction is probably caused by coalescing oriented cracks and rupture of conducting paths system during the macrocrack formation. Similar anisotropies in $\Delta\rho_a$ were recorded for all tested rocks.

3.3 Influence of various factors on the physical precursors

3.3.1 Influence of moisture

Moisture extended the second strain stage B recorded in the parameter $\mathcal{J}$ (Fig. 7 Ib). The trend of the AE was affected by moisture (Fig. 7 IIb). We noted more intense AE activity for the moistened rocks. The V_P and V_S velocities marked a sensible drop during the second stage of rock deformation at 0.4-0.65 t/t_f. Some anomalies were found in the variations of longitudinal V_P and transversal V_S wave velocities during the stage of progressive macrofailure (Fig. 7 III-IVb). Also, we noticed a growth and relaxation for tens of minutes in duration in the electrical parameter ΔU during the second and the third stages of deformation, reflecting an infiltration, a twinning processes and an evaporation of water within the cracks structures. Significant increases in ΔU have been recorded during the second stage of deformation macrofailures in all samples. This dynamics in ΔU did not correlate with variations in the applied stress and indicated the process of crack genesis and activation energy. The anomalies in electrical structure continued for dozens of minutes after a rock sample macrofailure testifying to a low velocity of relaxation processes. The main strain steps recorded during the fracture process of dry samples have not changed. However, the strength of moistened rocks decreased by 20 to 30% compared to that of same dry rocks. The decrease of strength of the moistened rock sample is caused by the increase of cracking rate and the adsorption phenomena. Signals of max AE amplitude arise and reach the maxima in the second stage of deformation. Figure 10 gives an example of acoustic signals recorded in granite sample in dry and moistened conditions. It is possible to conclude that AE activization appeared as result of pore collapse, and water diffusion inside the active cracks. The gradual migration of pore water to the microcracks field increased the stress around microdefects leading to an increase in the rate of cracking and their rapid development. The obtained AE behavior is similar

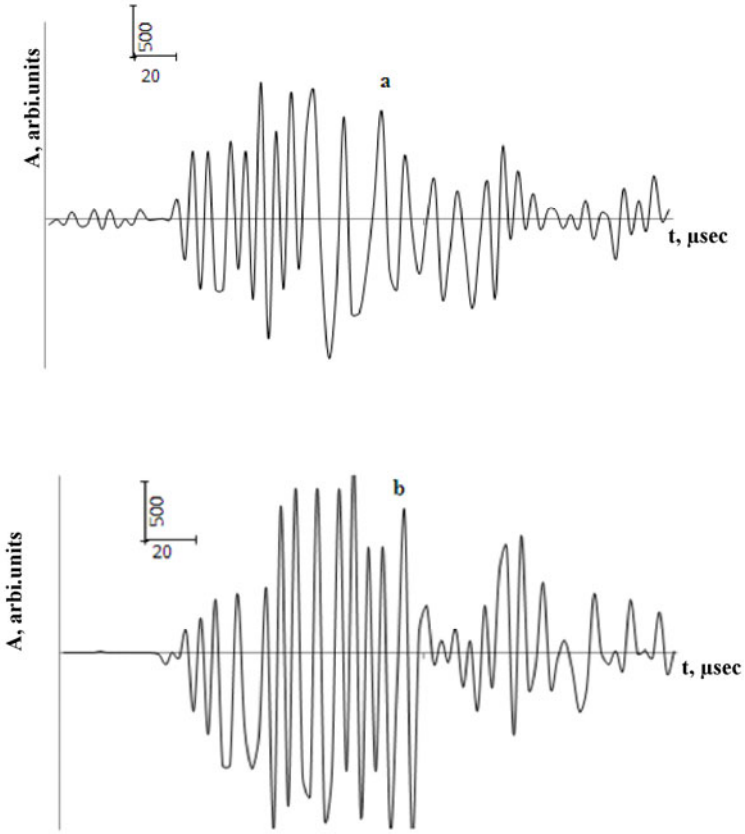


Fig. 10. AE signals recorded during the deformation of dry (a) and moistened (b) granite samples.

to the activity of earthquake swarms recorded in seismically active regions (Sobolev *et al*. 2010). In general, the moisture of crystalline rocks influenced the physical precursors of failure.

3.3.2 Influence of the deformation regime

Figure 7 I-Vc represents the experimental results during the variation of the deformation regime when the vertical load P was stabilized constant at stage (0.8 to 0.9)t_f and the horizontal load H decreased gradually until it became zero. We noted shrinkage in the third stage, C, of the deformation parameter $\mathcal{J}$, a strong increase of AE, an increase in AE during the second and third strain stages, an increase in V_P during the first stage of deformation and its decrease during the two last stages, non-noticeable variations of V_S. Electrical potential ΔU did not show a clear trend and its variation was generally

unremarkable, except for the pyrophyllite and concrete rock samples. As a result, a marked appearance of physical precursors in V_S and ΔU parameters was attenuated at this stage

3.3.3 Influence of uniaxial load and relatively high strain rate

Figure 7 I-Vd shows the experimental results when rocks samples were loaded under uniaxial stress at constant high strain rate ($10^{-6}s^{-1}$). A nearly linear trend of $\mathcal{J}$ parameter was recorded at the first strain stage, A, and its constancy at the second and third stages, B and C, exponential increase of AE during the stages, B and C, constancy in V_p during the second and the third stages of deformation, a decrease in V_S was recorded for the pyrophyllite during the third strain stage, ambiguous variations of V_S were unregistered for the rest of rocks and insignificant changes in ΔU compared to previous cases. It was difficult to identify the final stage of macrocracking by the parameters $\mathcal{J}$, V_P, and ΔU, however it was possible to identify the initial and the final stages of the failure by the AE. The measurements showed that duration of an EMR single pulse did not exceed 10 mcs.

3.3.4 Influence of temperature

The effect of temperature has an impact on the deformation of the rock. Results on temperature effect on the rock failure precursors were described in our previous work (Baddari *et al.* 2012). The brittle shear failure predominated at high temperatures. The $\mathcal{J}$ curve drifts leftward with the increase of temperature (Fig. 11 II-IIIa). This fact indicates that the increase of temperature led to an increase in the quantity of active cracks. The transition to the cracking phase was relatively fast at temperatures of 300-400 °C testifying to the brittle character of its rupture (Fig. 9 Ib, c). A shrinking of the maximum of the curve $\mathcal{J}$, with respect to the t/t_f-axis, was generally marked at high temperatures, for the rock samples, reflecting the weak occurrence of the quasi-plastic deformation of the sample. The V_P/V_S ratio at high temperatures (Fig. 11 II-IIIb) marked a significant decrease during the microfissuration stage. An extremum narrowing has been noted at 300 and 400 °C for all samples at the phase 0.4-05 t/t_f. The high temperature has increased the electrical parameters $\Delta\rho$ at 400 °C (Fig. 11 IIIc) and decreased ΔU at 300-400 °C (Fig. 11 II-IIId). However, at a temperature of 400 °C (Fig. 12), the amplitudes and periods of the EMR pulses have reached the maximum in the middle of the second stage of strain-stress rock state**,** which differs from the variation of these parameters under the conditions of room temperature. During the formation phase of isolated microcracks, the amplitude and the pulse period of EMR are much greater than those of AE. The

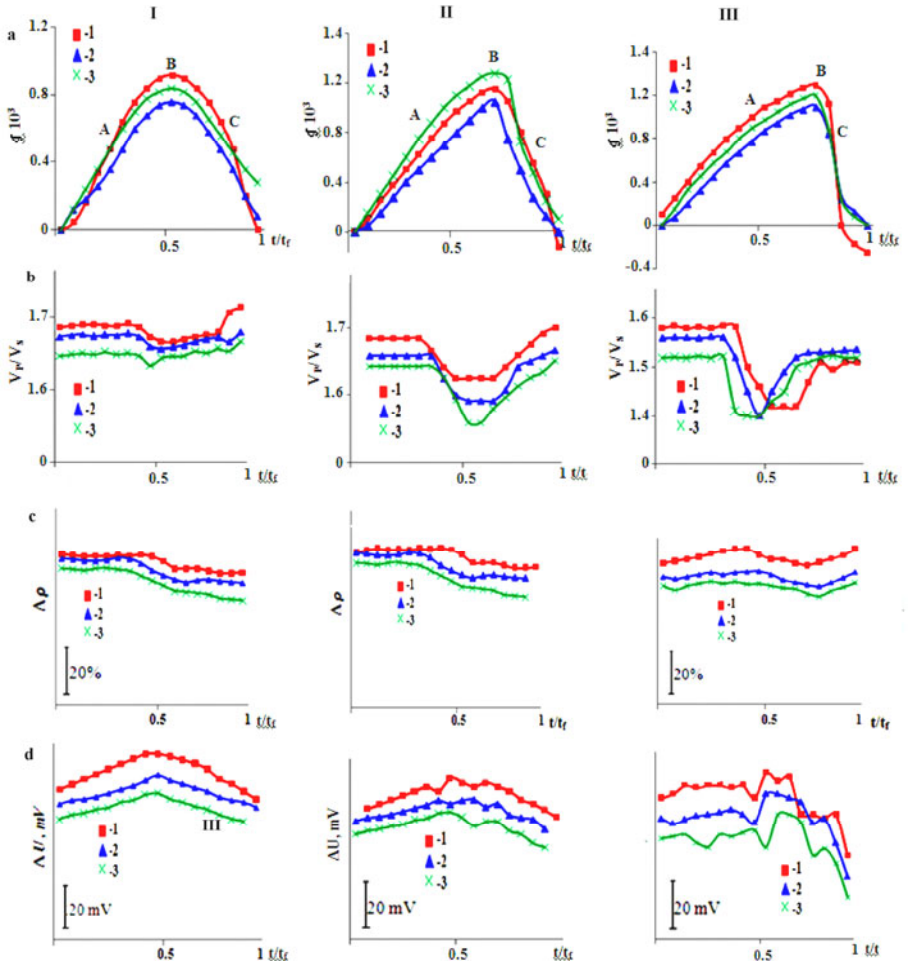


Fig. 11. Influence of temperatures on the physical precursors of the failure: first tensor invariant of flat deformation (a), V_P/V_S ratio (b), apparent electrical resistivity (c), and electrical potential (d). I-III – data obtained at T = 100-200, 300, and 400 °C respectively. 1 – granite, 2 – diabase, 3 – pyrophyllite.

synchronized recording of EMR and AE intensities showed that their variations have been influenced by applied temperature. At 400 °C EMR and AE events have been recorded practically at 0.3 [t/t_f] reflecting an early sample microfissuration (Fig. 13). In conclusion II curves of AE and EMR activities were drift rightward with increased temperatures. The gradual rise of load during the first phase of experiment was characterized by a low intensity of EMR and AE at 25 °C. The beginning of microcracking process clustering at

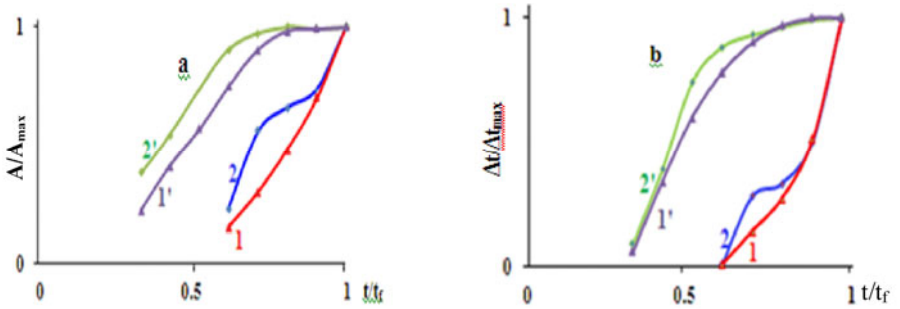


Fig. 12. Variations of the normalized amplitudes and the periods of AE (1) and EMR (2) at 25 °C and AE (1') and EMR (2') signals obtained at 400 °C for pyrophyllite sample.

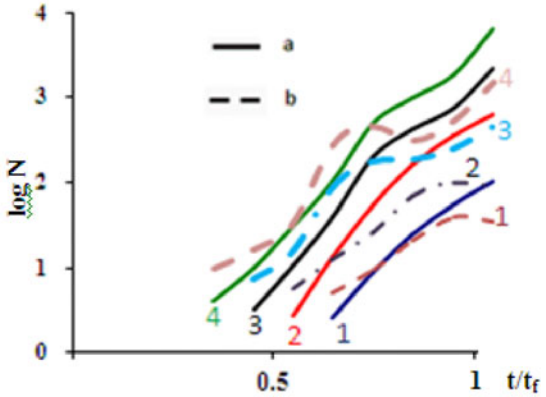


Fig. 13. The synchronized recording of AE (a) and EMR (b) intensities obtained at different temperatures. 1 – 25 °C, 2 – 100 °C; 3 – 300 °C, 4 – 400 °C for pyrophyllite sample.

t/t_f = 0.6-0.8 was characterized by a rise in EMR and AE intensities. Strong EMR signals were recorded in this phase reflecting the growth and relaxation dynamics of the stress distribution in the sample. Microscopic measurements and integrated recording of AE and EMR realized in the granite sample showed that small defects of 0.8-2.2 mm in size produced short acoustic signals (<500 μs duration), and longer EMR pulses (>1 ms), however shear fractures of 15-40 mm trigger longer AE pulses and short EMR ones. This result allows to conclude that large shear cracks seems have low EMR generating capacity. In general, the increase of the temperature of the deformed rock accelerated the microcracking accumulation and increased the EMR and AE intensities.

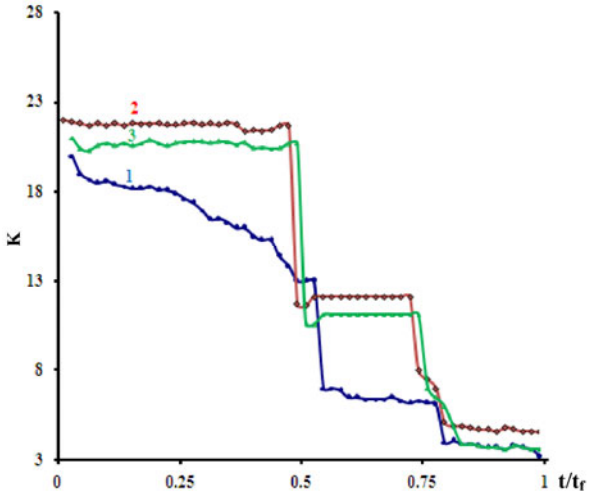


Fig. 14. Variations of parameter K at various temperatures for the dry diabase sample at 25 °C (1), 400 °C (2), and moistened sample (3).

3.4 The coalescence criterion

The coalescence criterion K kept a relative same morphology under various strain conditions as it has been obtained in Baddari *et al.* (2012). For the dry sample loaded at room temperature, the decrease of K appears much slower than in the case of moistened rocks and high temperature. The moisture and the increase in temperature generated a slip in the variations stages of K and an early transition to the irreversible deformation stage (Fig. 14, curves 2-3) which has been expanded with increasing temperature. The macrofracturing phase The macofracturing phase became faster and shorter at 0.5 [t/t_f]. We obtained practically the same variation form of K for all rock samples.

3.5 Statistical analysis of the joint physical field measurements

For convolution, several physical parameters were processed using natural orthogonal functions. In keeping with Baddari *et al.* (1999), and Sobolev (1995), the complex parameter was calculated by

$$S_k = \sum U_{ij} \varphi_{ik} \ , \tag{14}$$

where U_{ij} are the normalized values of the i-th physical parameter in the j-th point of measuring stress P, φ_{ik} are the eigenvectors of the correlation matrix of the measured physical parameters. Equation 15 is a result of convolution of a set of k physical parameters through a complex parameter S_k. Details on the chosen statistical method are explained in Baddari *et al.* (1999). Equa-

tion 14 is a result of convolution of a set of physical parameters through a complex parameter S_k.

To find out regularities in variation of complex parameters S_k, the measurements data for each of the tested samples were split into two groups. The first included points and routes measured in macrofailure areas, the second – those of the areas with no registered macrocracks. Accordingly, correlation matrixes for each group were compiled, reflecting changes in different physical parameters, and generalized values of eigenvectors were found. Six physical parameters have been considered: facet plane strain tensor ($\mathcal{J}$), longitudinal elastic wave velocity (V_P), EMR and AE intensities (N_{EMR} and N_{AE}), electrical resistivity and electrical potential changes ($\Delta\rho$ and ΔU). Computation of correlation matrixes showed that certain parameters are closely correlated regardless of the type of a sample tested. The correlation matrixes reveal invariably high correlation coefficients between the acoustic parameters N and V_P, on the one hand, and electrical $\Delta\rho_a$ and ΔU – on the other, no matter what the material. The study of six complex parameters S_k (k = 1,…, 6) computed by formula 14 with concrete values of eigenvectors φ_{ik} taken into account showed that bulk of information (95-98%) comes from the first two complex parameters – S_1 and S_2, while the equations for both parameters for the areas that were or were not brought to macrofailure differ greatly.

Despite a considerable dissimilarity in mechanical properties of samples, the complex parameters S_2 prove morphologically identical. Parameter S_2 is bay-shaped in form, which makes it more promising in terms of prognosis. Figure 15 gives changes in the complex parameter S_2 for portions brought to

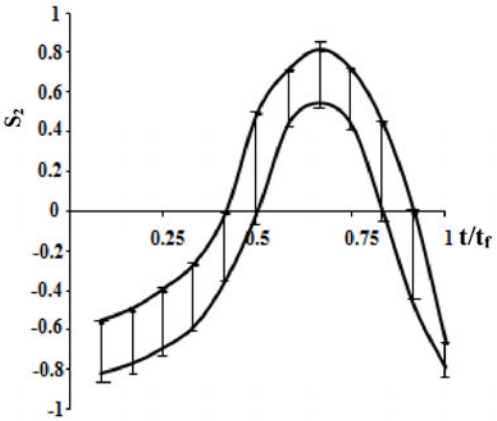


Fig. 15. Variations of the complex parameter S_2 illustrating the succession of the formation of the macrofailure in the rock sample. Same shaded is the area of variation of S2 for the studied rock samples and stress-strain conditions.

failure at different experimented strain conditions. The analysis of the complex parameter S_2 behaviour in macrofailure zone enabled to point out the following dependence of the S_2 value and its curve on strain and failure. Negative values of S_2 in the initial stage of the graph are characteristic of quasi-elastic deformations. The second stage in its behaviour is marked by a gradual rise in positive values up to maximal, which corresponded to microcracks concentration up to critical. The third, and the final stage, which manifests itself in a sharp fall of S_2 following the maximum, testifies to the process of large cracks formation along with a rise in unstable deformation in the failing portions of the samples. Identification of this stage by graph S_2 may serve as a macrofailure precursor. These three characteristics are valid for all the studied rock types. The analysis of S_2 morphology for the series of four tests showed good stability of this parameter (Fig. 15). Thus, the conditions of rock deformation affected individual physical precursors and did not qualitatively change the evolution of rock micro and macrofailure.

4. DISCUSSION OF RESULTS AND CONCLUSIONS

The use of large rock samples in our tests permits simulation of the fracturing process and utilization of a dense network for observation of the space – time variation of several physical fields. The experiments show that the reproducibility of measurements depends on the high resolution of the used laboratory equipment, the methods of measurements and the distribution of measurement network on different areas of rock sample. The errors of the acquired data for different physical fields are acceptable. The experiments were performed on solid rock samples with artificial defects acting as stress concentrators, which allowed simulation of internal shear failure. Each rock sample is characterized by a specific crystalline textural structure, with some strength distribution of grain contacts, grain interiors and intergrain zones. The process of deformation and failure of such a structure under load are accompanied by changes of these distributions caused by destruction of bonds as the weakest contacts and formation of microfailures.

After experiments using dry samples, a structural discreetness of the rock sample has been obtained. The tested rocks were divided into separate blocks by echelons of cracks of different size and orientations. Statistical processing of the geometric average sizes showed the ratio of the neighboring failures average sizes L_i/L_{i-1} is close to 2.7-3. In accordance with Sadovsky *et al*. (1991) and Sobolev (1995), the average ratio of adjacent basic dimensions of the seismogenic blocks after a seismic main shock in the geophysical environment is $L_i/L_{i-1} = 3.33$, which testifies to the self-similarity of the seismic process at different scales, and consequently the idea that structural discreetness of the earth and its discontinuity are one and the same thing. It should be noted that faultings in the earth crust after an earthquake

are also accompanied by volume units and widths of the faults may reach tens or hundreds of meters (Kuksenko *et al*. 2009).

The transition from a low level to a higher one occurs when the crucial concentration of cracks of appropriate dimensions is attained in the failure focus. This concentration depends on the sizes of forming cracks. The transition from micro to macrocracks in loaded object takes place upon reaching a certain crucial concentration of cracks of appropriate dimensions in the failure source. The critical factor, here, is the crack concentration parameter K (Baddari *et al.* 1996, Zavyalov 2006). This parameter characterizes the cracks mutual interaction of and their ability to merge via local stress field. $K = 3.8$ constitutes the threshold of the migration of the rock to the progressive macrocrack process.

The failure development process in dry samples follows three main stages: nonuniform and isolated spatial distribution of microcracks, nucleation of macrofailure source, propagation and expansion of the macrofailure. Non homogeneous distribution of the rock spatial crystalline structure subject to load leads to a mosaic distribution of EMR and AE characteristics. As a result, the crack scale effect has been observed in the EMR and AE structure behaviors. At the formation phase of isolated microcracks, the amplitude and the pulse period of EMR are much greater than those of AE. The transition of the material to the third phase of deformation (the progressive macro failure) triggers off a reverse relationship: the pulse quantity, as well as the amplitudes and periods, increases highly. The most important emissions activities have been recorded in the influence zone of the stress concentrators, where the main crack has been localized at the end of the experiment. An inequality in the intensities of AE and EMR has been noted. We note that EMR is more sensitive to the generation of tensile cracks than shear ones. Acoustic and electromagnetic quiescences have been recorded during different phases of sample deformation. This phenomenon has been recorded during different phases of rock deformation. The recorded quiescence is probably related to the scenario preceding various cracking levels transitions and consequently an avalanche in the course of expanding failure.

The evolution of the crack is controlled by the stability and interaction of ruptures in the stress strain field. This transformation is a result of reorganization in the rock of ionic and electronic physical subsystems, which must lead to characteristic changes of the mechanical and electrical properties depending on the strain-stress state of rock. The experimental results testify to important fluctuations in mechanical, and electrical parameters recorded in rock samples under various strain conditions. These fluctuations have certain common characteristics such as the transition through extremums and a developed dispersion in physical precursors during the prefailure stage. Observation and analysis of all the precursors enabled us to forecast the time of

main failure to within 60-20 min. The last precursors to be recorded were observed a few minutes before the final macrofailure in strain and electrical parameters. Basing on the character of changes of physical precursors, the rock has been divided in the zones of increased and decreased danger of the brittle failure, *i.e.*, into separate blocks, according to echelons of cracks of different sizes and directions. The analysis of the velocity variations showed, in some experiments, that if a zone is characterized by a low velocity, then a domain of high velocity will be formed around it and *vice versa*. It was found that above about one half of the failure strength the average velocity became a non-linear function of the load. In all samples, the velocity V_P reached maximum values at 0.4-0.5 [t/t_f], possibly due to the developing fracture nucleus (Fig. 7 IIIa). This maximum in V_P has been observed shortly before for the other samples. Increases in V_P have been recorded just before the main failure formation in the end of experiments. V_S marked practically the same trend as V_p but weaker (Fig. 7 IVa). Maximum values for V_S have been noted for the pyrophyllite and concrete samples at 0.4-0.5 [t/t_f]. Immediately after, V_S increased and decreased till 0.9 [t/t_f], reflecting a beginning of accumulation of shear damages at this stage. Strong increases in V_S have been recorded prior to macrofailure of diabase, pyrophyllite and concrete samples, which testifies to the heterogeneous field of strains resulting by the main failure expansion and the medium division into areas of unstable and elastic deformation. Anisotropies have been recorded in velocities in the directions x and y, parallel and perpendicular respectively to the direction of the main load P (Fig. 8). A considerable degree of elastic anisotropy is developed prior to failure. This is due to oriented microcracks, which play a major role in the failure process itself. The observed velocity anisotropy was found to be associated with changes in micro cracks density and distribution pattern. We noticed an increase in anisotropy prior to the macro failure. The maximum of this increase coincided with the dynamic propagation of the macrofailure structure. Local variations of apparent resistivity $\Delta\rho_a$ were detected. In certain zones, resistivity anisotropy was found to develop over time (Fig. 9). The anisotropy detected at the initial stage of loading is associated with the failure of the preexisting defects. As the stress increases, the portion of progressively smaller defects among the failed defects increases. The anisotropy can be qualitatively explained by the development of oriented cracks. The decrease of $\Delta\rho_a$ could be due to expansion of conducting paths system in room humidity (40-60%) experiment and crack accumulation within rock sample. The increase in $\Delta\rho_a$ is probably caused by coalescing failures, the rupture of the channels of current conduction and the formation of the macrofailure (Baddari *et al.* 1999, Ponomarev 1987, Sobolev 1995, Stopiński *et al.* 1991). It can be concluded that the variation of $\Delta\rho_a$ in the area of incipient macrofailure are upper in value and often of

the opposite polarity compared with the stable areas variations. The variations of $\Delta\rho_a$ depend on the surface of the studied area and localization of the measurement sensors in relation to the mechanical stress concentration, the orientations of cracks and the path of the macrofailure propagation. We can deduce that stressed rocks prior to macrofailure at different scales should be characterized by a heterogeneous field of strains. Polarity in $\Delta\rho_a$ changes registered in different positions, with regard to incipient failure zones, may be explained by the heterogeneities of the rock structure and the rock medium division into areas of unstable and elastic deformation. Laboratory and *in situ* investigations revealed that separate active strain regions are formed prior the macrofailure (Baddari and Frolov 2010, Sobolev 1995, Stopiński *et al.* 1991).

As the sample was loaded, the spontaneous electric field ΔU changed significantly (Fig. 7Va). We have recorded some cases of initiation and disintegration of local electric anomalies when the load varied in the range of time (0.5-0.9) t/t_f . Increases in ΔU have been recorded in all cases except in the pyrophyllite sample just before the end of the experiment. Some anomalies reached a maximum of about 30 mV at the beginning of final failure stage. Its lifetime has changed from 20 to 50 min, the reason being probably attributed to electric relaxation phenomenon. When the ΔU remained constant, we did not observe failures occurring at a higher scale. The fast transition of the failure to high scale, which led to the formation of the macrofailures and the redistribution of the stress in the rock sample, corresponded to important variation of the electric field. A nonlinear relation between the mechanical and electric fields after loading, the variations of the surface electric field significantly decreased. A number of features observed in the electric disturbances seem to be explainable qualitatively within the framework of ionic migration mechanisms.

The variations in elastic wave velocities and the electric parameters must, to a certain extent, mirror the strain-stress state of a rock and the development of anisotropy in it. The anisotropy of strength and the distributions of mechanical stresses in deformed rock massif are often a reason of local dominance of brittle destruction, resulting in catastrophical phenomena (earthquakes, rockbursts, collapses, *etc*.).

The features of the failure process refer to various conditions of rock mechanical loading. It was obtained that the negative feedback of loading of rock samples enabled to prolong the macrofailure growth stage and to study in details some regularities in main crack evolution. The macrofailure growth observed in these experiments was studied in our previous work (Baddari and Frolov 2010). It was established that moisture was accompanied by increasing the number of microcracks sets, and consequently intense AE and EMR activities in deformed rocks. The detailed observation of phys-

ical precursor variations shows that moisture increases the strain rate. The fast drop of the parameter K at 0.5 [t/t_f] (Fig. 14) testifies that the fracturing begins and develops more intensively in the rock sample under the actions of the loading and moisture. Before loading, water fills all defects (pores, capillaries, pre-existed cracks) in the rock material. Under the action of external load, during the second stage of deformation, when fracture nucleuses arise and begin to grow, the intergranular space increases causing, the diffusion of water in the rock body. We can admit that the velocity of water diffusion can be higher that the dilatancy one. The water fills the opening cracks, causes much stresses on their surfaces, induces an electrokinetic effect, affects their mechanical features, develops more cracks clusters, and causes the formation of new damages in the sample. The molecules of water play, moreover, appreciably hydrolytic role in the fracture of the interatomic bonds and changes in the physical properties in the rock material. It can be assumed that the water may have a crucial effect, rather than crack genesis and activation energy. The early transition to irreversible deformation stage and the expansion of the latter has been recorded in the first flat deformation tensor $\mathcal{J}$ and the crack concentration parameter K in the case of the moistened rocks. The moisture content of 0.3-5%, depending on nature of each rock, significantly increases the second phase (stage B of $\mathcal{J}$) of deformation with any constant deformation rate. At a lower stress level, the wet rock has a plastic deformation higher than that for a dry sample. The pore pressure can be approximately described by Terzaghi's law (Terzaghi 1925). The fluid pressure has not been controlled, and consequently the effective pressure ($P_{\text{eff}} = P - P_{\text{H2O}}$) has not been measured in our tests. According to Althaus *et al.* (1994), the resulting pressure differences can easily reach values of some hundreds of MPa and locally exceed by far the applied confining pressure. We admit that effective pressure calculated by the Terzaghi's law increases the crack rate. We admit that pore fluid pressure shift the Mohr-circle to the left and closer to the Mohr–Coulomb failure envelop. The second stage of moistened rock deformation led to decreases in wave velocities, and increases followed by decreases in the electric potential. It should be noted that there is a relative increases of ΔU at the last phase of the second strain stage for the granite and the pyrophyllite samples. When the effect of moisture has been weakened as a result of its large dispersion over the sample body, its evaporation and reduction the potential energy in the end of this stage, the rock passed to a fast unstable and progressive macrofailure at the last deformation stage.At this stage the elastic wave velocities and the electric potential tend to decrease. As a result, the water saturation plays active role in lowering the strength of the rock sample, decreases its elastic limit, Young's modulus, affects its internal structure, and we obtain finally large damage accumulation

in the condition of lower loading compared with the case of dry rocks. Similar idea is claimed by Kuksenko *et al*. (2011).

During variation of the deformation regime (Fig. 5 I-Vc), we noted shrinkage in stage *C* of the rock. We can deduce that the prediction of dynamic failure on recorded variations of physical parameters cannot always be universal. It depends on loading regime, the nature of the used sample, and its degree of anisotropy.

In case of different strain rates, it should be noted that there is a relation between the final stage of macrofailure focus development and the strain rate of a loaded rock. The stepwise failure localization did not occur during the rock deformation with a fast strain rate ($10^{-6}s^{-1}$). In studied deformed rocks, the failure growth depended on the growth rate, as the growth rate is low. It was not possible to forecast in details the onset of macrofailure and to determine the instant of time at which the transition from the first stage to the second begins. Therefore, it should be said that the precursors time prediction is inversely proportional to the strain rate. Under conditions prevailing in some seismoactive regions, the rate of the strain is in order of $10^{-14}s^{-1}$ (Baddari *et al.* 1999, Sobolev 1995). The uncertainly in forecasting the time, extrapolated from our laboratory results to the earth scale, may be in the order of few months to a few years.

The rise in temperature from 25 up to 400 °C showed a variation of physical precursors of failure, especially at the intermediate and final stages of the rock failure. The high temperature showed a noticeable effect on the period and amplitude of the impulses (Fig. 12), as well as on the intensity of the acoustic and especially electromagnetic emissions (Fig. 13), which reflected changes in the corresponding microcracking process. This was manifested in the fast decrease of the parameter *K* (Fig. 14), which is probably due to the decrease in the plasticity of the grains and the transition to a more brittle rupture of the inter-grains contacts as a result of gravitational evaporation of the rock moisture and explosion of liquid or gaseous inclusions and microfracture (Baddari and Frolov 1990, Corrêa and Nascimento 2005, Lavrov and Shkuratnik 2005, Panin 1985). The applied temperatures in these experiments are different from the temperatures of recrystallization and other second-type transition phases in the studied rocks. The variation of physical precursors is certainly related to the quantity and changes of the state of adsorbed fluids in the pores of polycrystalline field. In general, an increase of temperature has led to the predominance of the brittle fracture mechanism, which led to reduction of the plastic limit and its extension in time following the formation of spatially separate cracks. In our opinion, high applied temperatures were accompanied by an increase of the proportion of intercrystalline failure and brittle crystal boundaries microcracks. For well-studied brittle materials it was noted that subcritical crack propagation rate

was determined by the rate of chemical reaction at the crack tip (Beeler 2004, Atkinson and Meredith 1987). The final stage of macrofailure genesis rupture was accelerated while the temperatures became significant.

The analysis of a set of physical fields makes the precursor identification more unambiguous and practically dependable. The obtained results testify to the considerable fluctuations in strain, acoustic, electromagnetic, and electrical parameters recorded in natural rocks and concrete material under various strain conditions. Based on these experiments, the form of these fluctuations differs not only for physically different fields, the nonhomogeneity of the stress-strain distribution and the anisotropy of the rock medium but also for realizations of an individual parameter registered in different frequency ranges, bases, and orientations. The differences in the time changes of the physical precursors caused by micro and macrofailure allow to assume that all investigated parameters are not equivalent. Various strain conditions change the tendency of each physical parameter. The complex statistical analysis obtained after convolution of measurement results allowed to reliably forecast the different failure stages independently of the composition, nature, and experimental conditions. The increase of S_2 to a maximum, followed by a sharp decrease can be used as a predictor of macrofailure in particular parts of the rock mass. The result provided in Fig. 15 makes S_2 more promising in terms of prognosis. The strain, the acoustic, the coalescing and the complex parameters appear to be useful in using them as physical precursors of fracture in arid, geothermal, and vadoze zones. The wave velocities, the electrical, the coalescing and the complex parameters can be used in order to investigate the fracture precursors in geothermal region. The obtained results point out the efficiency of laboratory modelling of seismic process in the earth crust.

Acknowledgments. The authors would like to express their gratitude to Dr. Mathilde Adelinet and to the two anonymous referees for their helpful criticism, fruitful and valuable comments.

References

Adushkin, V.V., and S.B. Turuntaev (2005), Anthropogenic processes in the earth's crust. **In:** *Risks and Catastrophes*, INEK, Moscow.

Althaus, E., A. Friz-Töpfer, Ch. Lempp, and O. Natau (1994), Effects of water on strength and failure mode of coarse-grained granites at 300 °C, *Rock Mech. Rock Eng.* **27**, 1, 1-21, DOI: 10.1007/BF01025953.

Ammon, Ch.J., H. Kanamori, and Th. Lay (2008), A great earthquake doublet and seismic stress transfer cycle in the central Kuril islands, *Nature* **451**, 561-565, DOI: 10.1038/nature06521.

Asatryan, Kh.O., G.A. Sobolev, and V.A Mansurov (1993), Development of block hierarchy and acoustic emission in rock at all-sided compression. **In**: *Modeling of the Seismic Process and Earthquake Precursors*, Vol. 1, RAS, 17-20 (in Russian).

Atkinson, B.K., and P.G. Meridith (1987), Experimental fracture mechanics data for rocks and minerals. **In:** B.K. Atkinson (ed.), *Fracture Mechanics of Rock,* Academic Press, San Diego, CA, 477-525.

Baddari, K., and A.D. Frolov (1990), Influence of temperature on physical precursors of rock failure, *Izv. Geol. Prospect.* **9**, 102-108 (in Russian).

Baddari, K., and A.D. Frolov (2010), Regularities in discrete hierarchy seismo-acoustic mode in a geophysical field, *Ann. Geophys.-Italy* **53**, 5-6, 31-42, DOI: 10.4401/ag.4725.

Baddari, K., G.A. Sobolev, and A.D. Frolov (1996), Similarity in seismic precursors at different scales, *CR Acad. Sci. II A* **323**, 755-763.

Baddari, K., G.A. Sobolev, A.D. Frolov, and A.V. Ponomarev (1999), An integrated study of physical precursors of failure in relation to earthquake prediction, using large scale rock blocks, *Ann. Geophys.-Italy* **42**, 5, 771-787.

Baddari, K., A.D. Frolov, V. Tourtchine, and F. Rahmoune (2011), An integrated study of the dynamics of electromagnetic and acoustic regimes during failure of complex macrosystems using rock blocks, *Rock Mech. Rock Eng.* **44**, 3, 269-280, DOI: 10.1007/s00603-010-0130-5.

Baddari, K., A.D. Frolov, V. Tourtchine, S. Makdeche, and F. Rahmoune (2012), Effect of temperature on physical precursors of block rock failure, *Acta Geophys*. **60**, 4, 1007-1029, DOI: 10.2478/s11600-012-0038-4.

Bahat, D., A. Rabinovitch, and V. Frid (2005), *Tensile Fracturing in Rocks: Tectonofractographic and Electromagnetic Radiation Methods*, Springer Verlag, Berlin.

Beeler, N.M. (2004), Review of the physical basis of laboratory-derived relations for brittle failure and their implications for earthquake occurrence and earthquake nucleation, *Pure Appl. Geophys.* **161**, 9-10, 1853-1876, DOI: 10.1007/s00024-004-2536-z.

Benson, P.M., B.D. Thompson, P.G. Meredith, S. Vinciguerra, and R.P. Young (2007), Imaging slow failure in triaxially deformed Etna basalt using 3D acoustic-emission location and X-ray computed tomography, *Geophys. Res. Lett.* **34**, 3, DOI: 10.1029/2006GL028721.

Bizzarri, A., and M. Cocco (2006), A thermal pressurization model for spontaneous dynamic rupture propagation on a three-dimensional fault: 2. Traction evolution and dynamic parameters, *J. Geophys. Res.* **111**, B05304, DOI: 10.1029/2005JB003864.

Brace, W.F., and D.L. Kohlstedt (1980), Limits on lithospheric stress imposed by laboratory experiments, *J. Geophys. Res.* **85**, B11, 6248-6552, DOI: 10.1029/JB085iB11p06248.

Cai, M., and D. Liu (2009), Study of failure mechanism of rock under compressive-shear loading using real-time laser holography, *Int. J. Rock Mech. Min.* **46**, 1, 59-68, DOI: 10.1016/j.ijrmms.2008.03.010.

Chen, F., D.Y. Chen, Q.-P. Cao, S.-J. Yu, D.-J. Xu, C.-X. Chen, Y.-L. Yu, and J.-H. Sheng (1993), Study on the property of apparent resistivity changes of rock samples by in situ shear and friction test, *Acta Seismol. Sinica* **6**, 3, 721-279, DOI: 10.1007/BF02650411.

Corrêa, C.C., and R.S.V. Nascimento (2005), Study of shale-fluid interactions using thermogravimetry, *J. Therm. Anal. Calorim.* **79**, 2, 295-298, DOI: 10.1007/s10973-005-0052-8.

Darot, M., and T. Reuschlé (2000), Acoustic wave velocity and permeability evolution during pressure cycles on a thermally cracked granite, *Int. J. Rock Mech. Min.* **37**, 7, 1019-1026, DOI: 10.1016/S1365-1609(00)00034-4.

David, C., B. Menéndez, and M. Darot (1999), Influence of stress-induced and thermal cracking on physical properties and microstructure of La Peyratte granite, *Int. J. Rock Mech. Min.* **36**, 4, 433-448, DOI: 10.1016/S0148-9062(99)00010-8.

Dresen, G., S. Stanchits, and E. Rybacki (2010), Borehole breakout evolution through acoustic emission location analysis, *Int. J. Rock Mech. Min.* **47**, 3, 426-435, DOI: 10.1016/j.ijrmms.2009.12.010.

Evans, K.F. (2005), Permeability creation and damage due to massive fluid injections into granite at 3.5 km at Soultz: 2. Critical stress and fracture strength, *J. Geophys. Res*. **110**, B4, DOI: 10.1029/2004JB003169.

Fortin, J., S. Stanchits, S. Vinciguerra, and Y. Guéguen (2011), Influence of thermal and mechanical cracks on permeability and elastic wave velocities in a basalt from Mt. Etna volcano subjected to elevated pressure, *Tectonophysics* **503**, 60-74, DOI: 10.1016/j.tecto.2010.09.028.

Frid, V., A. Rabinovitch, and D. Bahat (2003), Fracture induced electromagnetic radiation, *J. Phys. D. Appl. Phys*. **36**, 13, 1620-1628, DOI: 10.1088/0022-3727/36/13/330.

Glover, P.W.J., J.B. Gomez, P.G. Meredith, S.A. Boon, P.R. Sammonds, and S.A.F. Murrell (1996), Modelling the stress-strain behaviour of saturated rocks undergoing triaxial deformation using complex electrical conductivity measurements, *Surv. Geophys.* **17**, 3, 307-330, DOI: 10.1007/BF01904046.

Gupta, H.K. (2005), Artificial water reservoir-triggered earthquake with special emphasis at Koyna, *Curr. Sci. India* **88**, 10, 1628-1631.

Heap, M.G., S. Vinciguerra, and P.G. Meredith (2009), The evolution of elastic moduli with increasing crack damage during cyclic- stressing of basalt from Mt. Etna volcano, *Tectonophysics* **471**, 1-2, 153-160, DOI: 10.1016/j.tecto.2008.10.004.

Heap, M.J., P. Baud, P.G. Meredith, S. Vinciguerra, A.F. Bell, and I.G. Main (2011), Brittle creep in basalt and its application to time-dependent volcano deformation, *Earth Planet. Sci. Lett.* **307**, 1-2, 71-82, DOI: 10.1016/j.epsl.2011.04.035.

Jouniaux, L., K. Masuda, X. Lei, O. Nishizawa, K. Kusunose, L. Liu, and W. Ma (2001), Comparison of the microfracture localization in granite between fracturation and slip of a preexisting macroscopic healed joint by acoustic emission measurements, *J. Geophys. Res.* **106**, B5, 8687-8698, DOI: 10.1029/2000JB900411.

Jouniaux, L., M. Zamora, and T. Reuschlé (2006), Electrical conductivity evolution of non-saturated carbonate rocks during deformation up to failure, *Geophys. J. Int.* **167**, 2, 1017-1026, DOI: 10.1111/j.1365-246X.2006.03136.x.

Kadomtsev, A.G., E.E. Damaskinskaya, and V.S. Kuksenko (2011), Fracture features of granite under various deformation conditions, *Phys Solid State* **53**, 9, 1876-1881, DOI: 10.1134/S1063783411090150.

Kuksenko, V.S. (2005), Diagnostic and forecasting of breakage of large-scale objects, *Phys. Solid State* **45**, 5, 812-816, DOI: 10.1134/1.1924837.

Kuksenko,V.S., Kh.F. Makhmudov, V.A. Mansurov, U. Sultonov, and M.Z. Rustamova (2009), Changes in structure of natural heterogeneous materials under deformation, *J. Min. Sci.* **45**, 4, 355-358, DOI: 10.1007/s10913-009-0044-3.

Kuksenko, V.S., E.E. Damaskinskaya, and G. Kadomtsev (2011), Fracture of granite under various strain conditions, *Izv. Phys. Solid Earth* **47**, 10, 879-885, DOI: 10.1134/S1069351311100053.

Lacidogna, G., A. Carpinteri, A. Manuello, G. Durin, A. Schiavi, G. Niccolini, and A. Agosto (2011), Acoustic and electromagnetic emissions as precursor phenomena in failure processes, *Strain* **47**, Suppl. s2, 144-152, DOI: 10.1111/j.1475-1305.2010.00750.x.

Lavrov, A.V., and V.L. Shkuratnik (2005), Deformation- and fracture-induced acoustic emission in rocks, *Acoust. Phys.* **51**, Suppl. 1, 2-11, DOI: 10.1134/1.2133948.

Lockner, D. (1993), The role of acoustic emission in the study of rock, *Int. J. Rock Mech. Min.* **30**, 7, 883-899, DOI: 10.1016/0148-9062(93)90041-B.

Lockner, D.A., and J.D. Byerlee (1986), Changes in complex resistivity during creep in granite, *Pure Appl. Geophys.* **124**, 4-5, 659-676, DOI: 10.1007/BF00879603.

Lockner, D.A., and S.A. Stanchits (2002), Undrained poroelastic response of sandstones to deviatoric stress change, *J. Geophys. Res.* **107**, B12, 2553, DOI: 10.1029/2001JB001460.

Lockner, D.A., J.D. Byerlee, V.S. Kuksenko, and A.V. Ponomarev (1986), Stick slip, charge separation and decay, *Pure Appl. Geophys.* **124**, 3, 601-608, DOI: 10.1007/BF00877218.

Lockner, D.A., J.D. Byerlee, V.S. Kuksenko, A. Ponomarev, and A. Sidorin (1991), Quasi-static fault growth and shear fracture energy in granite, *Nature* **350**, 6313, 39-42, DOI: 10.1038/350039a0.

Panin, V.E. (1985), *Structural Levels of Solid Bodies' Deformation*, Nauka, Novosibirsk (in Russian).

Ponomarev, A.V. (1987), The study of variations in the electrical state of rocks as applied to the search for earthquake precursors, Ph.D. Thesis, Moscow (in Russian).

Ponomarev, A.V., A.D. Zavyalov, V.B. Smirnov, and D.A. Lockner (1997), Physical modelling of the formation and evolution of seismically active fault zones, *Tecnophysics* **277**, 57-81, DOI: 10.1016/S0040-1951(97)00078-4.

Pozzi, J.-P., and L. Jouniaux (1994), Electrical effects of fluid circulation in sediments and seismic prediction, *CR Acad. Sci. II A* **318**, 1, 73-77.

Reuschlé, T., S.G. Haore, and M. Darot (2006), The effect of heating on the microstructural evolution of La Peyratte granite deduced from acoustic velocity measurements, *Earth Planet. Sci. Lett*. **243**, 3-4, 692-700, DOI: 10.1016/j.epsl.2006.01.038.

Rudajev, V., J. Vilhelm, J. Kozák, and T. Lokajíček (1996), Statistical precursors of instability of loading rock samples based on acoustic emission, *Int. J. Rock Mech. Min.* **33**, 7, 743-748, DOI: 10.1016/0148-9062(96)00023-X.

Sadovsky, M.A., L.G. Bolkhovitinov, and V.F. Pisarenko (1991), *Deformation of the Geophysical Medium and Seismic Process*, Nauka, Moscow (in Russian).

Schubnel, A., and Y. Guéguen (2003), Dispersion and anisotropy of elastic waves in cracked rocks, *J. Geophys. Res*. **108**, B2, 1978-2012, DOI: 10.1029/2002JB001824.

Schubnel, A., B.D. Thompson, J. Fortin, Y. Guéguen, and R.P. Young (2007), Fluid induced rupture experiment on Fontainebleau sandstone: Premonitory activity, rupture propagation, and aftershocks, *Geophys. Res. Lett.* **34**, 19, DOI: 10.1029/2007GL031076.

Shearer, P.M. (1999), *Introduction to Seismology*, Cambridge Univ. Press, Cambridge.

Simpson, D.W., W.S. Leith, and C.H. Sholz (1988), Two types of reservoir- induced seismicity, *Bull. Seismol. Soc. Am.* **78**, 6, 2025-2040.

Smirnov, V.B., and A.V. Ponomarev (2004), Seismic regime relaxation properties from in situ and laboratory data, *Izv. Phys. Solid Earth* **40**, 10, 807-816.

Smirnov, V.B., A.V. Ponomarev, and A.D. Zavyalov (1995), Acoustic structure in rock samples and the seismic process, *Izv. Phys. Solid Earth* **31**, 1, 38-58.

Smirnov, V.B., A.V. Ponomarev, P. Bernard, and A.V. Patonin (2010), Regularities in transient modes in the seismic process according to the laboratory and natural modelling, *Izv. Phys. Solid Earth* **46**, 2, 104-135, DOI: 10.1134/S1069351310020023.

Sobolev, G.A. (1995), *Fundamental of Earthquake Prediction*, ERC, Moscow.

Sobolev G.A., and A.V. Ponomarev (2003), *Physics of Earthquakes and Precursors*, Nauka, Moscow (in Russian).

Sobolev, G.A., and A.V. Ponomarev (2011), Dynamics of fluid-triggered fracturing in the modeles of a geological medium, *Izv. Phys. Solid Earth* **47**, 10, 902-918, DOI: 10.1134/S1069351311100119.

Sobolev, G.A., A.V Ponomarev, Yu.Ya. Maibuk, N.A. Zakrzhevskaya, V.I. Ponyatovskaya, D.G. Sobolev, A.A. Khromov, and Yu.V. Tsyvinskaya (2010), The dynamics of the acoustic emission with water initiation, *Izv. Phys. Solid Earth* **46**, 2, 136-153, DOI: 10.1134/S1069351310020035.

Soloviev, S.P., and A.A. Spivak (2009), Electromagnetic signals generated by the electric polarization during the constrained deformation of rocks, *Izv. Phys. Solid Earth* **45**, 4, 347-355, DOI: 10.1134/S1069351309040077.

Stanchits, S.A., D.A. Lockner, and A.V. Ponomarev (2003), Anisotropic changes in P wave velocity and attenuation during deformation and fluid infiltration of granite, *Bull. Seismol. Soc. Am.* **93**, 4, 1803-1822, DOI: 10.1785/0120020101.

Stopiński, W., A.V. Ponomaryov, and V. Los (1991), The dynamics of rupture in porous media, *Pure Appl. Geophys*. **136**, 1, 29-47, DOI: 10.1007/BF00878886.

Terzaghi, K. (1925), *Principles of Soil Mechanics: A Summary of Experimental Studies of Clay and Sand*, McGraw-Hill, New York.

Thompson, B.D., R.P. Young, and D.A. Lockner (2006), Fracture in Westerly granite under AE feedback and constant strain rate loading: Nucleation, quasi-static propagation, and the transition to unstable fracture propagation, *Pure Appl. Geophys*. **163**, 5-6, 995-1019, DOI: 10.1007/s00024-006-0054-x.

Vinciguerra, S., C. Trovato, P.G. Meredith, and P.M. Benson (2005), Relating seismic velocities, thermal cracking and permeability in Mt. Etna and Iceland basalts, *Int. J. Rock Mech. Min.* **42**, 7-8, 900-910, DOI: 10.1016/j.ijrmms.2005.05.022.

Wan, Z.J., Y.S. Zhao, Y. Zhang, and C. Wang (2009), Research status quo and prospection of mechanical characteristics of rock under high temperature and high pressure, *Procedia Earth Planet. Sci*. **1**, 1, 565-570, DOI: 10.1016/j.proeps.2009.09.090.

Xu, X.L., F. Gao, X.M. Shen, and H.P. Xie (2008), Mechanical characteristics and microcosmic mechanisms of granite under temperature loads, *J. China Univ. Min. Technol*. **18**, 3, 413-417.

Zang, A., F.C. Wagner, S. Stanchits, Ch. Janssen, and G. Dresen (2000), Fracture process zone in granite, *J. Geophys. Res*. **105**, B10, 23651-23661, DOI: 10.1029/2000JB900239.

Zavyalov, A.D. (2006), *Intermediate Term Earthquake Prediction*, Nauka, Moscow.

Zhang, L.Y., X.B. Mao, and A.H. Lu (2009), Experimental study of the mechanical properties of rocks at high temperature, *Sci. China Ser. E* **52**, 3, 641-646, DOI: 10.1007/s11431-009-0063-y.

Zhurkov, S.N. (1984), Kinetic concept of the strength of solids, *Int. J. Fract. Mech.* **26**, 4, 295-307, DOI: 10.1007/BF00962961.

Zhurkov, S.N., V.S. Kuksenko, V.A. Petrov, V.N. Savelev, and U.S. Sultanov (1980), Concentration criterion of rock volume fracture. **In**: *Physical Processes in Earthquake Sources*, Nauka, Moscow, 78-86 (in Russian).

Utsu, T. (2002), Statistical features of seismicity. **In:** W.H.K. Lee, H. Kanamori, P.C. Jennings, and C. Kisslinger (eds.), *International Handbook of Earthquake and Engineering Seismology*, Part A, Academic Press, San Diego, 719-732, DOI: 10.1016/S0074-6142(02)80246-7.

Utsu, T., Y. Ogata, and R.S. Matsu'ura (1995), The centenary of the Omori formula for decay law of afterschock activity, *J. Phys. Earth* **43**, 1-33, DOI: 10.4294/jpe1952.43.1.

Received 14 May 2013
Received in revised form 18 November 2013
Accepted 27 November 2013

DE GRUYTER OPEN

Acta Geophysica
vol. 63, no. 1, Feb. 2015, pp. 103-124
DOI: 10.2478/s11600-014-0217-6

Results of Crust and Mantle Soundings in Central and Northern Europe in the 21st Century (Review)

Vladimir Yu. SEMENOV

Institute of Geophysics, Polish Academy of Sciences, Warszawa, Poland
e-mail: sem@igf.edu.pl

A b s t r a c t

The first decade of 21st century is characterized by the appearance of new approaches to deep induction soundings. The theory of magnetovariation and magnetotelluric soundings was generalised or corrected. Spatial derivatives of response functions (induction arrows) were obtained for the ultra-long periods. New phenomena have been detected by this method: secular variations of the Earth's apparent resistivity and the rapid changes of induction arrows over the last 50 years. The first one can be correlated with the number of earthquakes, and the second one – with geomagnetic jerks in Central Europe. The extensive studies of geoelectrical structure of the crust and mantle were realized in the frame of a series of international projects. New information about geoelectrical structures of the crust in Northern Europe and Ukraine was obtained by deep electromagnetic soundings involving controlled powerful sources. An influence of the crust magnetic permeability on the deep sounding results was confirmed.

Key words: induction soundings, conductivity, crust, mantle, Europe.

1. INTRODUCTION

The end of the 20th century was characterized by an increasing interest in the deep induction soundings of the Earth mantle, particularly by the magnetovariation (MV) method (*e.g.*, Roberts 1984, 1986, Schultz and Larsen 1987, 1990, Semenov 1989, Schultz 1990, Schultz *et al.* 1993, Olsen 1992,

1998, 1999a, Semenov and Jóźwiak 1999, Schmucker 1999a, b) and using satellite data (*e.g.*, Oraevsky *et al.* 1993, Olsen 1999b). A feature of deep soundings is the lack of direct verification of the obtained results except of the ultra-deep borehole data in the Earth's crust, reaching 12 km depth. Therefore, the correctness of the induction sounding theory plays a dominant role in such investigations. For example, it is obvious that the model of the source field in the common form of a "plane wave" will not be valid for long periods as well as without sphericity of the Earth taken into account (Schultz and Zhang 1994).

In retrospect, it is surprising that some oldest works in geoelectricity were broader than the traditional "plane wave" model. The concept of impedance was introduced in the early 1930s by Leontovich in Russia. Then his student Rytov (1940) published the mathematical model in France and Russia, the first approximations of which are applied now as magnetotelluric (MT) and generalized magnetovariation (MV) soundings. Later a similar work was published by Wait (1954). Leontovich (1948) has considered limits of applicability of the Rytov model and finally Senior and Volakis (1995) found a small error in that work. The simplified model suggested by Tikhonov (1950) and Cagniard (1953) certainly had a great success for the exploration of mineral resources by the MT method. Two traditional MV methods for estimation impedances and tippers separately required a separation of the observed field to "normal" and "anomalous" parts; that is a vague procedure. Solution of the problem has been already incorporated in the Rytov model that was pointed out by Guglielmi and Gokhberg (1987), while the simplified MV method was already developed and used by Banks (1969), Berdichevsky *et al.* (1969), and Schmucker (1970). Approaches considered by Bates *et al.* (1976), Woods and Lilley (1979), Kuckes (1973), and Kuckes *et al.* (1985) were closer to Rytov's one. The corrected concept for the induction sounding was proposed by Shuman (1999). The theoretical transformations of impedance matrix (called "tensor" with overstatement because we do not know exactly what kinds of field sources formed the measured signal in each direction) to an "resistivity azimuthal tensor" was obtained theoretically by Reilly (see Weckmann *et al.* 2003) and Semenov (1988). As a result, we have polar diagrams of the apparent resistivity instead of impedances. They can be not equal, as shown below.

The deep induction soundings request impedances in a wide period range to obtain reliable results. To satisfy this request, the joint inversions of local MT impedances with regional MV responses were tested by Semenov (1988), Egbert and Booker (1992), Schultz *et al.* (1993) and Semenov and Rodkin (1996). This approach was used for the induction soundings in the frame of the following international projects: Baltic Electromagnetic Array Research (BEAR, 1998-2002), Central Europe Mantle Electrical Structure

(CEMES, 2001-2003), Electro-Magnetic soundings of Trans-European Suture Zone" (EMTESZ, 2003-2005), Electro-Magnetic Mini Arrays (EMMA, 2005-2008) in Fennoscandia, Fennoscandian Electrical conductivity from soundings with Natural and Controlled Sources (FENICS, 2007-2009), and Lithospheric Structure of TESZ by Magneto-Variation Soundings (LS-MVS, 2009-2012) including territories of Belarus, Czech Republic, Finland, Germany, Hungary, Norway, Poland, Romania, Russia, Slovakia, Sweden, and Ukraine. The national and global investigations have been already reviewed by Korja (2007) and Kuvshinov (2012).

2. THEORY AND MODELING

In the beginning of 21st century, a number of important theoretical works were published devoted to basic problems of the deep MV and MT soundings. The first problem concerned the deep MV sounding. The common model with the plane wave did not contain a vertical component of the magnetic field, which only occurred in the presence of heterogeneity in the medium. However, the long-period field of an auroral electrojet can be considered as a plane wave in the middle latitudes (Vanyan *et al.* 2002) but naturally with the vertical component. The deep soundings are also possible using the source, but there is a problem of separating the observed field into parts related and not related to the presence of inhomogeneities without knowledge about medium properties. Solution to this problem follows from several old and recent works.

Following Rytov (1940) and Wait (1954), the scalar response function $C(\mathbf{r}, \omega)$ and its gradient can be found simultaneously from the boundary condition derived by Guglielmi and Gokhberg (1987):

$$i\omega B_z = Z(\omega, \mathbf{r}) \cdot \operatorname{div} B_\tau + \operatorname{grad} Z(\omega, \mathbf{r}) \cdot \boldsymbol{B}_\tau \ , \qquad (1)$$

where B_z and $\boldsymbol{B}_\tau$ are the Fourier amplitudes of the "observed" vertical and tangential component of magnetic field, respectively; $Z(\omega, \mathbf{r}) = E_x/B_y$ is the scalar impedance, which can be recalculated in a response function $C(\omega, \mathbf{r}) = Z/i\omega$. Here E_x is the electric field component orthogonal to B_y, ω is the angular frequency, i is the imaginary unit, and $\mathbf{r}$ is the position vector. Relation 1 without the term $\operatorname{grad} Z \bullet \boldsymbol{B}_\tau$ is known as the horizontal spatial gradient (HSG) method and used for soundings of the laterally homogeneous media (Logvinov 2002). Its variant for the linearly polarized field generated by the magnetospheric ring currents was named the geomagnetic depth sounding (GDS) method (Olsen 1992) and used in the spherical geomagnetic reference system (r, φ, θ) for soundings of the mid- and lower mantle: $B_\mathbf{r} = C(\omega, \mathbf{r}) \cdot 2 \cdot B_\theta / R \cdot \operatorname{tg} \theta$. Here R is the Earth's radius; θ is the co-latitude. Note that the response function $C(\omega, \mathbf{r})$ in the GDS method has a direction

coinciding with an averaged direction of the induced currents, *i.e.*, along the geomagnetic longitudes and thus could be considered as a vectoral value. A new interesting transfer function $D(\omega) \sim B_\varphi(\omega)/B_\theta(\omega)$ for the GDS method has been proposed by Fujii and Schultz (2002) along with the conventional response $C(\omega) \sim B_r(\omega)/B_\theta(\omega)$. The physical sense of the response $D(\omega)$ is fixing the observed deviations of horizontal field from the geomagnetic coordinate system. The function $D(\omega) \neq 0$ can be caused by some inhomogeneity in the conductive Earth or changing of conductivity due to changes of a relative position of ring currents and the Earth, during the magnetic storms for example.

Following Schmucker (2003), the magnetic field $\mathbf{B}(t)$ may be separated into "normal" ($_\mathrm{n}$) and "anomalous" ($_\mathrm{a}$) parts. The "normal" part can be applied for soundings of a laterally homogeneous medium while the "anomalous" part is used to fix anomalous zones in media characterized by the complex tippers $a(\omega, \mathbf{r})$ and $b(\omega, \mathbf{r})$ in the empirical relation $B_{z\mathrm{a}} = a \cdot B_x + b \cdot B_y$. Because the "observed field" $B_z = B_{z\mathrm{n}} + B_{z\mathrm{a}}$, this and HSG equations were combined by Schmucker into the single one:

$$B_z = C\left(\omega, \mathbf{r}\right) \cdot \mathrm{div}\, \boldsymbol{B}_{\mathrm{n}\tau} + c \cdot B_x + d \cdot B_y \ . \tag{2}$$

The response function was considered here as a tensor. Equation 2 is similar to Eq. 1 if $\mathrm{div}\, \boldsymbol{B}_\tau \equiv \mathrm{div}\, \boldsymbol{B}_{n\tau}$. Besides, the tippers $a(\omega, \mathbf{r})$, $b(\omega, \mathbf{r})$, and the functions $c(\omega, \mathbf{r}) = \partial C/\partial x$, $d(\omega, \mathbf{r}) = \partial C/\partial y$ are essentially different and accordingly the conventional induction arrows $\boldsymbol{C}_u$, $\boldsymbol{C}_v$ differ from the vectors $\boldsymbol{S}_u$ and $\boldsymbol{S}_v$, based on c and d values. Relation 2 has been applied to the mantle soundings by Schmucker (2003).

The problem has been considered by Shuman and Kulik (2002), and Shuman (2007). New precise boundary conditions were suggested:

$$B_z = C\,\mathrm{div}\, \boldsymbol{B}_\tau + \mathrm{grad}\, C \cdot \boldsymbol{B}_\tau + C_0^* \,\mathrm{div}\left(\boldsymbol{B}_\tau^* \times \mathbf{n}\right) + \mathrm{grad}\, C_0^* \left(\boldsymbol{B}_\tau^* \times \mathbf{n}\right) . \tag{3}$$

Here the asterisk (*) means the complex conjugate value; C and C_0 are two different functions, $\mathbf{n}$ is the unit vector normal to the surface. The question may arise: why Eq. 3 is so long? Equations 1-3 are the impedance boundary conditions on the Earth's surface (Senior and Volakis 1995). The boundary condition 3 is based on a more rigorous mathematical approach determining new impedances from the "vectoral" impedance boundary condition (VIBC) introduced by Aboul-Atta and Boerner (1975). They used the theorem that the electromagnetic fields inside a medium can be found uniquely if the tangential fields are known on its "full closed surface" (Harrington 1961). The VIBC are generalizing the complex Fourier spectra $\boldsymbol{E}(\omega)$ and $\boldsymbol{B}(\omega)$ – actually included in the boundary conditions. In other words, two projections of the complex vector $\boldsymbol{E}_\tau(\omega)$ on the complex vector $\boldsymbol{B}_\tau(\omega)$ require two orthogo-

nal bases in a 2D unitary vector space and consequently two impedances must be considered simultaneously (Shuman 2003).

Omitting the part with response C_0 in the Eq. 3 and rewriting the expression 2 in the corresponding form with gradient of $C(\omega, \mathbf{r})$ there can be noticed a similarity of all three approaches (1, 2, and 3). The neglect of C_0 is caused by the necessity of determining six unknown transfer functions from the single equation in practice. It is seen that the physical sense of tippers and induction arrows is the spatial derivatives of the response function along the horizontal axis. By the way, the approaches 2 and 3 were presented by Prof. Schmucker and Prof. Shuman separately on the same day at the international conference in Poland (in 2001). Audience, including the author of this work, did not recognize at first the similarity of these reports.

Impedances depend in general on the particular sounding methods, on the properties of the conductive medium including their magnetic permeability, as well as on the adopted model of the space in the theory, the latter aspect being important while considering the way in which the induced currents are locking, even at infinity.

Moreover, the impedances depend on the exciting field that originates from sources of several different types in the period range of hours to 11 years used for the mantle electromagnetic soundings. Thus, quoting Guglielmi and Gokhberg (1987) we can say that "the Earth does not have its own impedance".

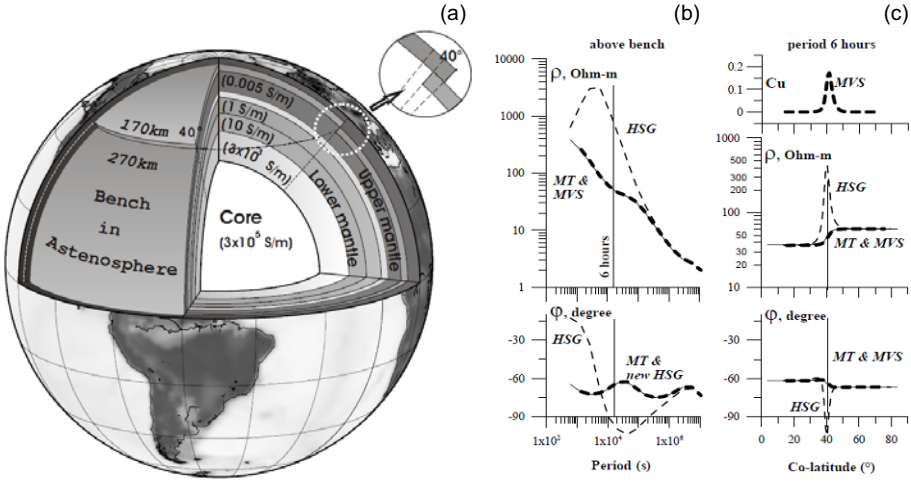


Fig. 1: (a) The model of the Earth's interiors with the surface shell conductance; (b) the apparent resistivity and impedance phase spectra obtained by MT – the thin line, HSG (GDS) – the thin dash line, and MVS (GMV) – the thick dash line soundings just above the uplift of the mantle conductive layer; (c) the profile results along co-latitudes with real induction arrow (Cu) (after Vozar and Semenov 2010).

The question arises: how great is the difference between the results obtained by different induction methods over heterogeneous media? The spherical modeling (Kuvshinov *et al.* 2005) was applied (Vozar and Semenov 2010) along the profile crossing the essential bench at co-latitude 40° whose top changed from 170 to 270 km (Fig. 1a). The modeling of different types of responses above 2D and 3D mantle inhomogeneities has shown that the various methods can give mutually inconsistent results in the apparent resistivities. The results can reach an order for apparent resistivity modules exactly above the center of anomaly while phase difference can reach tenths of degrees (Fig. 1c). Besides, induction vectors are depending on spherical latitudes. However, the MT and GMV sounding methods have shown the identical apparent resistivities above the inhomogeneity!

3. METHODS AND SOURCES

Let us note that the magnetotelluric impedance $Z(\omega)$ for a fixed ω is a ratio $\boldsymbol{E}(\omega)/\boldsymbol{B}(\omega)$ because the measured magnetic field is the magnetic induction $\mathbf{B}(t)$, but not the field intensity $\mathbf{H}(t)$, as evidenced by its unit – nanotesla (equal to 1 γ – off-system unit). Exactly this unit had been required by the IAGA in the resolution no. 3 of "Transactions…" (1973). Note that the impedance at a fixed period is a "functional" of the conductivity (Berdichevsky and Zhdanov 1981). It means that many distributions of conductivity can correspond to the impedance value found for a fixed ω. This fact does not contradict the theorem of uniqueness of the inverse solution for induction soundings proved for "infinite input data" (Rokitjansky 1982).

Estimations of impedances in practice are based on the random process theory assuming impedances as transfer functions between spectra of observed field components. These transfer functions are usually considered as scalar or matrix $[2 \times 2]$ values, sometimes even $[3 \times 3]$ (Dmitriev and Berdichevsky 2002). The last generalized heuristic approach leads to a similar result as obtained by Becken and Pedersen (2003). Such an approach is mixing both modes for the corresponding impedances that complicate analysis of data over inhomogeneous media (Semenov and Shuman 2010). Separating of the modes for 3D case has been discussed by Becken *et al.* (2008). Note that the impedances can be found in the time domain too (Nowożyński 2004).

The publication by Weckmann *et al.* (2003) brings us back to the problem of transforming the MT impedance matrix to apparent resistivity tensor elements ρ_{ij}^{*} for a laterally anisotropic medium. Widespread approach is to consider the impedance as a tensor, two and sometimes four elements of which are recalculated into the same quantity of the scalar apparent resistivities even keeping indexes of the impedance tensor. Result of such a procedure can look very strange. It is more natural to consider "resistivity"

of media as an "azimuthal tensor" and impedances as a matrix (for MT), vector (for GDS) or scalar (for GMV). The theoretical evidence transforming an impedance matrix to the resistivity tensor made by both Reilly (see Weckmann *et al.* 2003) and Semenov (2000) are identical:

$$\rho^*_{xx}(\omega) = \left(Z^2_{xy} - Z_{xx}\, Z_{yy}\right)\mu/\omega\,; \quad \rho^*_{xy}(\omega) = Z_{xx}\left(Z_{yx} - Z_{xy}\right)\mu/\omega\,;$$
$$\rho^*_{yy}(\omega) = \left(Z^2_{yx} - Z_{xx}\, Z_{yy}\right)\mu/\omega\,; \quad \rho^*_{yx}(\omega) = Z_{yy}\left(Z_{xy} - Z_{yx}\right)\mu/\omega\,. \tag{4}$$

Of course, reconsidering long-standing postulates is a thankless job. But frequent recording of $\rho^*_{xy} \sim Z^2_{xy}$ is without sense: asterisk *xy* marks a minor element of resistivity tensor while it marked major element of impedance. The relation following from the theory is $\rho^*_{xx} \sim Z^2_{xy} - Z_{xx}\, Z_{yy}$. To obtain a scalar resistivity $\rho^*_{xx} \sim Z^2_{xy}$ or $\rho^*_{yy} \sim Z^2_{yx}$ the value of $Z_{xx}\, Z_{yy}$ must be minimized. Such directions were named the preferential ones; they may be not orthogonal (Fig. 2). The apparent resistivity modules shown here are only shifted at both preferential directions and their phases are similar (Fig. 2b), while they are different for the orthogonal principal directions. Moreover, the spectra of apparent resistivities obtained from the impedance matrix by common and theoretical transformation 4 can be essentially different from their 1D inversion models. A careful study of the anisotropic media with the arbitrarily directed tensor of conductivity has been made by Pek (2002) and Pek and Santos (2002).

The problem of deep soundings arises: how an obtained local MT resistivity tensor can be combined with the regional GMV or *Sq* scalar resistivity or with the GDS continental vectoral one? In order to combine

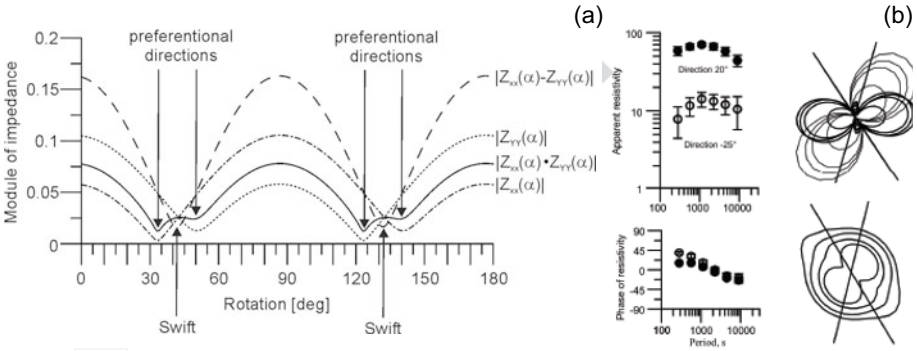


Fig. 2: (a) Choice of two preferential directions (min $|Z_{xx}\, Z_{yy}|$) in comparison with the orthogonal principal directions chosen by the Swift's procedure (min $|Z_{xx} - Z_{yy}|$); (b) example of apparent resistivities for the observatory HRB (after Semenov *et al.* 2008).

them, we have to assume that a medium is rather homogeneous at great depths. But in fact the currents of different sources are induced and locked in different ways inside the Earth and may contain different information about its structure. This difference is clearly visible for the spherical model in Fig. 1 due to the different sources and methods including the *Sq* variations (Semenov *et al*. 2013). This problem still requires further efforts to study the deep irregularities in the mantle.

The problem of a lack of coincidence between inversion models of different teams in the international projects obtained from the same initial apparent resistivities is arising too. If it concerns 1D modeling, their results can be transformed to the monotonically increasing conductance with depth. This approach gives possibility to investigate large areas presenting results as schemes of conductance at a fixed depth or depths to a fixed conductance even without subsurface parts for comparison (Semenov and Jóźwiak 2006).

Besides, it was established that secular variations of the Earth apparent resistivity estimated by two GMV and GDS methods can reach 20% of the mean measured values (Fig. 3). These variations are well correlated with the number of earthquakes in the seismically active areas of Central Europe. The depths of the earthquakes are less than 40 km where sources of such variations are situated. So the internal as well as changing external sources can essentially disturb the sounding results during long observations. Besides, the induction arrows can change their directions and their values during a couple of years. This phenomenon coincides with the appearance of geomagnetic jerks in Central Europe (Petrishchev and Semenov 2013).

Another registered phenomenon is connected with the high magnetic permeability near the Curie point in the crust (Kiss *et al.* 2005). The influence of this effect on the deep soundings was considered by Szarka *et al.* (2007).

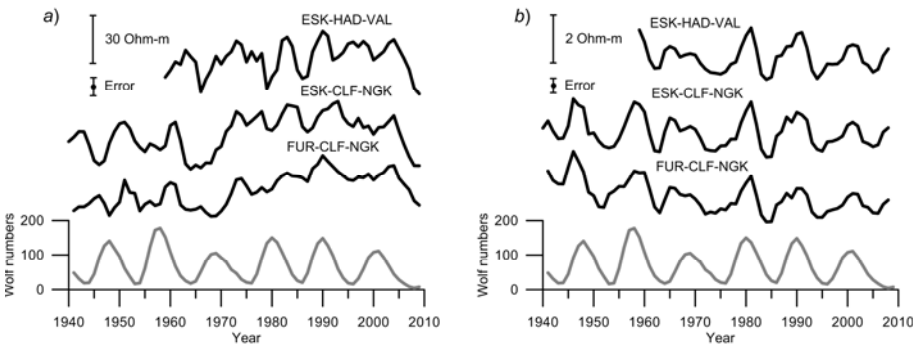


Fig. 3. Comparisons of the Earth's apparent resistivities with $T = 8.8$ hours (a), and 30 days (b) observed by different groups of observatories marked by their codes (after Petrishchev and Semenov 2013).

4. RESULTS OF DEEP SOUNDINGS

The final results of electrical conductivity studies of the Earth's crust at the Ukrainian and Fennoscandian shields were published by Ingerov *et al.* (1999) and Korja *et al.* (2002). Both studies show considerable heterogeneity in the crust conductance in these regions, reaching few orders of magnitude. The example of the crust conductance in the Ukraine is presented in Fig. 4.

The extensive deep geoelectric studies in Northern Europe were realized in the frame of two international projects: BEAR (1998-2002) with its deep sounding continuation – EMMA (2005-2008) in Fennoscandia. These measurements have been done using *c.a.* 50 MT stations, displayed permanently over the 150 × 150 km network on the territory of Fennoscandian shield. Four countries have taken part in the investigations. The obtained crustal conductance up to 60 km depth for Northern Europe is shown in Fig. 4b. The fixed conductance contrast reaches six orders and shows extremely high electrical heterogeneity of the crust in the Fennoscandian shield (Korja *et al.* 2002). The large regions of high resistance are surrounded by relatively narrow, highly conducting zones with conductances reaching dozens of kS. The origin of crustal anomalies can be connected with electronically conducting sulfide and carbon bearing structures (Zhamaletdinov 1996). The upper mantle conductance was estimated from the BEAR data during the EMMA project. It was shown that at depths of 150-300 km the conductance reaches 4-5 kS that may be expressed as "an asthenosphere conducting layer" (Sokolova and Varentsov 2007).

The second project was CEMES (2001-2003) at the territories of seven countries. The long-time MT measurements have been carried out at eleven

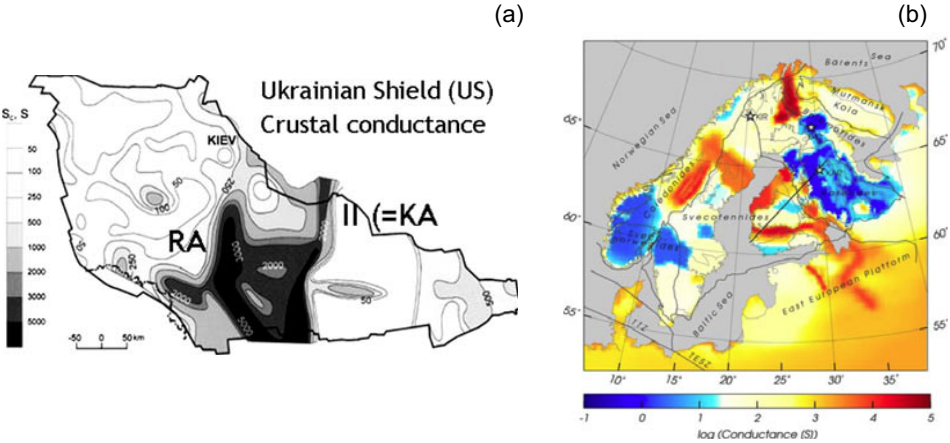


Fig. 4. Schemes of the crust conductances on the Ukrainian (a) and Fennoscandian (b) shields (after Ingerov *et al*. 2005 and Korja *et al.* 2002, respectively.

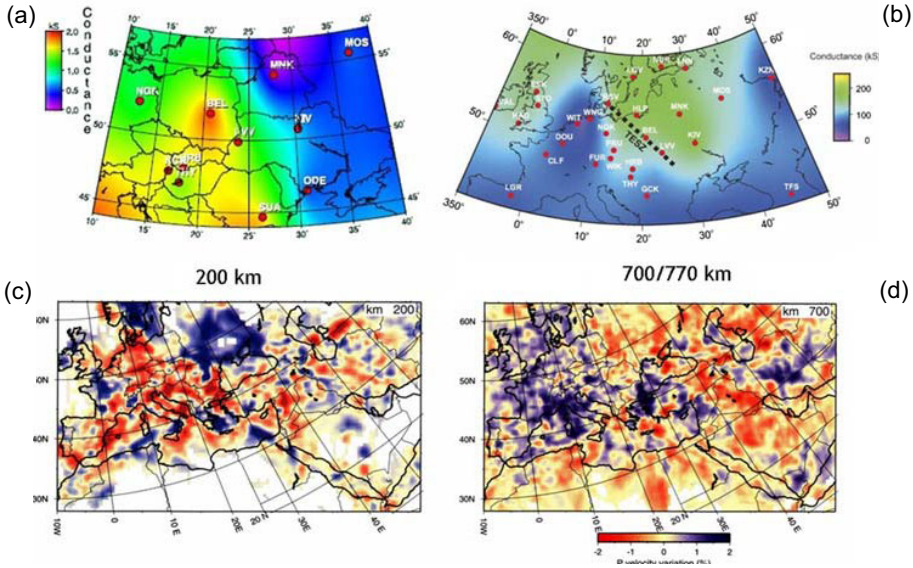


Fig. 5. Comparison of the conductance distributions in kS: depths of 50-200 km (a) and depths of 50-770 km (b) with seismic *P* velocity variations (%) at 200 km (c) and 700 km (d) (after Korja 2007).

geomagnetic observatories of Central Europe and their sounding results were combined with the MV soundings obtained at the same observatories using the historical hourly data (Fig. 5a).

Besides, the MV impedances for the periods 4 hours – 11 years estimated by six authors for 35 European observatories were collected. The precise selection of their results and subsequent combination allowed applying 1D inversion modeling to estimate the regional mantle conductance at a depth of 770 km beneath chosen observatories. The scheme of interpolation is shown in Fig. 5b.

These results have shown that the Trans-European Suture Zone (TESZ) coincides with the depth gradients of the 1 kS conductance in the upper mantle. The same effect has been observed along TESZ for the conductance gradient at a depth of 770 km in the mid-mantle. However, the conductance is increasing to the west in the upper mantle while in the mid-mantle it is increasing to the east. This reversing occurs by seismic data at about 600-800 km and has been confirmed by deep sounding results obtained along profile Germany–Belarus (Fig. 6).

The international project EMTESZ (2003-2005) was carried out on the Polish and German territories. Two long magnetotelluric profiles along the seismic ones LT-7 (Guterch *et al.* 1994) and P2 (Janik *et al.* 2002) were

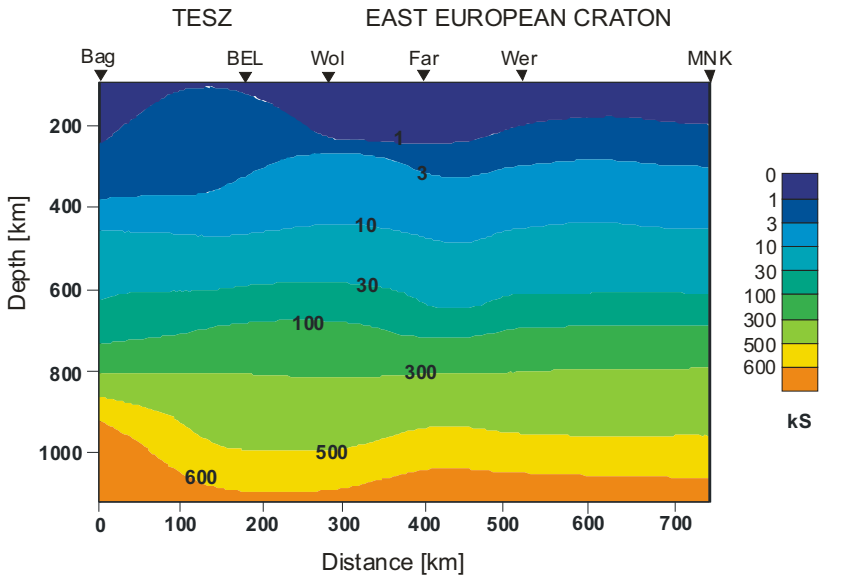


Fig. 6. Smoothed conductance deep structure along profile between geomagnetic observatories NGK (Germany) and MNK (Belarus) (after Semenov and Jóźwiak 2005).

studied by the MT method in the wide period range (0.1-10^4 s). The resistivity cross-sections up to upper mantle depths were published by Ernst *et al.* (2008).

Depths of the upper mantle were investigated separately combining three MT soundings in the TESZ-center of the profiles with apparent resistivities obtained at the nearest geomagnetic observatory Belsk situated in the TESZ too. The observed sounding result is a rarity: two curves of apparent resistivity in orthogonal directions are reaching each other at the period of one day (Fig. 7b). The obtained effect can be explained by the laterally anisotropic layer at depths of 30-40 km.

The next project was the FENICS (2007-2009). Two mutually orthogonal industrial power transmition lines of 110 and 120 km lengths (Fig. 8) with the generator of 200 kW were used to estimate of the transversal resistance T [$\Omega \cdot m^2$] of the lithosphere in the frequency range 0.1-200 Hz (Zhamaletdinov *et al.* 2011). Spacing between transmitters and receiving points reached 700 km. Besides, the soundings with the Magneto-Hydro-Dynamic (MHD) generator "Khibiny" of 80 MW power were carried out for the investigations (Zhamaletdinov 2005). The Moho depth gradient of the deepest anomalous zone in Europe (Grad *et al.* 2007) corresponds to the gradient of the transversal resistance.

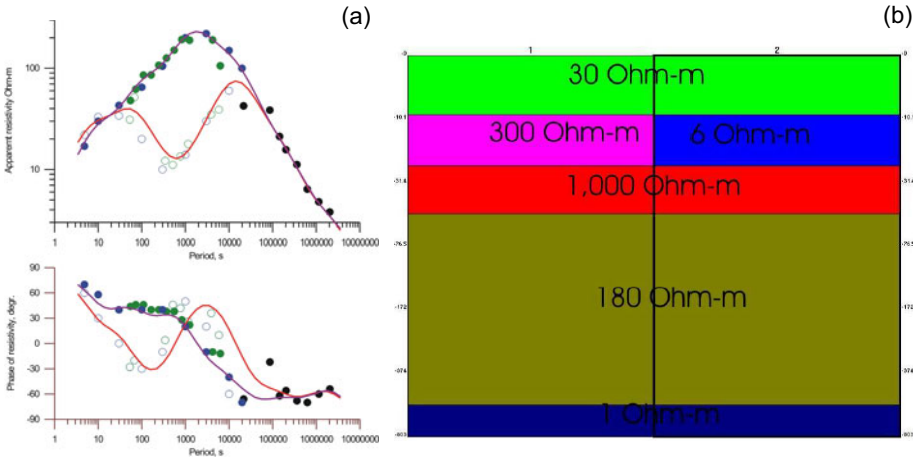


Fig. 7. The combined MT and MV apparent resistivities at two preferential directions across (black) and alongside (red) in the center of EMTESZ profiles (a) and their both 1D inversion models (b) (after Semenov *et al.* 2005).

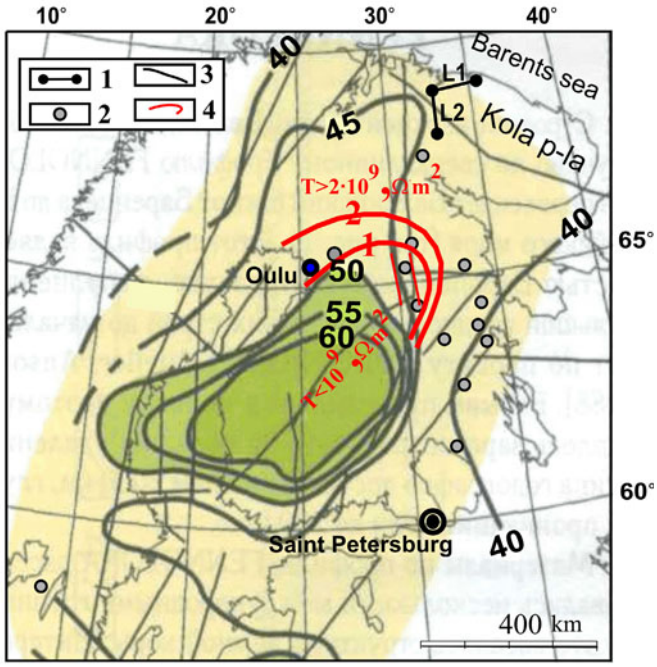


Fig. 8. The gradient of transversal resistance T [$\Omega \cdot m^2$] of the crust at the depth interval 10-60 km (red lines) coinciding with boundary of the deepest Moho zone in Europe (black lines) (after Zhamaletdinov 2011, Zhamaletdinov *et al.* 2011).

The fifth international project was the LS-MVS (2009-2012). Five countries took part in these investigations. New MVS method was tested successfully in Central Europe using the data of geomagnetic observatories (Semenov *et al.* 2011). Thus, distributions of the induction arrows as well as response functions were estimated in Poland for the period range from 3 hours to one day. This new information is analyzed now.

The original method to analyze the spatial distribution of induction arrows has been proposed recently by Jóźwiak (2012). Known arrows were recalculated to the scalar tippers and then into a ratio of the horizontal field at different points relative to a point at infinity. The Hilbert's transformation was used for this aim. These red zones are characterized by high conductance (Fig. 9). The obtained results are in good agreement with the geological knowledge.

Recently, sounding results using data of the global geomagnetic observatory net were published by Praus *et al.* (2011). A review of electromagnetic study of lithospheric structure around the TESZ was made by Jóźwiak (2013).

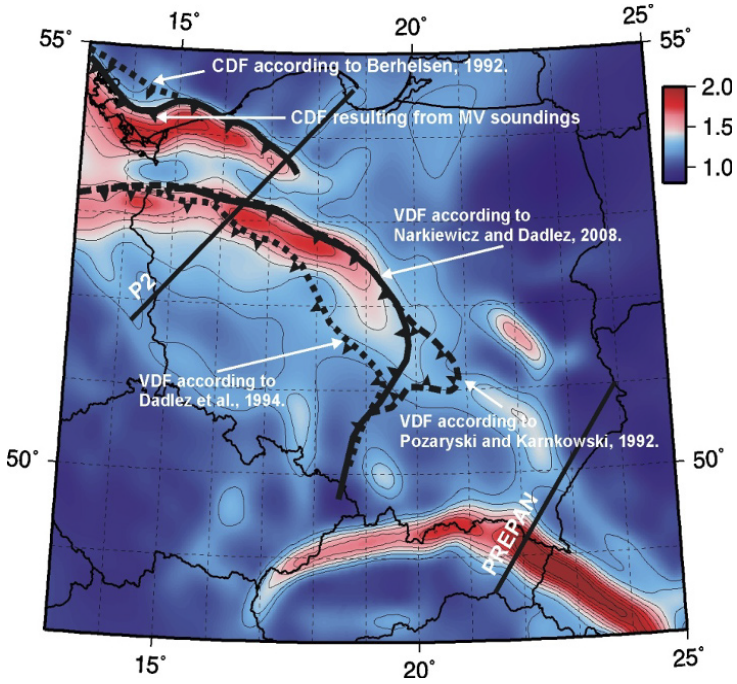


Fig. 9. Scheme of the spatial distribution of the conductive structures (red) basing on the hypothetical locations of the Caledonian and Variscan deformation fronts (right, scale in log of Siemens) (after Jóźwiak 2012).

5. CONCLUSIONS AND DISCUSSION

The development of generalized magnetovariation sounding theory significantly expanded possibilities of induction soundings in the period range from hours to a couple of days. The regional investigations by this method allow estimating additionally gradients of response functions including information about electrical inhomogeneities in the upper mantle (up to about 500 km) that was not made earlier. The induction sounding results are not stable: their long period variations can be caused by internal sources while variations with shorter periods by an abrupt change of the external field source. It is a reason why the results of formal inversions may be changeable in time. Besides, the forward spherical modeling has shown that sounding results using different sources and sounding methods above deep inhomogeneities can be essentially different. However, the sounding results of GMV and MT methods remain close even over significant irregularities in the Earth.

The methodic peculiarities of the mantle soundings are connected with combinations of two induction methods with tensor (MT) and scalar (GMV) or vectoral (GDS) apparent resistivities. The choice of destinations in the MT soundings to compare with the MV ones is promising if the resistivity of medium assumed to be a tensor rather than the impedance. It was shown that the conductance is most reliable for comparison between sounding results obtained by different investigators. An attractive but expensive method is that of controlled sources (like MGD generator) used in Russia for the deep soundings practically on a non-conductive surface. A lot of problems connected with natural sources disappear in such soundings, but requirements concerning the theory of the methods are not reduced.

Four of the five international projects in Central and Northern Europe performed areal studies. The results of the BEAR experiment increased the knowledge about the electrical structure of the Earth's crust at the Fennoscandian Shield (Korja *et al.* 2002). But another goal of the experiment – the search for a possible existence of the asthenosphere in the upper mantle – did not get a clear answer. Moreover, different research groups have come to some conflicting conclusions from the same experimental data. According to the St. Petersburg group an intermediate conductive layer associated with partial melting of rocks is clearly recorded in the depth range of 200-400 km and its average value of the longitudinal conductance reaches 8 kS (Vardanyants and Kovtun 2009). This conclusion is not contrary to the interpretation of the seismic data which has a slightly different depth about 100-150 km (Abramovitz *et al.* 2002). On the other hand, the interpretations made by other creative teams on the basis of the same experimental data assume absence of any asthenosphere under the Fennoscandian shield

or it is extremely weak on the background of the confidence limits (Varentsov *et al*. 2002, Sokolova and Varentsov 2007). The last interpretation of the BEAR data coincides with interpretation made by Vanyan *et al*. (2002). Discrepancies between the interpretations may be explained by sharp electrical inhomogeneity of the crust at that resistive region and high latitudes with complicated source field. It would be interesting to take into account the deep borehole in Karelia, where extremely solid rocks were met at the depth about 12 km.

The two main results of the CEMES project have fixed gradients of the total conductance in the mantle at 300 and 770 km depths, coinciding with the TESZ. Now physical explanation of this effect is absent. It could be attributed to the methodological inaccuracies or errors of observations, but a similar inverse pattern was observed on the results of seismic tomography (Piromallo and Morelli 2003). Reverses of the overabundant mass densities were also established at about 700 km depth by a rigorous analysis of the satellite data (Martinec and Pěč 1990). These phenomena require additional investigations as well as influence of the Earth magnetic permeability on the MV soundings and analysis of induction arrows at the period range of 3-30 hours.

The FENICS project has allowed estimating the position of the transversal resistance gradients changing twice at the boundary of the deepest (about 50-60 km) area of the Moho in Northern Europe. These northern boundaries (gradient zones) of electrical and gravity anomalies coincide. Besides, the 1D inversions show the lower crust resistivity (about 10^5 Ω·m), which is at least two orders more than for other shields in Europe and Canada. This high resistivity is coinciding well with the laboratory investigations for depths of 10-50 km (Zhamaletdinov *et al.* 2011) but is not consistent with previous studies (Vanyan *et al.* 2001) and other inversion results (Sokolova and Varentsov 2007).

Acknowledgments. The research was partially supported within statutory activities No. 3841/E-41/S/2014 of the Ministry of Science and Higher Education of Poland and by National Science Centre (NCN), Poland, grant No. 2011/01/B/ST1O/07046. Author is thankful to prof. A.A. Zhamaletdinov for discussions.

References

Aboul-Atta, O.A., and W.M. Boerner (1975), Vectoral impedance identity for the natural dependence of harmonic fields on closed boundaries, *Canadian J. Phys*. **53**, 15, 1404-1407, DOI: 10.1139/p75-179.

Abramovitz, T., H. Thybo, and E. Perchuć (2002), Tomographic inversion of seismic P- and S-wave velocities from the Baltic Shield based on FENNO-LORA data, *Tectonophysics* **358**, 1-4, 151-174, DOI: 10.1016/S0040-1951 (02)00422-5.

Banks, R.J. (1969), Geomagnetic variations and the electrical conductivity of the upper mantle, *Geophys. J. Roy. Astr. Soc.* **17**, 5, 457-487, DOI: 10.1111/ j.1365-246X.1969.tb00252.x.

Bates, R.H.T., W.M. Boerner, and G.R. Dunlop (1976), An extended Rytov approximation and its significance for remote sensing and inverse scattering, *Opt. Commun.* **18**, 4, 421-423, DOI: 10.1016/0030-4018(76)90285-6.

Becken, M., and L.B. Pedersen (2003), Transformation of VLF anomaly maps into apparent resistivity and phase, *Geophysics* **68**, 2, 497-505, DOI: 10.1190/ 1.1567217.

Becken, M., O. Ritter, and H. Burkhardt (2008), Mode separation of magnetotelluric responses in three-dimensional environments, *Geophys. J. Int.* **172**, 1, 67-86, DOI: 10.1111/j.1365-246X.2007.03612.x.

Berdichevsky, M.N., and M.S. Zhdanov (1981), *Interpretation of Anomalies Alternating Electromagnetic Field of the Earth*, Nedra, Moscow (in Russian).

Berdichevsky, M.N., L.L. Vanyan, and E.B. Fainberg (1969), The frequency sounding of the Earth using spherical analysis results of geomagnetic variations, *Geomagn. Aeron.* **9**, 372-374 (in Russian).

Cagniard, L. (1953), Basic theory of the magneto-telluric method of geophysical prospecting, *Geophysics* **18**, 3, 605-635, DOI: 10.1190/1.1437915.

Dmitriev, V.I., and M.N. Berdichevsky (2002), A generalized model of impedance, *Izv. Phys. Solid Earth* **38**, 10, 897-903.

Egbert, G.D., and J.R. Booker (1992), Very long period magnetotellurics at Tucson observatory: Implications for mantle conductivity, *J. Geophys. Res*. **97**, B11, 15099-15112, DOI: 10.1029/92JB01251.

Ernst, T., H. Brasse, V. Cherv, N. Hoffmann, J. Jankowski, W. Jóźwiak, A. Kreutzmann, A. Neska, N. Palshin, L.B. Pedersen, M. Smirnov, E. Sokolova, and I.M. Varentsov (2008), Electromagnetic images of the deep structure of the Trans-European Suture Zone beneath Polish Pomerania, *Geophys. Res. Lett.* **35**, 15, L15307, DOI: 10.1029/2008GL034610.

Fujii, I., and A. Schultz (2002), The 3D electromagnetic response of the Earth to ring current and auroral oval excitation, *Geophys. J. Int*. **151**, 3, 689-709, DOI: 10.1046/j.1365-246X.2002.01775.x.

Grad, M., T. Tiira, and ESC Working Group (2007), The Moho depth of the European plate, European Seismological Commission, Warsaw–Helsinki, http://www.seismo.helsinki.fi/mohomap/; http://www.igf.fuw.edu.pl/mohomap/.

Guglielmi, A.V., and M.B. Gokhberg (1987), On the magnetotelluric sounding in the seismically active areas, *Izv. Phys. Solid Earth* **33**, 11, 122-123 (in Russian).

Guterch, A., M. Grad, T. Janik, R. Materzok, U. Luosto, J. Yliniemi, E. Lück, A. Schulze, and K. Förste (1994), Crustal structure of the transition zone between Precambrian and Variscan Europe from new seismic data along LT-7 profile (NW Poland and eastern Germany), *C. R. Acad. Sci. II Paris* **319**, 12, 1489-1496.

Harrington, R.F. (1961), *Time-Harmonic Electromagnetic Fields*, McGraw-Hill, New York, 480 pp.

Ingerov, A.I., I.I. Rokityansky, and V.I. Tregubenko (1999), Forty years of MTS studies in the Ukraine, *Earth Planets Space* **51**, 10, 1127-1133.

Janik, T., J. Yliniemi, M. Grad, H. Thybo, T. Tiira, and POLONAISE P2 Working Group 1 (2002), Crustal structure across the TESZ along POLONAISE'97 seismic profile P2 in NW Poland, *Tectonophysics* **360**, 1-4, 129-152, DOI: 10.1016/S0040-1951(02)00353-0.

Jóźwiak, W. (2012), Large-scale crustal conductivity pattern in Central Europe and its correlation to deep tectonic structures, *Pure Appl. Geophys.* **169**, 10, 1737-1747, DOI: 10.1007/s00024-011-0435-7.

Jóźwiak, W. (2013), Electromagnetic study of lithospheric structure in the marginal zone of East European Craton in NW Poland, *Acta Geophys.* **61**, 5, 1101-1129, DOI: 10.2478/s11600-013-0127-z.

Kiss, J., L. Szarka, and E. Prácser (2005), Second-order magnetic phase transition in the Earth, *Geophys. Res. Lett.* **32**, 24, L24310, DOI: 10.1029/2005GL024199.

Korja, T. (2007), How is the European lithosphere imaged by magnetotellurics? *Surv. Geophys.* **28**, 2-3, 239-272, DOI: 10.1007/s10712-007-9024-9.

Korja, T., M. Engels, A.A. Zhamaletdinov, A.A. Kovtun, N.A. Palshin, M.Yu. Smirnov, A.D. Tokarev, V.E. Asming, L.L. Vanyan, I.L. Vardaniants, and the BEAR Working Group (2002), Crustal conductivity in Fennoscandia – a compilation of a database on crustal conductance in the Fennoscandian Shield, *Earth Planets Space* **54**, 5, 535-558.

Kuckes, A.F. (1973), Relations between electrical conductivity of a mantle and fluctuating magnetic fields, *Geophys. J. Roy. Astr. Soc.* **32**, 1, 119-130, DOI: 10.1111/j.1365-246X.1973.tb06523.x.

Kuckes, A.F., A.G. Nekut, and B.G. Thompson (1985), A geomagnetic scattering theory for evaluation of the Earth structure, *Geophys. J. Roy. Astr. Soc.* **83**, 2, 319-330, DOI: 10.1111/j.1365-246X.1985.tb06489.x.

Kuvshinov, A.V. (2012), Deep electromagnetic studies from land, sea, and space: Progress status in the past 10 years, *Surv Geophys.* **33**, 1, 169-209, DOI: 10.1007/s10712-011-9118-2.

Kuvshinov, A., H. Utada, D. Avdeev, and T. Koyama (2005), 3-D modelling and analysis of Dst C-responses in the North Pacific Ocean region, revisited, *Geophys. J. Int.* **160**, 2, 505-526, DOI: 10.1111/j.1365-246X.2005.02477.x.

Leontovich, M.A. (1948), On approximate boundary conditions for an electromagnetic field on the surface of highly conductive bodies. **In:** *Issledovania po Rasprostraneniu Radiovoln*, Acad. Sci. USSR, Moscow, 5-12 (in Russian).

Logvinov, I.M. (2002), Applying the horizontal spatial gradient method for the deep conductivity estimations in the Ukraine, *Acta Geophys. Pol*. **50**, 4, 567-573.

Martinec, Z., and K. Pěč (1990), The influence of the core-mantle boundary irregularities on the mass density distribution inside the Earth. **In:** A. Vogel, C.O. Ofoegbu, R. Gorenflo, and B. Ursin (eds.), *Geophysical Data Inversion. Methods and Applications*, *Proc. 7th Int. Math. Geophys. Seminar, 8-11 February 1989, Free University of Berlin*, 233-256, DOI: 10.1007/978-3-322-89416-8_15.

Nowożyński, K. (2004), Estimation of magnetotelluric transfer functions in the time domain over a wide frequency band, *Geophys. J. Int*. **158**, 1, 32-41, DOI: 10.1111/j.1365-246X.2004.02288.x.

Olsen, N. (1992), Day-to-day C-response estimation for Sq from 1 cpd to 6 cpd using the Z:Y-method, *J. Geomag. Geoelectr.* **44**, 6, 433-447, DOI: 10.5636/jgg.44.433.

Olsen, N. (1998), The electrical conductivity of the mantle beneath Europe derived from C-responses from 3 to 720 hr, *Geophys. J. Int*. **133**, 2, 298-308, DOI: 10.1046/j.1365-246X.1998.00503.x.

Olsen, N. (1999a), Long-period (30 days – 1 year) electromagnetic sounding and the electrical conductivity of the lower mantle beneath Europe, *Geophys. J. Int.* **138**, 1, 179-187, DOI: 10.1046/j.1365-246x.1999.00854.x.

Olsen, N. (1999b), Induction studies with satellite data, *Surv. Geophys*. **20**, 3-4, 309-340, DOI: 10.1023/A:1006611303582.

Oraevsky, V.N., N.M. Rotanova, T.N. Bondar, D.Yu. Abramova, and V.Yu. Semenov (1993), On the radial geoelectrical structure of the mid-mantle from magnetovariational sounding using MAGSAT data, *J. Geomagn. Geoelectr.* **45**, 11-12, 1415-1423.

Pek, J. (2002), Spectral magnetotelluric impedances for an anisotropic layered conductor, *Acta Geophys. Pol.* **50**, 4, 619-643.

Pek, J., and F.A.M. Santos (2002), Magnetotelluric impedances and parametric sensitivities for 1-D anisotropic layered media, *Comput. Geosci.* **28**, 8, 939-950, DOI: 10.1016/S0098-3004(02)00014-6.

Petrishchev, M.S., and V.Yu. Semenov (2013), Secular variations of the Earth's apparent resistivity, *Earth Planet Sci. Lett*. **361**, 1-6, DOI: 10.1016/j.epsl.2012.11.027.

Piromallo, C., and A. Morelli (2003), P wave tomography of the mantle under the Alpine-Mediterranean area, *J. Geophys. Res.* **108**, B2, 2065, DOI: 10.1029/2002JB001757.

Praus, O., J. Pěčova, V. Červ, S. Kováčiková, J. Pek, and J. Velímský (2011), Electrical conductivity at mid-mantle depths estimated from the data of Sq and long period geomagnetic variations, *Stud. Geophys. Geod.* **55**, 2, 241-264, DOI: 10.1007/s11200-011-0014-5.

Roberts, R.G. (1984), The long-period electromagnetic response of the Earth, *Geophys. J. Roy. Astr. Soc.* **78**, 2, 547-572, DOI: 10.1111/j.1365-246X.1984.tb01963.x.

Roberts, R.G. (1986), The deep electrical structure of the Earth, *Geophys. J. Int*. **85**, 3, 583-600, DOI: 10.1111/j.1365-246X.1986.tb04534.x.

Rokitjansky, I.I. (1982), *Geomagnetic Investigation of the Earth's Crust and Mantle*, Springer, Berlin, 381 pp.

Rytov, S.M. (1940), Skin-effect calculations by the disturbance method, *J. Exp. Theor. Phys.* **10**, 2, 180-189 (in Russian).

Schmucker, U. (1970), *Anomalies of Geomagnetic Variations in the Southwestern United States*, University of California Press, Berkeley, 165 pp.

Schmucker, U. (1999a), A spherical harmonic analysis of solar daily variations in the years 1964-1965: response estimates and source fields for global induction – I. Methods, *Geophys. J. Int.* **136**, 2, 439-454, DOI: 10.1046/j.1365-246X.1999.00742.x.

Schmucker, U. (1999b), A spherical harmonic analysis of solar daily variations in the years 1964-1965: response estimates and source fields for global induction – II. Results, *Geophys. J. Int.* **136**, 2, 455-476, DOI: 10.1046/j.1365-246X.1999.00743.x.

Schmucker, U. (2003), Horizontal spatial gradient sounding and geomagnetic depth sounding in the period range of daily variations. **In:** A. Hördt and J.B. Stoll (eds.), *Protokoll über das 20. Kolloquium elektromagnetische Tiefenforschung, 29.09-3.10.2003, Königstein, Deutschland*, 228-237.

Schultz, A. (1990), On the vertical gradient and associated heterogeneity in mantle electrical conductivity, *Phys. Earth Planet. In.* **64**, 1, 68-86, DOI: 10.1016/0031-9201(90)90006-J.

Schultz, A., and J.C. Larsen (1987), On the electrical conductivity of the mid-mantle – I. Calculation of equivalent scalar magnetotelluric response functions, *Geophys. J. Int.* **88**, 3, 733-761, DOI: 10.1111/j.1365-246X.1987.tb01654.x.

Schultz, A., and J.C. Larsen (1990), On the electrical conductivity of the mid-mantle – II. Delineation of heterogeneity by application of extremal inverse solu-

tions, *Geophys. J. Int.* **101**, 3, 565-580, DOI: 10.1111/j.1365-246X.1990.tb05571.x.

Schultz, A., and T.S. Zhang (1994), Regularized spherical harmonic analysis and the 3-D electromagnetic response of the Earth, *Geophys. J. Int.* **116**, 1, 141-156, DOI: 10.1111/j.1365-246X.1994.tb02133.x.

Schultz, A., R.D. Kurtz, A.D. Chave, and A.G. Jones (1993), Conductivity discontinuities in the upper mantle beneath a stable craton, *Geophys. Res. Lett.* **20**, 24, 2941-2944, DOI: 10.1029/93GL02833.

Semenov, V.Yu. (1988), Analysis of magnetotelluric data during the anisotropic media sounding, *Russ. Geol. Geophys*. **10**, 121-125 (in Russian).

Semenov, V.Yu. (1989), Evaluation of mantle conductivity beneath northern hemisphere continents, *Izv. – Phys. Solid Earth* **25**, 3, 221-226 (in Russian).

Semenov, V.Yu. (2000), On the apparent resistivity in magnetotelluric sounding, *Izv. – Phys. Solid Earth* **36**, 1, 99-100.

Semenov, V.Yu., and W. Jóźwiak (1999), Model of the geoelectrical structure of the mid- and lower mantle in the Europe–Asia region, *Geophys. J. Int*. **138**, 2, 549-552, DOI: 10.1046/j.1365-246X.1999.00888.x.

Semenov, V.Yu., and W. Jóźwiak (2005), Estimation of the upper mantle electric conductance at the Polish margin of the East European platform, *Izv. – Phys. Solid Earth* **41**, 4, 326-332.

Semenov, V.Yu., and W. Jóźwiak (2006), Lateral variations of the mid-mantle conductance beneath Europe, *Tectonophysics* **416**, 1-4, 279-288, DOI: 10.1016/j.tecto.2005.11.017.

Semenov, V.Yu., and M. Rodkin (1996), Conductivity structure of the upper mantle in an activ subduction zone, *J. Geodynamics* **21**, 4, 355-364, DOI: 10.1016/0264-3707(95)00038-0.

Semenov, V.Yu., and V.N. Shuman (2010), Impedances for induction soundings of the Earth's mantle, *Acta Geophys*. **58**, 4, 527-542, DOI: 10.2478/s11600-010-0003-z.

Semenov, V.Yu., T. Ernst, K. Nowożyński, J. Pek, and EMTESZ WG (2005), Estimation of the deep geoelectrical structure beneath TESZ in NW Poland, *Publs. Inst. Geophys. Pol. Acad. Sci*. **C-95**, 386, 63-65.

Semenov, V.Yu., J. Pek, A. Ádám, W. Jóźwiak, B. Ladanyvskyy, I.M. Logvinov, P. Pushkarev, J. Vozar, and Experimental Team of CEMES project (2008), Electrical structure of the upper mantle beneath Central Europe: Results of the CEMES project, *Acta Geophys.* **56**, 4, 957-981, DOI: 10.2478/s11600-008-0058-2.

Semenov, V.Yu., B. Ladanivskyy, and K. Nowożyński (2011), New induction sounding tested in Central Europe, *Acta Geophys*. **59**, 5, 815-832, DOI: 10.2478/s11600-011-0030-4.

Semenov, V.Yu., M. Hvoždara, and J. Vozar (2013), Modeling of deep magnetovariation soundings on the rotating Earth, *Acta Geophys.* **61**, 2, 264-280, DOI: 10.2478/s11600-012-0086-9.

Senior, T.B.A., and J.L. Volakis (1995), *Approximate Boundary Conditions in Electromagnetics*, IEE Press, London, 353 pp.

Shuman, V.N. (1999), Scalar local impedance conditions and the impedance tensor in processing and interpretation of a magnetotelluric experiment, *Geophys. J.* **19**, 361-385.

Shuman, V. (2003), The general theory of geoelectromagnetic sounding systems accounting the electrodynamics of spherical sources. **In:** *3DEM III Workshop, February 2003, Adelaide, USA*, ASEG Ext. Abstr., 1-7, DOI: 10.1071/ASEG2003_3DEMab015.

Shuman, V.N. (2007), Imaginary surface vectors in multidimensional inverse problems of geoelectrics, *Izv. – Phys. Solid Earth* **43**, 3, 205-210, DOI: 10.1134/S1069351307030044.

Shuman, V., and S. Kulik (2002), The fundamental relations of impedance type in general theories of the electromagnetic induction studies, *Acta Geophys. Pol*. **50**, 4, 607-618.

Sokolova, E.Yu., I.M. Varentsov, and BEAR Working Group (2007), Deep array electromagnetic sounding on the Baltic Shield: External excitation model and implications for upper mantle conductivity studies, *Tectonophysics* **445**, 1-2, 3-25, DOI: 10.1016/j.tecto.2007.07.006.

Szarka, L., A. Franke, E. Prácser, and J. Kiss (2007), Hypothetical mid-crustal models of second-order magnetic phase transition. **In:** *4th Int. Symp. on Three-Dimensional Electromagnetics, 27-30 September 2007, Freiberg, Germany.*

Tikhonov, A.N. (1950), On determining electrical characteristics of the deep layers of the Earth's crust, *Dokl. AN USSR* **73**, 2, 295-297 (in Russian).

Transactions… (1973), Transactions of the 2nd IAGA Scientific Assembly, Kyoto, Japan, 1973, *IAGA Bull*. **35**, 189 pp.

Vanyan, L., B. Tezkan, and N. Palshin (2001), Low electrical resistivity and seismic velocity at the base of the upper crust as indicator of rheologically weak layer, *Surv. Geophys.* **22**, 2, 131-154, DOI: 10.1023/A:1012937410685.

Vanyan, L.L., V.A. Kuznetsov, T.V. Lyubetskaya, N.A. Palshin, T. Korja, I. Lahti, and the BEAR Working Group (2002), Electrical conductivity of the crust beneath Central Lapland, *Izv. – Phys. Solid Earth* **38**, 10, 798-815.

Vardanyants, I.L., and A.A. Kovtun (2009), The study of the possible existence of asthenosphere on the territory of Fennoscandian shield by the BEAR data. **In:** *Complex Geological-Geophysical Models of Ancient Shields,* Geological Inst. of the Kola Science Centre, Russ. Acad. Sc., Apatity, 15-18.

Varentsov, I.M., M. Engels, T. Korja, M.Yu. Smirnov, and the BEAR Working Group (2002), The generalized geoelectric model of Fennoscandia: A challenging database for long-period 3D modeling studies within the Baltic

electromagnetic array research (BEAR) Project, *Izv. – Phys. Solid Earth* **38**, 11, 855-896 (in Russian).

Vozar, J., and V.Yu. Semenov (2010), Compatibility of induction methods for mantle soundings, *J. Geophys. Res*. **115**, B3, B03101, DOI: 10.1029/2009JB 006390.

Wait, J.R. (1954), On the relation between telluric currents and the Earth's magnetic field, *Geophysics* **19**, 2, 281-289, DOI: 10.1190/1.1437994.

Weckmann, U., O. Ritter, and V. Haak (2003), Images of the magnetotelluric apparent resistivity tensor, *Geophys. J. Int*. **155**, 2, 456-468, DOI: 10.1046/ j.1365-246X.2003.02062.x.

Woods, D.V., and F.E.M. Lilley (1979), Geomagnetic induction in Central Australia, *J. Geomag. Geoelectr.* **31**, 4, 449-458, DOI: 10.5636/jgg.31.449.

Zhamaletdinov, A.A. (1996), Graphite in the Earth's crust and electrical conductivity anomalies, *Izv. – Phys. Solid Earth* **32**, 4, 272-288.

Zhamaletdinov, A.A. (2005), Khibiny MHD experiment: The 30th anniversary, *Izv. – Phys. Solid Earth* **41**, 9, 737-742.

Zhamaletdinov, A.A. (2011), The new data on the structure of the continental crust based on the results of electromagnetic sounding with the use of powerful controlled sources, *Dokl. Earth Sci.* **438**, 2, 798-802, DOI: 10.1134/ S1028334X11060146.

Zhamaletdinov, A.A., A.N. Shevtsov, T.G. Korotkova, Yu.A. Kopytenko, V.S. Ismagilov, M.S. Petrishev, B.V. Efimov, M.B. Barannik, V.V. Kolobov, P.I. Prokopchuk, M.Yu. Smirnov, S.A. Vagin, M.I. Pertel, E.D. Tereshchenko, A.N. Vasil'ev, V.F. Grigoryev, M.B. Gokhberg, V.I. Trofimchik, Yu.M. Yampolsky, A.V. Koloskov, A.V. Fedorov, and T. Korja (2011), Deep electromagnetic sounding of the lithosphere in the Eastern Baltic (Fennoscandian) shield with high-power controlled sources and industrial power transmission lines (FENICS experiment), *Izv. – Phys. Solid Earth* **47**, 1, 2-22, DOI: 10.1134/S1069351311010149.

Received 5 December 2013
Received in revised form 20 January 2014
Accepted 22 January 2014

DE GRUYTER OPEN

Acta Geophysica
vol. 63, no. 1, Feb. 2015, pp. 125-153
DOI: 10.2478/s11600-014-0243-4

Study of Shallow Low-Enthalpy Geothermal Resources Using Integrated Geophysical Methods

Lara De GIORGI and Giovanni LEUCCI

IBAM – National Council of Research, Lecce, Italy
e-mail: g.leucci@ibam.cnr.it

Abstract

The paper is focused on low enthalpy geothermal exploration performed in south Italy and provides an integrated presentation of geological, hydrogeological, and geophysical surveys carried out in the area of municipality of Lecce. Geological and hydrogeological models were performed using the stratigraphical data from 51 wells. A ground-water flow (direction and velocity) model was obtained. Using the same wells data, the ground-water annual temperature was modeled. Furthermore, the ground surface temperature records from ten meteorological stations were studied. This allowed us to obtain a model related to the variations of the temperature at different depths in the subsoil. Integrated geophysical surveys were carried out in order to explore the low-enthalpy geothermal fluids and to evaluate the results of the model. Electrical resistivity tomography (ERT) and self-potential (SP) methods were used.

The results obtained upon integrating the geophysical data with the models show a low-enthalpy geothermal resource constituted by a shallow ground-water system.

Key words: low enthalpy, 3D geological and hydrogeological models, 3D high resolution geophysics.

1. INTRODUCTION

One of the key components crucial in economic and social development is energy. Most energy sources at some stage are dependent on ecosystem services (Williamson and McCormick 2008). How these forms of energy are harnessed and employed makes the energy use a critical environmental issue, as its sourcing, production, transmission, and consumption often impact the ecosystems. Ecosystems are also keys to help in meeting the growing energy demand. Thus, to increase future energy supplies in a sustainable way, ecosystems need to be well-managed.

The current energy exploration has resulted in environmental changes with significant local, regional and global effects. Biodiversity loss and global warming threats continue to place an increasing demand for environmentally sound energy systems. Presently, the shift is toward clean renewable energy sources among which geothermal energy is the most attractive because of its relatively benign nature (de Jesus 1997).

The term "geothermal energy" is understood not only as the heat contained within the Earth. Geothermal energy is more often used to "indicate that part of the Earth's heat that can, or could be recovered and exploited by man" (Dickson and Fanelli 2004). In general, geothermal energy is extracted from geothermal reservoirs. According to the *Encyclopedia of Physical Science and Technology* (Meyers 1992), geothermal reservoirs are defined as a "geometrically definable volume of permeable rock which contains a proven reserve of thermal energy, such as water or steam that can be extracted in a practical, economic way". In other words, a geothermal reservoir could be considered a geothermal system consisting of three main elements: a heat source, a reservoir, and a fluid, which is the carrier that transfers the heat (Dickson and Fanelli 2004). This heat has two primary origins: (i) the decay of natural radioactive isotopes, and (ii) the original heat from gravitational energy released during the earth's formation. This heat is continuously streaming towards the surface by conduction, radiation and advection via mass transport. On average, the temperature in the continental crust increases with depth by 30 °K/km, reaching high temperatures in the earth's interior (~900 °C at the crust-mantle boundary).

Such high-enthalpy resources (Lee 2001) can be used for electricity generation, and their availability is restricted to regions with high heat flow. On the other sides, when geothermal energy is characterized by low temperature (low-enthalpy resources), it is often tapped using specialized technology. According to the temperature range, the low-enthalpy resources can be utilized for direct heating of dwellings (*e.g.*, using aquifer systems) (*e.g.*, Muffler and Cataldi 1978, Hochstein 1990, Benderitter and Cormy 1990, Haenel *et al.* 1988). Cooler resources (< 30 °C for compression heat pumps, < 50 °C

for absorption heat pumps) are made accessible through the use of the so-called geothermal heat pump systems.

In the context of possible use of low-enthalpy geothermal resources, the aim of this paper is to investigate the groundwater flow system that characterizes the Municipality of Lecce (Apulia Region, southern Italy), in order to verify its potentiality as source of geothermal low enthalpy energy for its application in heating and cooling systems to be used in private buildings.

1.1 Low-enthalpy geothermal resources

The main advantage of geothermal energy at low or intermediate temperature range is related to its direct use, because this resource is more widespread and exists at economic drilling depths (Lund 2007). In addition, parts of a geothermal system could be used, for example, as an underground thermal energy storage system (Andersson 2007). The general strategy of geothermal resource development starts with the development of a conceptual model of the potential geothermal resource.

The review of existing data, especially on geology (structural geology, stress field, hydraulic transport properties, thermal rock properties, petrography, and mineralogy), temperature, heat flow, and geochemistry form the basis to identify the potential geothermal reservoir and to estimate its size and its heat content. Low enthalpy geothermal energy, however, can be obtained at any place with the help of a ground-source heat pump from the soil, rocks, and from groundwater. It has a high potential for heating and cooling of buildings, especially in temperate and polar zones. However, an optimization of low-enthalpy exploitation requires general knowledge on the heat flow and groundwater flow properties of the aquifer. In addition, new exploration data should be integrated into site-specific geological models, useful for geothermal wells location and path planning, the risk assessment concerning the borehole stability, and the simulation of the reservoir behavior (during production). Models and computer simulations to predict the reservoir behavior are important tools useful for a sustainable, reliable, and efficient energy provision, for the planning of further reservoir treatments, and for the consideration of environmental aspects related to the exploitation of geothermal energy. The drilling of new wells, which provide additional information on the local conditions and, thus, implement the model, should be carried out according to the need of geothermal wells and the site-specific conditions. The last step in the development of the geothermal resource requires the installation and management of the whole geothermal system.

1.2 Geothermal energy and ecosystem services

Geothermal energy is defined as a carbon-free form of energy. However, the use of geothermal resources in a particular location can affect the availability

of groundwater resource and the land-use due to urban sprawl. There is the recognition that the overexploitation of natural resources has gravely affected many ecosystem services (MEA 2005). Ecosystem services are defined as the benefits that ecosystems provide to people (Costanza *et al.* 1997). They have been classified in: (i) provisioning ecosystem services (the provision of food, water, *etc.*), (ii) regulating ecosystem services (*i.e.*, erosion control), (iii) supporting ecosystem services (*i.e.*, soil formation), and (iv) cultural ecosystem services (recreation and life fulfillment) (MEA 2005). Several ecosystem services have a direct linkage with groundwater storage, even if this interdependency is little considered in decision-making processes. Groundwater system can provide important regulating services, such as water regulation, water purification and, indirectly, erosion regulation, flood control, and climate regulation (Bergkamp and Cross 2006). In terms of water regulation, groundwater plays an important role in the hydrological cycle and it is strongly connected with surface water resources (Falkenmark and Rockström 2004, Morris *et al.* 2003, Shaw 1994) by sustaining the stream flow during dry periods (Ward and Robinson 1990). For what concerns water purification, groundwater hosts a biological component that purifies water from organic compounds and potential human pathogens (Herman *et al.* 2001). The indirect regulating services are related, for instance, to the regulation of soil erosion, by providing water to vegetation cover (Malanson 1993), and to the buffering of climate variability (FAO 2003). In addition, groundwater can be seen in terms of supporting ecosystem services because it provides the basis for the hydrological cycle (Ward and Robinson 1990) and water is essential for living organisms.

At the same time, ecosystems, like wetlands, forests, and some kinds of land-covers, provide services that are crucial to maintain groundwater systems. Changes in land-cover, for instance, can affect the volume of recharge into groundwater aquifers. Thus, several management approaches are now becoming available to address groundwater depletion and deterioration, mainly based on new technologies and participatory approaches.

2. STUDY AREA

The study area is the Municipality of Lecce in the Salento peninsula (Apulia region, southern Italy) (Fig. 1). Apulia region is the emerged part of a plate stretching between the Ionian Sea and the Adriatic Sea, which constitutes the foreland of both Apenninic and Dinaric orogens. It comprises a Variscan basement covered by a 3-5 km thick Mesozoic carbonate sequence (the Calcari delle Murge unit), and overlain by thin deposits of Paleogene (Bosellini *et al.* 1999, Bossio *et al.* 1992, 1998, 1999, Margiotta 1999, Margiotta and Ricchetti 2002), Neogene (Bossio *et al.* 1992, 1994, 1998), and Quater-

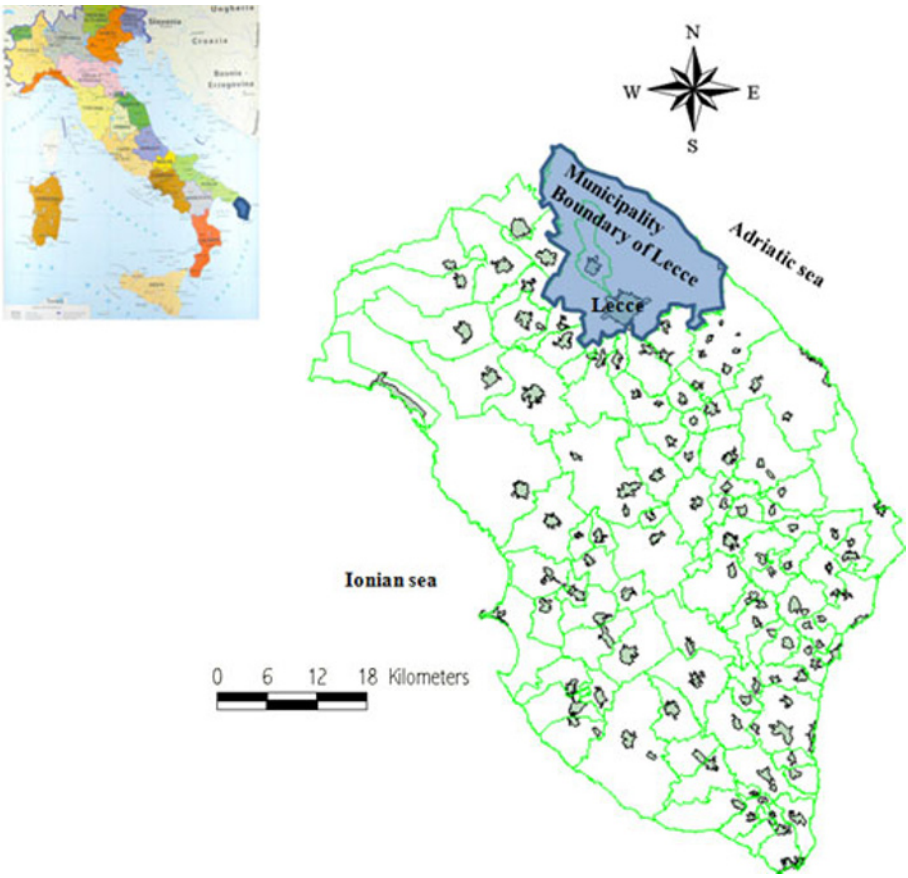


Fig. 1. The study area. Green lines indicate the city boundaries.

nary age (Bossio *et al.* 1987, D'Alessandro *et al.* 2004). The mid-southern part of Salento peninsula is marked by a wide endorheic area, bordered both toward the East and the West by degradated fault scarps. Marls, calcareous marls, and calcarenites belonging to several Pleistocene sedimentary cycles extensively crop out in the endorheic area. These deposits cover a stratigraphic sequence compound by calcareous units whose age is comprised between the Upper Cretaceous and the Upper Pliocene. The margins of the endorheic area are shaped on Lowest Pleistocene deposits (Bossio *et al.* 1987), up to 70 m thick, made by calcareous, bioclastic sandstones, locally clinostratified; they shade into bluish clayey marls toward centre of the area. The bottom of depressions is covered by thick sandy colluvial deposits. The most part of endorheic area is constituted by a flat surface, gently sloping northeastward, reaching 120 m of altitude at its SE part. The surface is

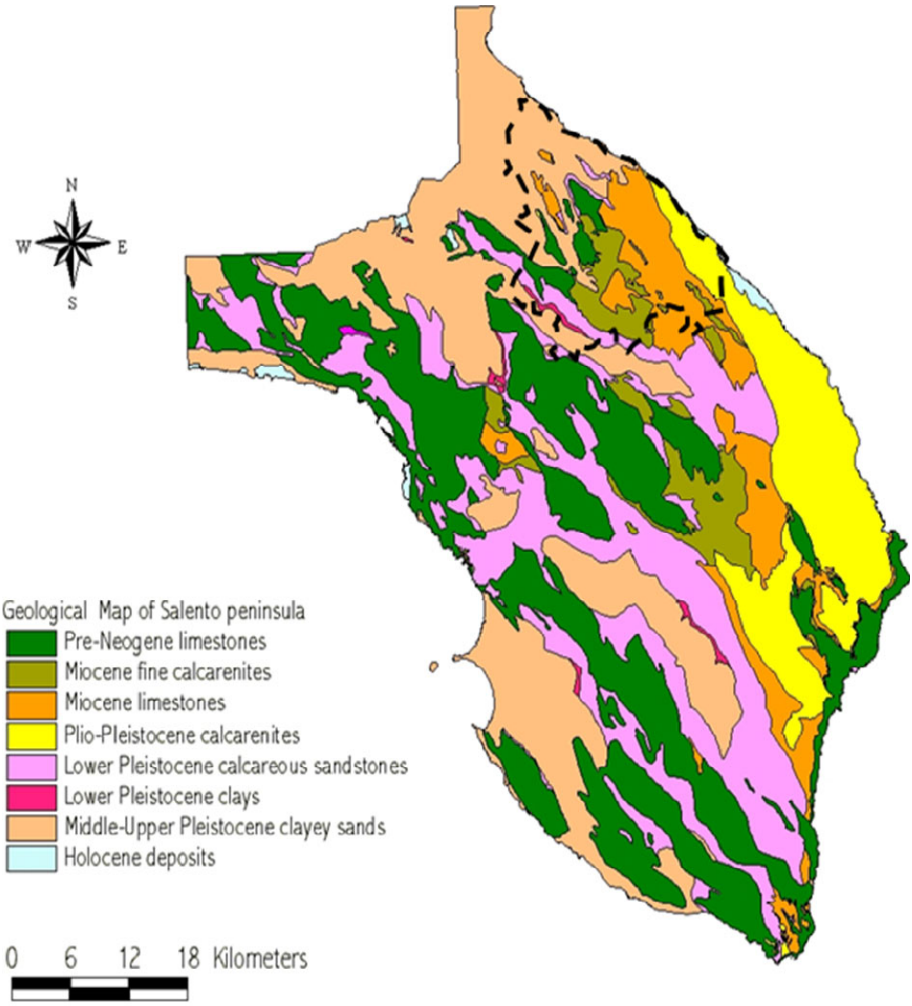


Fig. 2. Geological map of Salento penisula. The dashed dark line marks the boundaries of the study area.

shaped on white quartz sands that can be most likely referred to the Middle Pleistocene. To the north-west, a low relict cliff joins this surface to a wide marine terrace placed between 40-80 m of altitude whose deposits, made by coarse calcareous sandstones, lie transgressively on the Lower Pleistocene sandy and clayey deposits (Fig. 2) (D'Alessandro *et al.* 1994). The Salento peninsula is marked by a wide, deep aquifer hosted into the Mesozoic limestones, which remain on sea-water intruded from the nearby coastal areas, according to the Ghyben-Herzberg principle. However, a number of shallow

water tables occur in the most recent deposits. In particular, in the endorheic area several water tables can be found within the Lower Pleistocene calcareous sandstone and in the Middle Pleistocene sands even if their characteristics are not well known. However, a significant drainage from shallow water tables to the deep acquifer is most likely to occur along subvertical planes of greater hydraulic conductivity (Leucci *et al.* 2003).

In the study area marked with dashed dark line, on the basis of literature data (Margiotta and Negri 2004) the geological characterization is possible; in particular, the following eight litho-stratigraphical units were discriminated, in order from earliest to latest:

- Altamura limestone (upper Cretaceous),
- Galatone formation (late Oligocene),
- Lecce formation (late Oligocene – early Miocene),
- Pietra Leccese (Miocene-late Miocene),
- Andrano calcarenites (upper Miocene),
- Leuca formation (Pliocene),
- Uggiano la Chiesa formation (early-upper Pliocene),
- Gravina calcarenites (Quaternary).

Furthermore, according to the information extrapolated from several wells (Margiotta and Negri 2004), it is possible to highlight that there is a deep aquifer which is under pressure at the bed of the Miocene formation of Lecce stone, which shows a considerable thickness (over 50 m) and extending for several meters (also 75 m) under the sea level.

Margiotta and Negri (2004) highlighted the presence of four types of shallow aquifers located at different levels. The first three types are hosted in calcareous sediments and calcarenitic, in some cases marly and compact, which according to these characteristics are mainly low permeable. The last type of a shallow aquifer, instead, is hosted in sandy-type sediments, totally permeable, for cracking and porosity. The first aquifer is hosted in sediments of the Oligocene formation of Galatone and Lecce formation. This aquifer is also the most powerful discovery in the area because these surface waters are found from 25 m below sea level (b.s.l.) and up to 60 m b.s.l.

The second aquifer is hosted in sediments of Lecce stone and has a slightly smaller power than the last. This pitch was indeed found from a depth of 6 m above sea level (a.s.l.) up to 6 m b.s.l., with a total thickness of about 12 m. The third aquifer is hosted in Miocene sediments of the Calcarenite of Andrano, but it has a little thickness: 6 m. The latest aquifer is hosted in Plio-Pleistocene sediments of the Uggiano la Chiesa formation and Leuca formation. This aquifer is small and it has a thickness of about 5 m and is located entirely above the sea level.

3. MATERIALS AND METHODS

Using the software Groundwater Modeling System (GMS) a 3D geological model, containing lithological information, and a 3D groundwater model, containing water flow velocity, ground water thickness, and water flow direction, are performed. The scope of building a model is to simplify the field problem and organize the field data so that the system can be analyzed more readily (Anderson and Woessner 1992). The conceptualization of the model includes synthesis and framing up of data pertaining to geology, hydrogeology, hydrology, and meteorology.

The lithologic data were collected from 51 lithologs of drinking-water, and irrigation wells (Margiotta and Negri 2004) whose location is shown in Fig. 3. Depth of the lithologs ranges from 70 to 200 m b.g.l. Based on the acquired lithologs, a lithologic model was developed with Groundwater Modeling System (GMS). The lithologic solid model was created using a block-centric finite-difference grid. The gridding was performed through the eight nearest-neighbor methodology with 3D interpolation by average minimum distance. The solid model was developed to a depth of about 200 m b.g.l. The sensitivity of the model was tested by varying the horizontal and vertical spacing of the nodes, and an optimized final model, least sensitive to changes in spatial resolution (*i.e.*, smaller grid sizes) was built. The resolution of the final model was 100 m × 100 m × 2 m. The resulting dis-

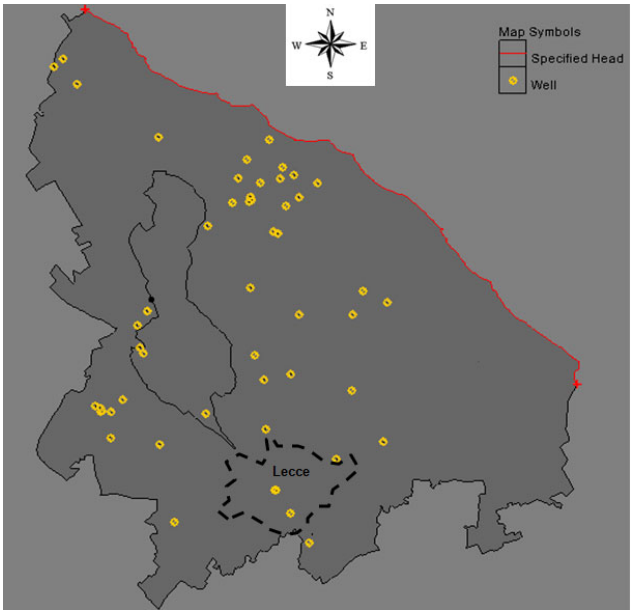


Fig. 3. Location of 51 wells inside the study area.

cretization consisted of 250 nodes × 250 nodes × 100 nodes, obtaining 6 250 000 solid model nodes, each with a voxel volume of 20 000 m^3. It should be noticed that the results of the modeling are not free from uncertainties, which could be reduced by increasing datapoints number, but they illustrate one of the most probable scenarios. The smallest scale of variation that the model is able to depict is equal to the resolution of one voxel.

Given the heterogeneity and complexity associated with the study area (*e.g.*, local-scale variations of aquifer properties, presence of numerous surface water bodies, *etc.*), some generalizations, simplifications, and assumptions were made to construct the groundwater flow model. The modeled study area is limited in the east by the Adriatic Sea, and it accounts for about 25 km in length and 25 km in width. The 3D constant-density groundwater flow was simulated by a block-centric, finite-difference grid model, using MODFLOW code inside GSM software. Furthermore, MODAEM code, which allows the analysis to the analytical elements (polygons, lines, and points), was used. MODAEM was used to extract information related to the depth and thickness of the aquifer inside the area under investigation.

The study area was horizontally discretized in the same way as described in the above paragraph. The maximum thickness of the model was 200 m. The model was vertically discretized in 5 layers. The layers had variable thickness as a result of wells stratigraphy. The 5th layer was defined as the basal no-flow boundary, consistent with the extensive clay layer at about 200 m below mean see level (m.s.l.). The layers were allowed to have seepage from the top and leakage through the base, making them hydraulically connected. The top-most layer and the second layer were defined as unconfined. The rest of the layers were defined as confined, as the water table was not expected to fall > 60 m (Leucci *et al.* 2003).

The conductivity was assigned to layers according to the modeled lithology. Vertical hydraulic conductivity (*kz*) was ranging from 0.86 to 0.64 m/d for the various types of calcarenite present in the study area and ranging from 2×10^{-2} to 0.8×10^{-2} m/d for the limestone (Sileo 2011). Other parameters were the horizontal anisotropy, the vertical anisotropy, the effective porosity, and the porosity.

What concerns anisotropy ratio, it is related to hydraulic conductivities in different directions. For example, vertical to horizontal hydraulic conductivity anisotropy ratio is given by *kr*/*kz*, where *kz* is the vertical hydraulic conductivity and *kr* is the radial (horizontal) hydraulic conductivity. Anisotropy in a horizontal plane is given by *kx*/*ky* where *kx* and *ky* are horizontal hydraulic conductivities in the *x* and *y* directions, respectively. In the case of the study area, hydraulic conductivity *kx* ranges from 0.36 to 0.067 m/d and for the various types of calcarenite present, and it ranges from 0.025 to 0.032 m/d for the limestone (Sileo 2011), while *ky* ranges from 0.31 to

0.054 m/d for the various types of calcarenite present, and it ranges from 0.013 to 0.018 m/d for the limestone (Sileo 2011). Consequently, the horizontal anisotropy ranges from 1.16 to 1.24 for the various types of calcarenite present in the study area, and it ranges from 1.92 to 1.78 for the limestone. On the other hand, the vertical anisotropy ranges from 0.39 to 0.094 for the various types of calcarenite present and from 0.95 to 0.31 for the limestone.

The total porosity is defined as the fraction of the bulk rock volume V that is not occupied by solid matter. It should be noticed that the porosity does not give any information concerning pore sizes, their distribution, and their degree of connectivity. Thus, rocks characterized by the same porosity could have widely different physical properties. An example of this might be a carbonate rock and a sandstone. Each could have a porosity of 0.2, but carbonate pores are very often unconnected resulting in its permeability being much lower than that of the sandstone. There is also an "effective" porosity defined as the ratio between the connected pore volume and the total volume (Juhasz 1986, Hill *et al.* 1979, Clavier *et al.* 1977). The most common definition of "effective" porosity is (Juhasz 1986, Hill *et al.* 1979, Clavier *et al.* 1977):

$$Fe = Ft - V_D \ , \tag{1}$$

where Ft is the total porosity of clean (clay free) sand, and V_D is the volume of dispersed clay in the sand pore space expressed as a fraction of the bulk volume. Sileo (2011) has performed experimental studies on the hydrological characteristics of the calcarenite, some of which are present in the study area, and she has found that the "effective" porosity ranges from 27 to 45% while the total porosity ranges from 32.4 to 54%.

No detailed previous data on the total recharge were available for the study area and, hence, it was one of the least certain input parameters. The total recharge inflow in the study area can be estimated on the basis of two fundamental sources: (i) precipitation, and (ii) irrigation return flow. In particular, the average annual precipitation was calculated using multiple year data (about 80) for six locations in and near the study, and it is 649.4 mm (Leucci *et al.* 2003). The total water from rain available for recharge can be termed as potential recharge (PR) and is defined as follows:

$$R - ET = PR = AR + SF \ , \tag{2}$$

where R is the total precipitation, ET is the evapotranspiration, SF is the surface overflow, and AR is the amount of meteoric water that actually recharges the aquifer. Accurate quantification of AR and SF requires long-term hydraulic head data and surface water stage measurements from multiple lo-

cations in the study area. Because such data were not available, *AR* was estimated indirectly by approximation of seasonal *PR* from meteoric data (80 years).

The *PR* was estimated as a difference of the total precipitation and the *ET* in a specific area. *ET* was calculated by the method of Pike (1964) as:

$$ET\,[\text{mm/month}] = \frac{P}{\left[1+\left(\frac{P}{PET}\right)^2\right]^{1/2}}\,, \tag{3}$$

where P is the average precipitation [mm/month], and *PET* is the potential evapotranspiration [mm/month]. Values of *PET* were approximated from the method of Malmström (1969):

$$PET\,[\text{mm/month}] = 40.9 \cdot K \cdot \tau\,, \tag{4}$$

where $k = 0.611 \exp(17.3\ \tau/\tau + 237.3)$ and τ is average monthly temperature (°C). The average *PR* calculated for the study area is 266 mm/year equal to 41% of the mean annual rainfall.

These calculated *PR* data were used to develop zonal recharge values by kriging them in the seasonal flow models.

In the studied area, hydraulic heads in observation wells ($n = 51$) fall in the range between 0.5 and 2.8 m because of extensive pumping. These withdrawals are replenished yearly, leading to similar groundwater levels at the beginning of each irrigation season. This indicates that the total annual recharge is equal to the difference in groundwater levels between wet and dry seasons.

Furthermore, the ground surface temperature records from ten meteorological stations were studied. The mean temperatures in a range of 15 years on the four seasons were examined. Soil temperature varies temporally and spatially; it is affected mainly by variations in air temperature and solar radiation. These variations influence the temperature at different depths in the subsoil. These variations at several depths in the subsoil can be estimated using a sinusoidal function (Hillel 1982, Marshall and Holmes 1988, Wu and Nofziger 1999). The variation of average soil temperature at different depths is described with the following sinusoidal function (Hillel 1982):

$$T(z,t) = \left[T_a + A_0 \exp(-z/d)\right] \sin\left[2\pi(t - t_0)/365 - (z/d) - (\pi/2)\right]\,, \tag{5}$$

where $T(z, t)$ is the soil temperature at time t [d] and depth z [m], T_a is the average soil temperature [°C], A_0 is the annual amplitude of the surface soil temperature [°C], d is the damping depth [m] of annual fluctuation, and t_0 is the time lag (days) from an arbitrary starting date (taken as 1 January in this

study) to the occurrence of the minimum temperature in a year. The damping depth is given by $d = (2D_h/\omega)^{1/2}$, where D_h is the thermal diffusivity, and $\omega = 2\pi/365\ d^{-1}$.

Amplitude is a parameter characterizing the seasonal variation of soil temperature around an average value. If the variation in temperature within a day is averaged out over many years, the seasonal amplitude is one-half the difference between this seasonal averaged maximum and seasonal averaged minimum temperatures within a season.

Damping depth is a constant characterizing the decrease in amplitude with an increase in distance from the soil surface. It is defined as $(2D_h/\omega)^{1/2}$, where D_h is the thermal diffusivity, and ω is the frequency of a temperature fluctuation. For example, in the annual fluctuation $\omega = 2\pi/365\ d^{-1}$.

Thermal diffusivity (Table 1) is the change in temperature produced in a unit volume by the quantity of heat flowing through the volume in unit time under a unit temperature gradient. It can be calculated from thermal conductivity and volumetric heat capacity.

Time lag is the number of days from an arbitrary starting date to the occurrence of the minimum temperature in a year.

Successively, the heat flow in the surveyed area was determined. For this purpose, the aquifer temperature method was used. The heat flow (q) is calculated using the following relation (de Lima Gomes and Hamza 2005):

$$q = \left[\left(T_{AQT} - T_0\right)\Big/\left(Z_{AQT} - Z_0\right)\right]\left[\left(Z_{AQT} - Z_0\right)\Big/\Sigma_i R_i h_i\right], \tag{6}$$

Table 1

Mean thermal diffusivity
at different depths (D'Arpa *et al.* 2012)

Depth [m]	Mean thermal diffusivity [10^{-5} m^2 s^{-1}]
0-5	0.123
6-10	0.303
11-20	0.829

Table 2

Mean thermal resistivity
at different depths (D'Arpa *et al.* 2012)

Depth [m]	Mean thermal resistivity [°C cm/W]
0-5	50.23
6-10	56.90
11-20	69.86

where T_{AQT} is the temperature at depth Z_{AQT}, R (Table 2) is the thermal resistivity of the layer with thickness h and n the number of layers, T_0 is the soil temperature.

In the study area, due to the needs of the municipality of Lecce, one zone was chosen in order to obtain detailed information about stratigraphy, groundwater depth, flow direction, and hydraulic conductivity (Fig. 11). Moreover, in this zone the results of the model flow were validated.

Resistivity imaging system, also called electrical resistivity tomography (ERT) uses 2D and 3D electrical imaging/topography surveys to map areas with complex geology. This kind of survey normally uses a large number of electrodes (may be 25 or more), connected to a multicore cable. The electrode spacing is normally a constant. A laptop computer together with electronic switching is used to automatically select the four relevant electrodes for each measurement (Griffiths and Barker 1993, Loke 2011). Results thus obtained are normally presented as resistivity 2D and 3D images with different colour combinations representing different resistivities, as shown later.

Furthermore, ERT method could be considered in order to visualize the change in water content in the subsurface and therefore it allows obtaining a direct measure of hydraulic conductivity. In fact, electrical resistivity is extremely sensitive to subsurface water content. Hydraulic conductivity could be measured using electrical time-lapse resistivity measurements (Binley *et al.* 2002, Deiana *et al.* 2007).

The self potential (SP) method is based on measuring the natural direct current potential between any two points on the ground (Telford *et al.* 1990). The potential differences are partly constant and partly fluctuating and are associated with electrical currents in the ground. Large anomalous potentials are often observed over sulphide and graphite ore bodies, magnetite and several other electrically conductive minerals (Lowrie 2007), and groundwater accumulations (Meisner 1962, Paul 1965). Selfpotential anomalies are also associated with water in subsurface structures and flow of water through the ground (Fig. 15). The streaming self potential of groundwater is usually indicated as a negative anomaly in the profile (Colangelo *et al.* 2006). Figure 10 represents horizontal groundwater flow (from right to left), which generates SP data as a linearly increasing line in the direction of flow. The slope of the line is a relative measure of the driving hydraulic gradient (Vichabian and Dale Morgan 2002).

4. RESULTS AND DISCUSSIONS

4.1 The geo-lithological model

The lithologic modeling suggests the presence of a very complex 3D hydrostratigraphic framework in the subsurface of the studied area. A detailed de-

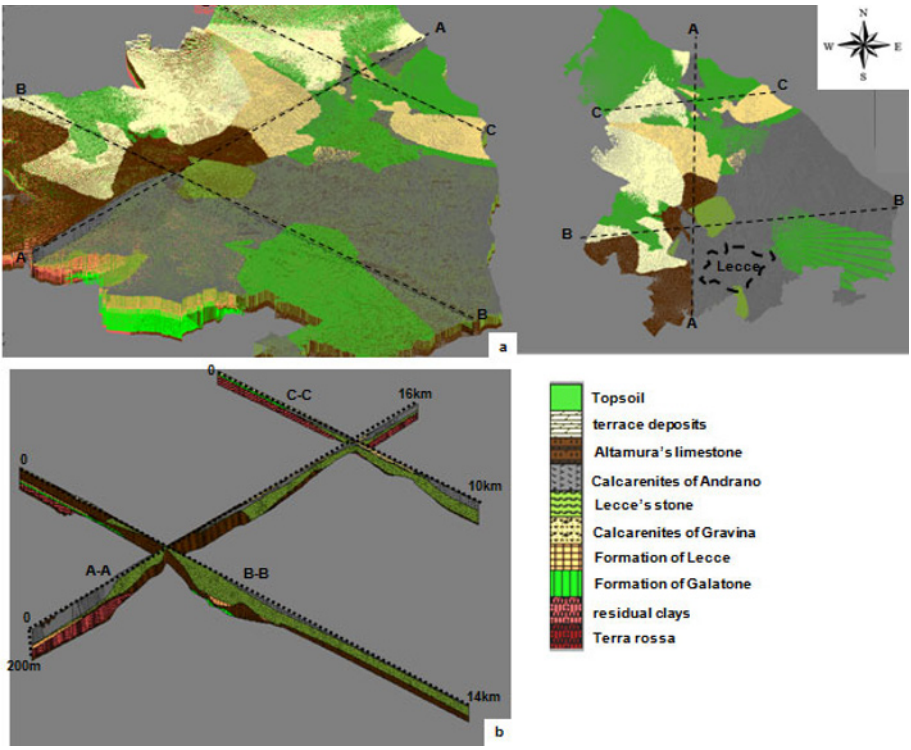


Fig. 4: (a) 3D lithological model, and (b) geological cross-sections.

scription of the study area at the block scale is provided here to illustrate the general spatial trends in aquifer thickness and spatial variability. The 3D lithologic model (Fig. 4a) with the cross-sections in Fig. 4b provide detailed depictions of the geology of the studied area.

4.2 The groundwater flow model

The sequence of modeling for the study area is shown in Figs. 5 and 6. Figure 5 shows the groundwater depth level (Fig. 5a) and groundwater thickness (Fig. 5b). As expected, the ground water level ranges between 1 and 58 m. In the south-western part of the study area, a greater depth of groundwater level is seen. In this zone, groundwater is at about 4.5 m a.s.l. It represents the second aquifer hosted in sediments of Lecce stone.

The velocity flow model (Fig. 6) suggests that it was topographically controlled. The groundwater flow velocity ranging from 5×10^{-8} to 14.5×10^{-8} m/s and its direction is essentially from the southern boundary of the study area to the north-west.

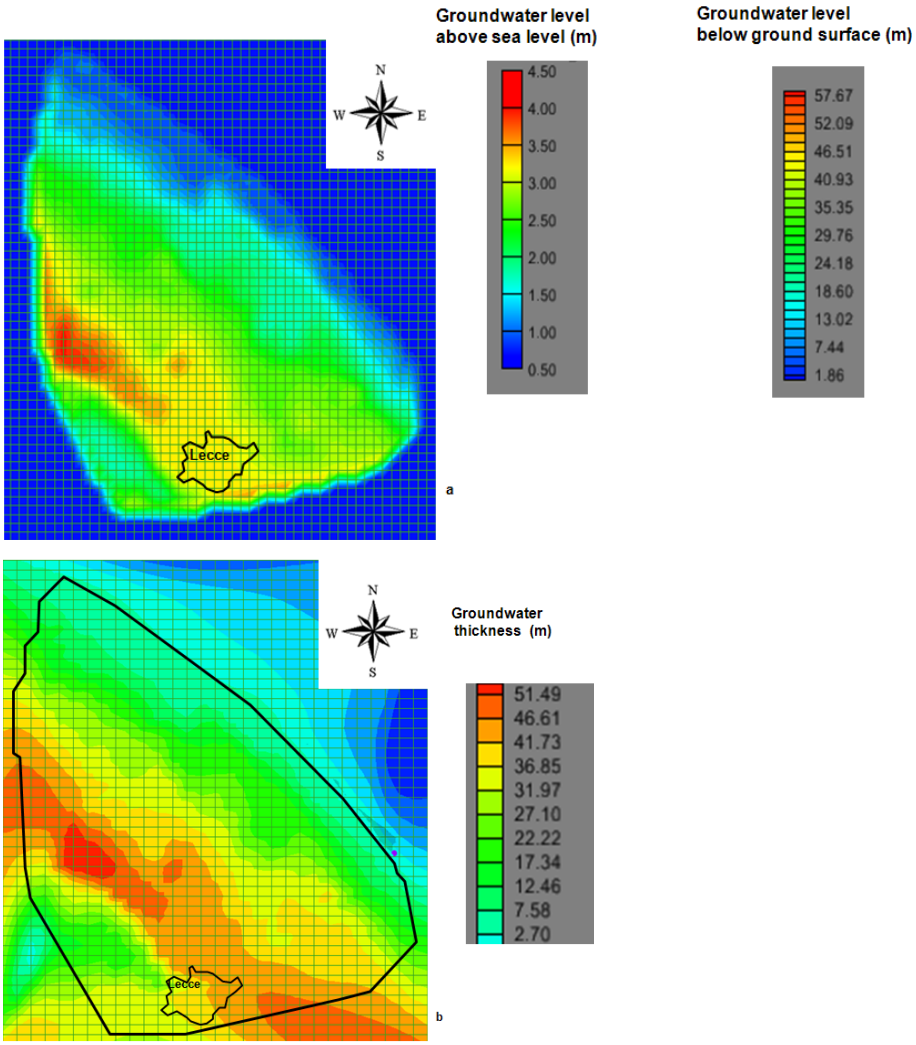


Fig. 5. The second aquifer hosted in sediments of Lecce stone: (a) groundwater depth level, and (b) ground water thickness.

4.3 The ground temperature model

Soil temperature data at different depths for the study area are shown in Fig. 7.

In autumn, the temperature ranged from 19 to 22°C at the soil surface, while at the 10 m depth it ranged from 20.5 to 24.6°C. At the 20 m depth it ranged from 18 to 21°C.

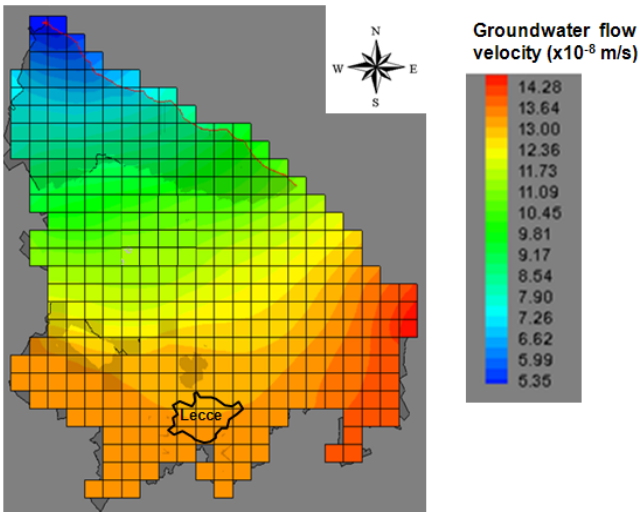


Fig. 6. Groundwater velocity flow model.

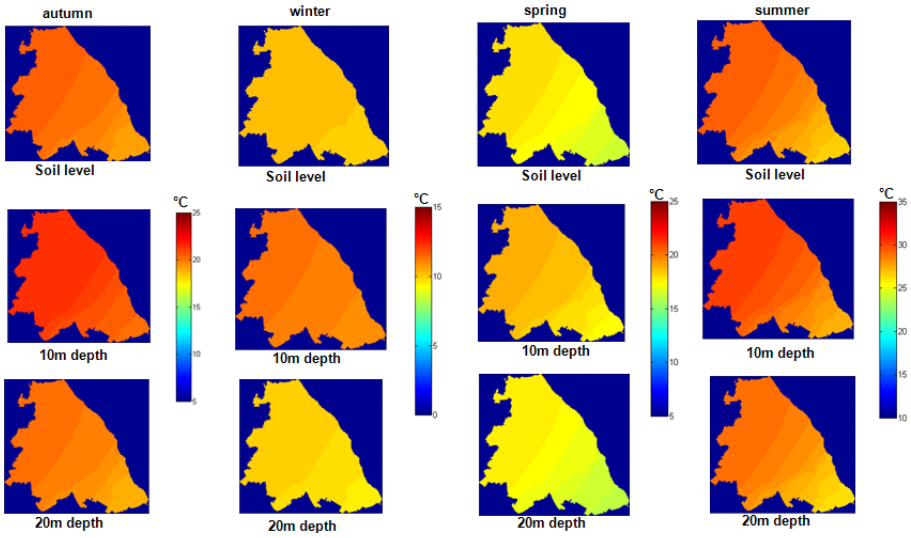


Fig. 7. Seasonal soil temperature map at different depths for the study area.

In winter, the temperature ranged from 10 to 12°C at the soil surface, while at the 10 m depth it ranged from 11 to 13°C. At the 20 m depth it ranged from 10 to 11°C.

In spring, the temperature ranged from 16 to 20°C at the soil surface, while at the 10 m depth it ranged from 17 to 21°C. At the 20 m depth it ranged from 16 to 19.4°C.

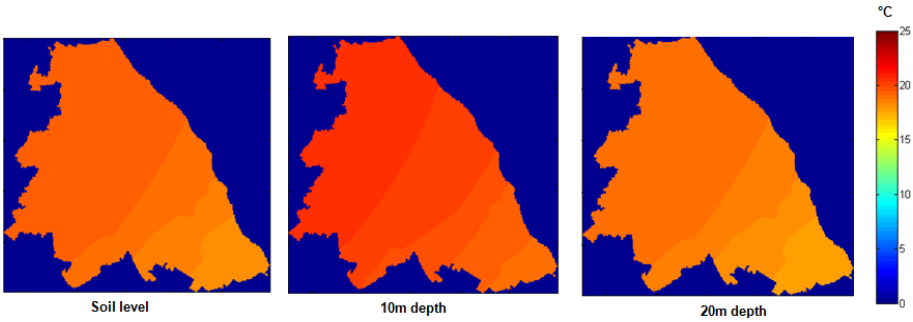


Fig. 8. Annual soil temperature map at different depths for the study area.

In summer, the temperature ranged from 26 to 33°C at the soil surface, while at the 10 m depth it ranged from 27 to 34°C. At the 20 m depth it ranged from 26 to 30°C.

The annual soil temperature data at different depths for the study area are shows in Fig. 8. Here, the temperature ranged from 16 to 22°C at the soil surface, while at the 10 m depth it ranged from 18 to 24°C. At the 20 m depth it ranged from 16 to 22°C.

4.4 The heat flow model

The first step in the heat flow determination was the evaluation of groundwater temperature in the study area. For this purpose, a geostatistical analy-

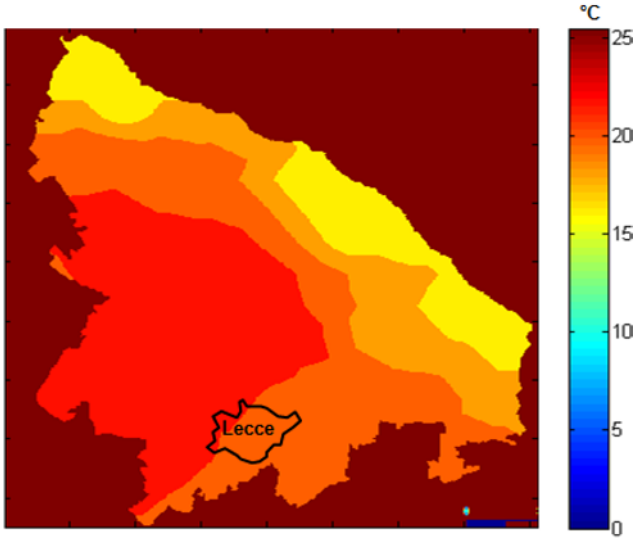


Fig. 9. Annual groundwater temperature map for the study area.

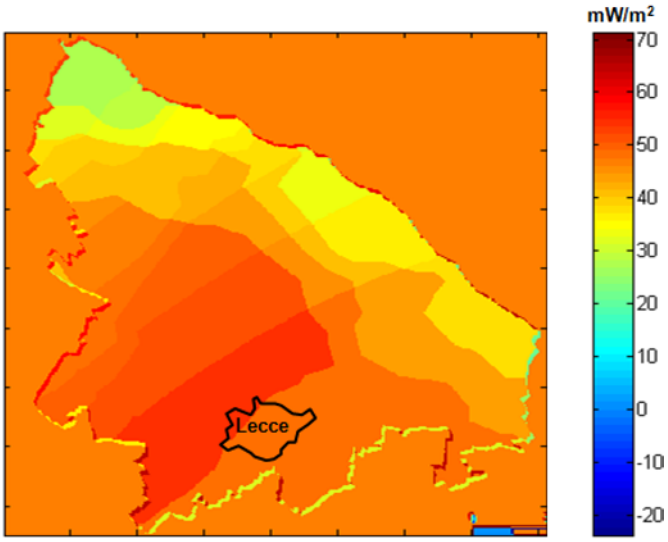


Fig. 10. Heat flow map for the study area.

sis was performed on the annual temperature data from the 51 wells used for the lithological model in Section 4.1. The groundwater temperature map is shows in Fig. 9. Here at the temperatures ranged from 16 to 22°C.

Based on this map and Table 2, Eq. 6 allows to make the heat flow map in the study area (Fig. 10). The heat flow map reveals that the heat flow is in the range of 25 to 65 mW/m^2. Heat flow values lower than 30 mW/m^2 were encountered on the sea line. Heat flow values higher than 50 mW/m^2 were found in the south western parts of the study area. It appears that the heat flow regime of the eastern part of the study area is distinctly different from that of the western part.

4.5 2D ERT time lapse measurements

In order to obtain detailed information about stratigraphy and groundwater depth in the selected zone (Fig. 11), the first geophysical measurement was the ERT time lapse in order to obtain information about the hydraulic conductivity.

The ERT data were collected in a straight-line array of 24 electrodes with 5 m spacing, W-E oriented. The Wenner array was chosen because it provides a good signal-to-noise ratio. Additionally, it is also highly sensitive to vertical changes in the subsurface resistivity. This makes the Wenner array a useful tool in studying the movement of the wetting front in time. The investigation depth was about 20 m, easily achievable with the chosen elec-

Fig. 11. The time-lapse ERT profile, 3D ERT area, and SP location in the selected zone.

trode configuration. The IRIS–SYSCAL Pro Switch 48 instrument was used to acquire electrical resistivity measurements.

To obtain 2D resistivity models, the field data were inverted using Res2Dinv software (Griffiths and Barker 1993, Loke and Barker 1996). In the processing of resistivity data, an inversion routine based on the smoothness-constrained least-squares method was implemented (deGroot-Hedlin and Constable 1990). The 2D model used by the inversion program divides the subsurface into a number of rectangular blocks, whose arrangement is linked to the distribution of points in pseudosections. The distribution and size of the blocks is generated automatically by the program, using the distribution of the data points as a rough guide. The depth of the bottom row of blocks is set to be approximately equal to the equivalent depth of investigation (Edwards 1977) of the data points. The optimization method basically tries to reduce the difference between the calculated and measured apparent resistivity values, by adjusting the resistivity of the model blocks. A forward modeling subroutine is used to calculate the apparent resistivity values, and a non-linear least-squares optimization technique is used for the inversion routine (Loke and Barker 1996). A measure of this difference is given by the root-mean squared (RMS) error.

The objective of these measurements was to obtain the field hydraulic conductivity. The water was inserted using a large ring infiltrometer installed directly at the x-axis of between 30 and 40 and between 80 and 90 m along the ERT profile above the ground surface in both tests. First measurements were performed before the water was inserted (Fig. 12a). The water pumps were used in order to insert water in the ring infiltrometer. A stratigraphy

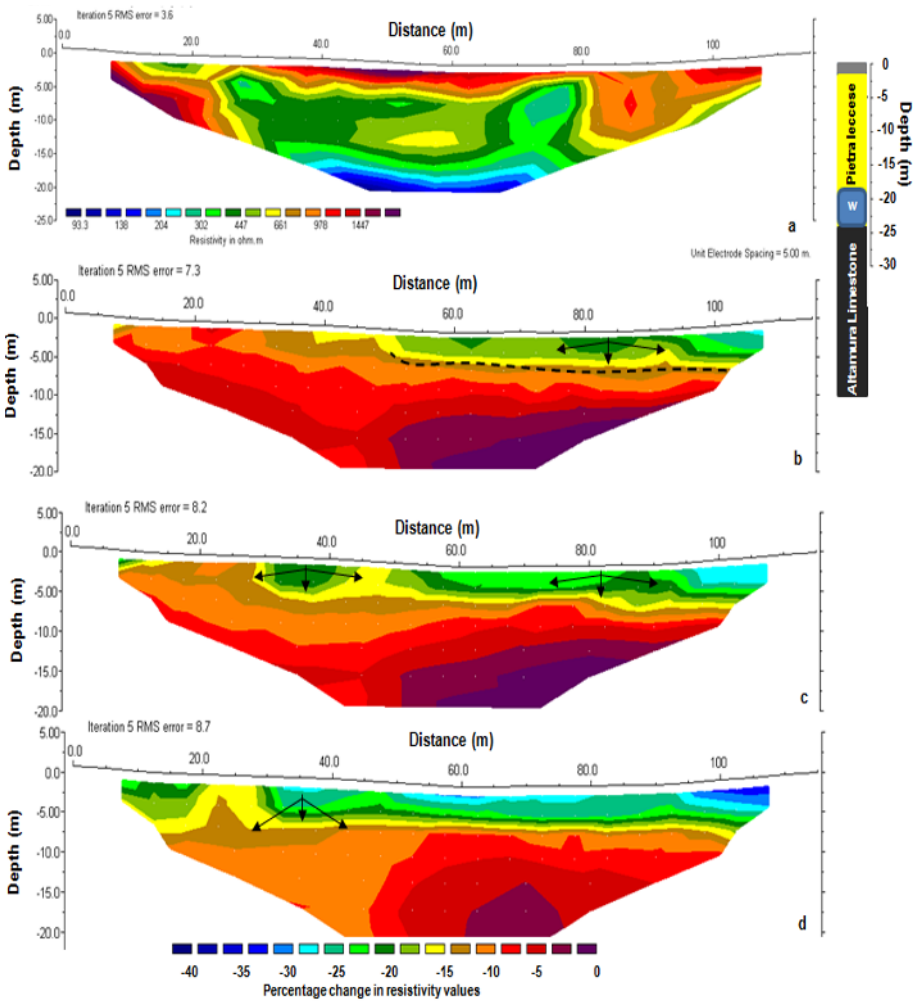


Fig. 12. Time lapse model: (a) resistivity distribution in dry condition, (b) percentage resistivity model during approximately 3 days of water infiltration, (c) percentage resistivity model during approximately 6 days of water infiltration, and (d) percentage resistivity model during approximately 22 days of water infiltration.

available and related to a well located near the investigated area allows to have a control on the results of the ERT profile. It is possible to notice a first layer composed by topsoil. The second layer is the Pietra Leccese, which was associated with the resistivity values ranging from 300 to 1400 ohm m (about 16 m thickness). The wide range of variability of the values of resistivity could be due to both the degree of fracturing and the variation of the water content within the Pietra Leccese. At about 18 m in depth, it is

possible to evidence the water table with resistivity values ranging from 90 to 110 ohm m.

During approximately 3 days of the water infiltration (Fig. 12b), the specific resistivity changed by about 17% while the infiltrating water has reached a depth of 4 m vertically and traveled 20 m horizontally. Probably, this is due to the high fracture degree.

In the sixth day of the water infiltration (Fig. 12c), the specific resistivity changes were of about 5-10% at depths between about 5 and 10 m.

In the twenty-second day of the water infiltration (Fig. 12d), the specific resistivity changes were of about 8% at depths between about 5 and 20 m.

During the experiment, the resistivity decreased, on the average, by 10% in the first 20 m depth. Therefore, water has traveled vertically an average of 0.9 m/d. Horizontally, the water has traveled an average of 7 m in twenty-two days (0.35 m/d). These values are in agreement with the values found experimentally on several calcarenite samples by Sileo (2011).

4.6 3D ERT measurements

A 48-channel Syscal-Pro Resistivity-meter was used for the survey. Resistivity field data were collected in a rectangular area of 235 by 60 m, along 5 m spaced parallel profiles oriented approximately W-E (Fig. 11). For each profile, 48 electrodes were used.

The multi-electrodes field procedure using the Wenner electrodes arrangement was applied to fulfill the resistivity measurements. The electrode separation was 5 m.

The measured data were processed by means of 3D inverse modelling software (ERTLab by Geostudy Astier). Its numerical core is based on tetrahedral Finite Elements (Zienkiewicz and Taylor 1989, Cifuentes and Kalbag 1992).

The 3D image of resistivity is shown in Fig. 13. The 3D visualization proposed in Fig. 13a shows the resistivity distribution of the upper 40 m in which it is possible to note:

(i) a zone of low resistivity (about 450-500 ohm m), from the surface to about 3 m in depth;

(ii) a zone in which resistivity values range between 700 and 850 ohm m, from about 3 m to about 10 m of depth;

(iii) a zone with resistivity values ranging between 300 and 450 ohm m, from about 10 to about 20 m of depth;

(iv) a water table zone located between about 20 and 40 m depth.

The 3D images of resistivity can easily be visualized by 3D contouring of iso-resistivity volumes (Fig. 13b). In this representation, the transparency function is defined by two threshold values of the resistivity, ρ_1 and ρ_2

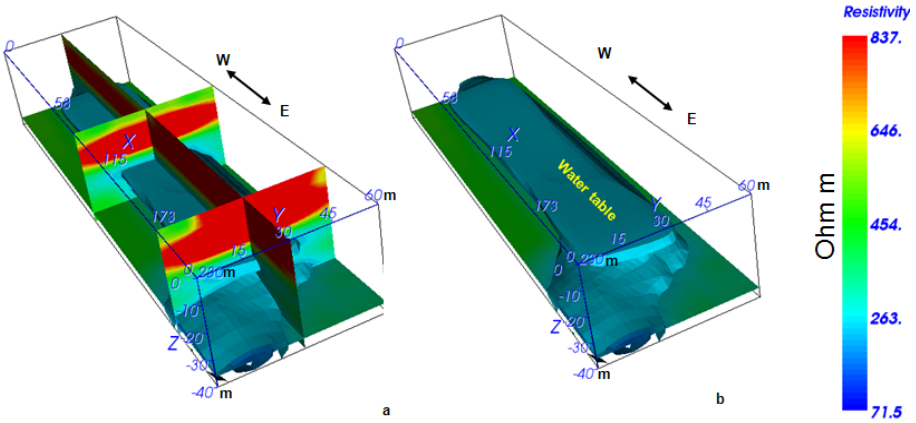


Fig. 13. The 3D ERT results: (a) resistivity distribution of the upper 40 m, and (b) iso-resistivity volume using a threshold values ranging from 70 to 110 ohm m.

($\rho_1 < \rho_2$). In the intervals $\rho < \rho_1$ and $\rho > \rho_2$, data are rendered as transparent; therefore, only the data in the interval $\rho_1 < \rho < \rho_2$ are visualized. In Fig. 13b, the resistivity data set is displayed with iso-resistivity volumes using a threshold value ranging from 70 to 110 ohm m. The continuous high-resistivity area, interpreted as water table, is more visible.

4.7 3D SP measurements

The SP measurements covered the 3D ERT surveyed area (Fig. 11). The self-potential signals were measured at the ground surface in a set of 432 measurement points located along nine parallel lines.

Each electrode (stainless and nonpolarizing) was placed inside a 10 cm deep hole, filled with a moistened bentonite and gypsum mixture to ensure good contact between the electrode and the ground, and stones were placed above the electrodes. Measurements of the self-potential signals were carried out with a Keithley 2701 multichannel voltmeter, and we used nonpolarizing Pb/PbCl2 (Petiau) electrodes (Perrier *et al.* 1997). The voltmeter was connected to a laptop computer where the data were recorded. All the electrodes were scanned during a period of 30 s.

The SP data were filtered with a low-pass filter in the frequency domain in order to avoid edge effects of space domain filters, so that high frequencies were eliminated and low frequencies were preserved (Aubanel and Oldham 1985). These data are used to build self-potential map shown in Fig. 14.

Self-potential anomalies vary in value according to their source (Reynolds 1998). If the value of self-potential is positive or negative of one-hundred millivolts, the cause is the movement of ground water. Is important

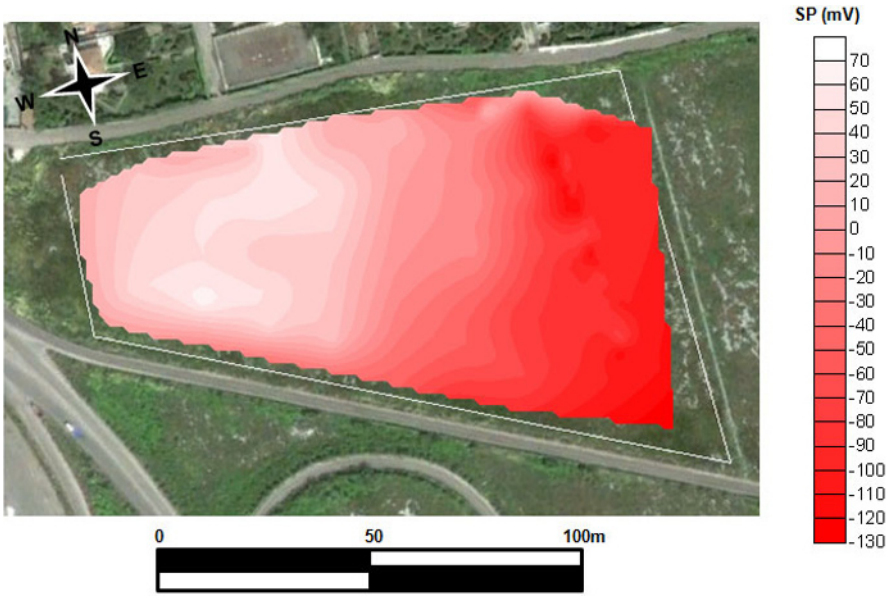


Fig. 14. The self-potential distribution map in the selected zone.

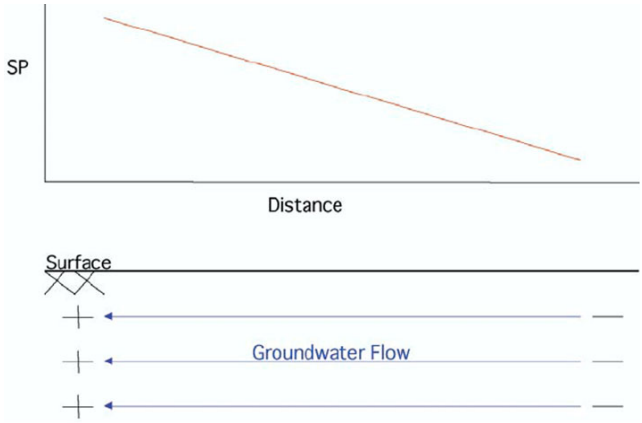


Fig. 15. Schematic response of self-potential distribution in the subsoil with associated water flow.

to note that the water flow direction is from negative to positive values of self-potential (Fig. 15). The result of this research shows that the self-potential values vary between –130 and 70 mV at a distance of approximately 230 m. This result can be interpreted as being due to a movement of groundwater due to the water table at about 20 m depth.

Figure 14 helps us to understand the complexity in local water flow direction that mainly confirms the predominant north-west direction and confirms the results of ground water flow model.

5. CONCLUSIONS

The proposed study allows the identification of areas suitable for the utilization of low enthalpy geothermal energy. For the studied area, data from stratigraphical information relative to 51 wells were collected. These data were integrated with geophysical data. This allowed the identification of the local geology, and particularly the depth of the groundwater table and the principal directions of the groundwater flow. The identification of the local geology (lithological classification) to determine subsoil physical characteristics (such as thermal resistivity, thermal diffusivity) related to the studied site was made. Furthermore, the ground surface temperature records from ten meteorological stations allowed us to model the variations of temperature at different depths in the subsoil.

The major findings of this research can be summarized as follows:

- The mean value of geothermal gradient was 0.2 °C/m;
- The ground-water temperature varies from 16 to 22 °C. It seem to influence the subsoil temperature variations at 20 m depth;
- The estimated heat flow ranges from 25 to 65 mW/m^2, with a mean value of 45 mW/m^2. It is higher in the northwestern part of the periphery of Lecce;
- The ground-water flow direction was north-west according to the hydroizohypses distribution of 51 analyzed wells;
- The velocity flow model suggests that it was topographically controlled and its velocity was ranging from 5×10^{-8} to 14.5×10^{-8} m/s;
- The geophysical survey makes it possible to give detailed information relating to the investigated zone that confirms the results of the model;
- This information allows establishing, within the same areas, the most suitable points for the installation of a geothermal low enthalpy plants.

References

Anderson, M.P., and W.W. Woessner (1992), *Applied Groundwater Modeling: Simulation of Flow and Advective Transport*, Academic Press Inc., San Diego, 381 pp.

Andersson, O. (2007), Aquifer thermal energy storage (ATES). **In:** H.Ö Paksoy (ed.), *Thermal Energy Storage for Sustainable Consumption*, Springer, Berlin Heidelberg, 155-176, DOI: 10.1007/978-1-4020-5290-3_8.

Aubanel, E.E., and K.B. Oldham (1985), Fourier smoothing without the fast Fourier transform, *Byte* **10**, 2, 207-218.

Benderitter, Y., and G. Cormy (1990), Possible approach to geothermal research and relative cost estimante. **In:** M.H. Dickson, and M. Fanelli (eds.), *Small Geothermal Resources*, UNITAR/UNDP Centre for Small Energy Resources, Rome, Italy, 61-71.

Bergkamp, G., and K. Cross (2006), Groundwater and Ecosystem Services: towards their sustainable use. **In:** *Proc. Int. Symp. on Groundwater Sustainability (ISGWAS), Alicante, Spain*, 177-193, http://aguas.igme.es/igme/ISGWAS/Ponencias%20ISGWAS/13-Bergkamp.pdf.

Binley, A., G. Cassiani, R. Middleton, and P. Winship (2002), Vadose zone flow model parameterisation using cross-borehole radar and resistivity imaging, *J. Hydrol*. **267**, 3-4, 147-159, DOI: 10.1016/S0022-1694(02)00146-4.

Bosellini, A., F.R. Bosellini, M.L. Colalongo, M. Parente, A. Russo, and A. Vescogni (1999), Stratigraphic architecture of the Salento coast from Capo d'Otranto to S. Maria di Leuca (Apulia, southern Italy), *Riv. Ital. Paleontol. S.* **105**, 3, 397-416.

Bossio, A., F. Guelfi, R. Mazzei, B. Monteforti, and G. Salvatorini (1987), Studies on the Neogene and Quaternary of Salento peninsula. III – Stratigraphy of the well of Poggiardo, *Quad. Ric. Centro Studi Geotecn. d'Ing. Lecce* **11**, 55-88 (in Italian).

Bossio, A., R. Mazzei, B. Monteforti, and G. Salvatorini (1992), Preliminary news about the Miocene of S. Maria al Bagno – S. Caterina, at Nardo (Lecce), *Paleopelagos* **2**, 99-107 (in Italian).

Bossio, A., F. Guelfi, R. Mazzei, B. Monteforti, and G. Salvatorini (1994), The Miocene succession of the Calcareniti of Andrano (Puglia, southern Italy), *Boll. Soc. Paleont. It*. **33**, 2, 249-255 (in Italian).

Bossio, A., D. Esu, L.M. Foresi, O. Girotti, A. Iannone, E. Luperto, S. Margiotta, R. Mazzei, B. Monteforti, G. Ricchetti, and G. Salvatorini (1998), Formation of Galatone, the new name for a lithostratigraphic unit of Salento (Puglia, southern Italy), *Atti Soc. Tosc. Sc. Nat. Mem. A* **105**, 151-156 (in Italian).

Bossio, A., L. Foresi, S. Margiotta, R. Mazzei, B. Monteforti, and G. Salvatorini (1999), Geological map of the north east of the province of Lecce, scale of 1: 25000; sector 7, 8, 10 scale 1: 10000, Università degli Studi di Siena (in Italian).

Cifuentes, A.O., and A. Kalbag (1992), A performance study of tetrahedral and hexahedral elements in 3-D finite element structural analysis, *Finite Elem. Anal. Des.* **12**, 3-4, 313-318, DOI: 10.1016/0168-874X(92)90040-J.

Clavier, C., G. Coates, and J. Dumanoir (1997), Theoretical and experimental bases for the dual-water model for interpretation of shaly sands. **In:** *Proc. 52nd Annual Meeting, Society of Petroleum Engineering, Denver, USA*, Rep. SPE-6859-PA, preprint 16 pp.

Colangelo, G., V. Lapenna, A. Perrone, S. Piscitelli, and L. Telesca (2006), 2D self-potential tomographies for studying groundwater flows in the Varco d'Izzo landslide (Basilicata, southern Italy), *Eng. Geol.* **88**, 3, 274-286, DOI: 10.1016/j.enggeo.2006.09.014.

Costanza, R., R. d'Arge, R. de Groot, S. Farber, M. Grasso, B. Hannon, K. Limburg, S. Naeem, R.V. O'Neill, J. Paruelo, R.G. Raskin, P. Sutton, and M. van den Belt (1997), The value of the world's ecosystem services and natural capital, *Nature* **387**, 253-260, DOI: 10.1038/387253a0.

D'Alessandro, A., G. Mastronuzzi, G. Palmentola, and P. Sansò (1994), Pleistocene deposits of Salento leccese (Southern Italy): problematic relationships, *Boll. Soc. Paleont. It.* **33**, 2, 257-263.

D'Alessandro, A., F. Massari, E. Davaud, and G. Ghibaudo (2004), Pliocene–Pleistocene sequences bounded by subaerial unconformities within foramol ramp calcarenites and mixed deposits (Salento, SE Italy), *Sediment. Geol.* **166**, 1-2, 89-144, DOI: 10.1016/j.sedgeo.2003.11.017.

D'Arpa, S., N. Zaccarelli, D.E. Bruno, G. Leucci, V.F. Uricchio, and G. Zurlini (2012), A geographically weighted regression model for geothermal potential assessment in mediterranean cultural landscape. **In:** *Proc. EGU General Assembly, 22-27 April 2012, Vienna, Austria*, 12432.

de Groot-Hedlin, C., and S. Constable (1990), Occam's inversion to generate smooth, two-dimensional models form magnetotelluric data, *Geophysics* **55**, 12, 1613-1624, DOI: 10.1190/1.1442813.

de Jesus, A.C. (1997), Environmental sustainability of geothermal development, *Energ. Source.* **19**, 1, 35-47, DOI: 10.1080/00908319708908830.

de Lima Gomes, A.J., and V.M. Hamza (2005), Geothermal gradient and heat flow in the state of Rio de Janeiro, *Rev. Brasil. Geofis*. **23**, 4, 325-347, DOI: 10.1590/S0102-261X2005000400001.

Deiana, R., G. Cassiani, A. Kemna, A. Villa, V. Bruno, and A. Bagliani (2007), An experiment of non-invasive characterization of the vadose zone via water injection and cross-hole time-lapse geophysical monitoring, *Near Surf. Geophys.* **5**, 3,183-194, DOI: 10.3997/1873-0604.2006030.

Dickson, M.H., and M. Fanelli (2004), *What is Geothermal Energy?* Instituo di Geoscienze e Georisorce, Pisa, Italy.

Edwards, L.S. (1977), A modified pseudosection for resistivity and IP, *Geophysics* **42**, 5, 1020-1036, DOI: 10.1190/1.1440762.

Falkenmark, M., and J. Rockström (2004), *Balancing Water for Humans and Nature: The New Approach in Ecohydrology*, Earthscan, London, 247 pp.

FAO (2003), *Groundwater Management – The Search for Practical Approaches*, Water Reports 25, Food and Agriculture Organization of the United Nations, Rome, Italy.

Griffiths, D.H., and R.D. Barker (1993), Two-dimensional resistivity imaging and modelling in areas of complex geology, *J. Appl. Geophys.* **29**, 3-4, 211-226, DOI: 10.1016/0926-9851(93)90005-J.

Haenel, R., L. Rybach, and L. Stegena (1988), Fundamentals of geothermics. **In:** R. Haenel, L. Rybach, and L. Stegena (eds.), *Handbook of Terrestrial Heat-Flow Density Determination*, Kluwer Academic Publ., Dordrecht, 9-57, DOI: 10.1007/978-94-009-2847-3_2.

Herman, J.S., D.C. Culver, and J. Salzman (2001), Groundwater ecosystems and the service of water purification*, Stanford Environ. Law J.* **20**, 479-495.

Hill, H.J., O.J. Shirley, and G.E. Klein (1979), Bound water in shaley sands – its relation to Qv and other formation properties, *The Log Analyst* **20**, 3, 3-19.

Hillel, D. (1982), *Introduction to Soil Physics*, Academic Press, NewYork.

Hochstein, M.P. (1990), Classification and assessment of geothermal resources. **In:** M.H. Dickson and M. Fanelli (eds.), *Small Geothermal Resources – A Guide to Development and Utilization*, UNITAR/UNDP Centre for Small Energy Resources, Rome, Italy, 31-59.

Juhasz, I. (1986), Assessment of the distribution of shale, porosity and hydrocarbon saturation in shaly sands. **In:** *Trans. Soc. Professional Well Log Analysts 10th European Formation Evaluation Symposium, Aberdeen, Scotland*, Ch. 15, paper AA.

Lee, K.C. (2001), Classification of geothermal resources by exergy, *Geothermics* **30**, 4, 431-442, DOI: 10.1016/S0375-6505(00)00056-0.

Leucci, G., S. Margiotta, S. Negri, L. Nuzzo, P. Sansò, G. Selleri, and A. Varola (2003), Integrated geophysical, geological and geomorphological investigations for study the impact of agricultural activities on a complex karstic area. **In:** *Proc. SAGEEP 2003, Environmental and Engineering Geophysical Society, 6-10 April 2003, Saint Antonio, USA*, 1162-1179.

Loke, M.H. (2011), Electrical imaging surveys for environmental and engineering studies. A practical guide to 2-D and 3-D surveys: RES2DINV Manual, IRIS Instruments, http://www.iris-instruments.com.

Loke, M.H., and R.D. Barker (1996), Rapid least-squares inversion of apparent resistivity pseudosections by a quasi-Newton method, *Geophys. Prospect*. **44**, 1, 131-152, DOI: 10.1111/j.1365-2478.1996.tb00142.x.

Lowrie, W. (2007), *Fundamentals of Geophysics*, Cambridge University Press, Cambridge.

Lund, J.W. (2007), Characteristics, development, and utilization of geothermal resources, *Geo-Heat Cent. Bull.* **28**, 2, 1-9.

Malanson, G.P. (1993), *Riparian Landscapes*, Cambridge Studies in Ecology, Cambridge University Press, Cambridge.

Malmström, V.H. (1969), A new approach to the classification of climate, *J. Geogr.* **68**, 6, 351-357, DOI: 10.1080/00221346908981131.

Margiotta, S. (1999), The contact between the formation of Galatone and formation of Lecce: stratigraphic and sedimentological evidence, *Atti Soc. Tosc. Sc. Nat. Mem. A* **106**, 73-77 (in Italian).

Margiotta, S., and S. Negri (2004), In search of water lost. New knowledge subsoil in Salento Lecce, Univ. degli Studi di Lecce (in Italian).

Margiotta, S., and G. Ricchetti (2002), Stratigraphy of oligomiocenici deposits of Salento (Puglia), *Boll. Soc. Geol. It.* **121**, 2, 243-252 (in Italian).

Marshall, T.J., and J.W. Holmes (1988), *Soil Physics*, 2nd ed., Cambridge University Press, New York, 374 pp.

MEA (2005), *Ecosystems and Human Well-Being: Wetlands and Water. Synthesis*, Millennium Ecosystem Assessment, World Resources Institute, Washington, D.C.

Meiser, P. (1962), A method of quantitative interpretation of selfpotential measurements, *Geophys. Prospect.* **10**, 2, 203-218, DOI: 10.1111/j.1365-2478.1962.tb02009.x.

Meyers, R.A. (ed.) (1992), *Encyclopedia of Physical Science and Technology*, Academic Press, San Diego.

Morris, B.L., A.R.L. Lawrence, P.J.C. Chilton, B. Adams, R.C. Calow, and B.A. Klinck (2003), Groundwater and its susceptibility to degradation: A global assessment of the problem and options for management, Early Warning and Assessment Report series, RS 03-3, United Nations Environment Programme, Nairobi, Kenya.

Muffler, P., and R. Cataldi (1978), Methods for regional assessment of geothermal resources, *Geothermics* **7**, 2-4, 53-89, DOI: 10.1016/0375-6505(78)90002-0.

Paul, M.K. (1965), Direct interpretation of self-potential anomalies caused by inclined sheets of infinite horizontal extensions, *Geophysics* **30**, 3, 418-423, DOI: 10.1190/1.1439596.

Perrier, F.E., G. Petiau, G. Clerc, V. Bogorodsky, E. Erkul, L. Jouniaux, D. Lesmes, J. Macnae, J.M. Meunier, D. Morgan, D. Nascimento, G. Oettinger, G. Schwarz, H. Toh, M.J. Valiant, K. Vozoff, and O. Yazici-Cakin (1997), A one-year systematic study of electrodes for long period measurements of the electric field in geophysical environments, *J. Geomagn. Geoelectr.* **49**, 11-12, 1677-1696, DOI: 10.5636/jgg.49.1677.

Pike, J.G. (1964), The estimation of annual run-off from meteorological data in a tropical climate, *J. Hydrol.* **2**, 2, 116-123, DOI: 10.1016/0022-1694(64)90022-8.

Reynolds, J.M. (1998), *An Introduction to Applied and Environmental Geophysics*, John Wiley & Sons Ltd., Chichester.

Shaw, E.M. (1994), *Hydrology in Practice*, 3rd ed., Chapman and Hall, London.

Sileo, M. (2011), Individuazione e caratterizzazione geologica, chimico-mineralogica e petrofisica di calcareniti tenere della Puglia e della Basilicata in relazione alle problematiche di provenienza e conservazione dei Beni Culturali, Ph.D. Thesis, University of Basilicata, Potenza, Italy (in Italian).

Telford, W.M., L.P. Geldart, and R.E. Sheriff (1990), *Applied Geophysics*, Cambridge University Press, Cambridge.

Vichabian, Y., and F.D. Morgan (2002), Self potentials in cave detection, *The Leading Edge* **21**, 9, 866-871, DOI: 10.1190/1.1508953.

Ward, R.C., and M. Robinson (1990), *Principles of Hydrology*, 3rd ed., McGraw–Hill Book Co., London.

Williamson, L., and N. McCormick (2008), Energy, ecosystems and livelihoods: understanding linkages in the face of climate change impacts, International Union for Conservation of Nature (IUCN), https://www.iucn.org/about/work/Initiatives/energy_welcome/index.cfm?uNewsID=1646.

Wu, J., and D.L. Nofziger (1999), Incorporating temperature effects on pesticide degradation into a management model, *J. Environ. Qual.* **28**, 1, 92-100, DOI: 10.2134/jeq1999.00472425002800010010x.

Zienkiewicz, O.C., and R.L. Taylor (1989), *The Finite Element Method: Basic Formulation and Linear Problems*, McGraw–Hill Book Co., London, 648 pp.

Received 23 August 2013
Received in revised form 9 February 2014
Accepted 18 March 2014

DE GRUYTER OPEN

Acta Geophysica
vol. 63, no. 1, Feb. 2015, pp. 154-175
DOI: 10.2478/s11600-014-0235-4

A Numerical Study of Effects of Valley-Weathering and Valley-Shape-Ratio on the Ground Motion Characteristics

Jay P. NARAYAN, Muhammad Y. ARAFAT, and KAMAL

Department of Earthquake Engineering, Indian Institute of Technology Roorkee, Roorkee, India; e-mail: jaypnfeq@iitr.ernet.in

A b s t r a c t

A study of combined effects of valley-weathering and valley-shape-ratio on the ground motion characteristics and associated differential ground motion (DGM) is documented in this paper. In order to properly quantify the weathering effects, a forth-order-accurate staggered-grid viscoelastic time-domain finite-difference program has been used for the simulation of *SH*-wave responses. Simulated results revealed that the de-focusing caused by valley is frequency-independent in contrast to the ridge-focusing. A decrease of average spectral amplification (ASA) with an increase of shape-ratio of the non-weathered triangular and elliptical valleys was obtained. Overall, the amplification and de-amplification pattern was larger in case of triangular valleys as compared to the elliptical valleys. It can be concluded that the dwelling within or near the top-corners of weathered valleys may suffer more damage as compared to their surroundings. A weathered triangular valley with large shape-ratio may cause unexpected damage very near its top-corners since both the ASA and DGM are largest.

Key words: viscoelastic time-domain response of valleys, finite-difference method, weathering and valley-shape-ratio effects, local site effects.

1. INTRODUCTION

This research work is stimulated with the changing structural scenario in the hilly areas. There is day by day increase of dwelling as well as the number of mega structures such as dams, bridges, and other life-lines across the valleys in the hilly areas. These mega structures may be subjected to non-uniform ground motions due to the spatial variability in ground motion caused by the valley topography. The earthquake records, post earthquake damage surveys, and simulations have revealed a ground motion amplification at the ridge top and de-amplification at the trough of a valley (Geli *et al.* 1988, Kawase and Aki 1990, Pedersen *et al.* 1994, Spudich *et al.* 1996, Narayan and Rai 2001, Narayan and Prasad Rao 2003, Narayan 2003, Kamalian *et al.* 2006). However, relatively less attention is given to the valley effects on the ground motion characteristics, may be due to the de-amplification of ground motion in the valleys. But theoretical studies have revealed that a valley may cause large ground motion amplification near its edges (Sextos *et al*. 2003, Zhou *et al.* 2010, Gao *et al.* 2012). The de-amplification and amplification patterns across a valley along with the geometry of the valley itself may cause intense differential ground motion (DGM). The DGM developed along the synclinal part of the valley may adversely affect the structures situated in valleys since it induces significant additional stresses in the structures as compared to the ones induced if the motions at the supports were considered to be identical. The extended structures across the valleys like dams, bridges, and other life-lines are largely affected by the DGM caused by the valley topography. So, the rapid development of long-span structures in the hilly areas calls for the quantification of amplification and de-amplification and DGM across the valley for the cost effective earthquake engineering.

In the past, seismic responses of homogeneous valleys of various shapes and size (triangular, U-shaped, trapezoidal, semi-sine, semi-circular, semi-elliptical, semi-cylindrical, *etc.*) have been simulated for the incident *SH*, *SV*, and *P* waves using different methods like Aki–Larner method, finite-difference (FD) method, finite-element methods, boundary element methods, and hybrid methods (Trifunac 1972, Wong 1982, Bouchon 1985, Hirai 1988, Sánchez-Sesma and Campillo 1991, Nguyen and Gatmiri 2007, Tsaur and Chang 2008, Gao *et al.* 2012). As far as we know, there are only few studies on the effects of valley-weathering on the ground motion characteristics. Zhao (2009, 2010) simulated the *P*- and *SV*-wave response of a weathered V-shaped valley for different weathering velocity and single shape-ratio using coupled method of finite and dynamic infinite-elements. Further, the weathering considered by Zhao (2010) was only along the slant part of the valley, although weathering thickness may be even larger on the horizontal/ gentle-slope part of the valley. It seems that almost no work is carried out to

quantify the combined effects of valley-weathering and valley-shape-ratio on the ground motion amplification and DGM across the valley.

This paper presents the combined effects of valley-shape, valley-shape-ratio, and the valley-weathering on the amplification of the *SH* wave as well as the DGM along the valley-flanks in details. It is verified whether valley de-amplification is frequency dependent or frequency independent. Snapshots at different moments are also computed to infer reflections from curved part of the valley, diffractions from the top corners of the valley, and the amplification patterns. A forth order accurate *SH*-wave time-domain viscoelastic FD program developed by Narayan and Kumar (2013) was used to simulate the responses of various considered models.

2. SALIENT ASPECTS OF THE USED FD COMPUTER PROGRAMS

A fourth order accurate *SH*-wave FD program developed by Narayan and Kumar (2013) is used for the simulation of responses of various valley models. This computer program is based on the staggered grid FD approximation of the viscoelastic *SH*-wave equations for the heterogeneous medium with a variable grid-size. The frequency-dependent damping in the time-domain FD simulations is incorporated based on the well known Generalized Maxwell Body – Emmerich and Korn (GMB-EK) rheological model (Emmerich and Korn 1987). A material independent anelastic function developed by Kristek and Moczo (2003) was used since it is preferable in case of material discontinuities in the FD grid (Narayan and Kumar 2013, 2014). The effective value of the unrelaxed modulus of rigidity μ_u and the density ρ at the desired motion in a grid cell is obtained using the harmonic and arithmetic means, respectively, to incorporate the material discontinuity (Moczo *et al.* 2002, Narayan and Kumar 2008). An improved vacuum formulation proposed by Zeng *et al.* (2012) is used as a free surface boundary condition. Zeng *et al.* (2012) reported that the improved vacuum formulation fully satisfies the free surface boundary condition by using an appropriate combination of the staggered-grid and a parameter averaging scheme proposed by Moczo *et al.* (2002). Both the sponge boundary condition of Israeli and Orszag (1981) and absorbing boundary condition of Clayton and Engquist (1977) were implemented on the model edges to avoid the edge reflections (Kumar and Narayan 2008).

3. VALLEY EFFECTS

It is well documented that the ground motion amplification caused by the ridge-focusing is frequency dependent (Geli *et al.* 1988). To find out whether valley-defocusing is frequency dependent, the *SH*-wave response of an elastic triangular valley (TRV) was computed using the incident horizon-

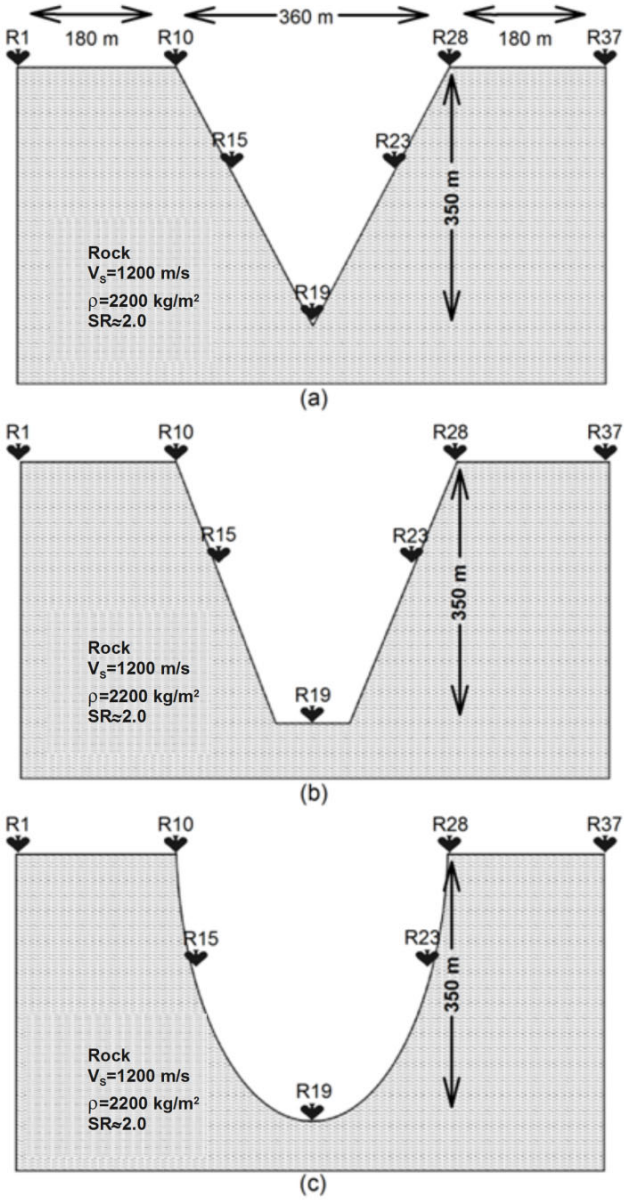


Fig. 1. Triangular (a), trapezoidal (b), and elliptical (c) valley models.

tal plane *SH*-wave front. Figure 1a shows a cross-section of a TRV model. The *S*-wave velocity and density for elastic rock are given in Table 1. The width and depth of the TRV model were taken as 360 and 350 m, respec-

Table 1

Rheological parameters for the elastic and viscoelstic rock and the visoelastic weathering material

Materials	Velocity and quality factor at F_R		Density [kg/m^3]	Unrelaxed modulus of rigidity [GPa]
	V_S [m/s]	Quality factor		
Elastic rock	1200	–	2200	3.168
Viscoelastic rock	1200	120	2200	3.240
Viscoelastic weathering	600	60	2000	0.733

tively. A plane horizontal *SH*-wave front was generated using various point sources at a depth of 0.50 km with respect to the trough of the TRV model. The shear stress σ_{ZY} in the form of a Ricker wavelet was used to generate the point sources along a horizon. The used Ricker wavelet with 4.0 Hz dominant frequency has a considerable spectral amplitude in the frequency band of 0.0 to 10.0 Hz. Time step and grid size were taken as 0.0005 s and 2 m, respectively. Seismic responses along the free surface were computed at 37 equidistant (20 m apart) receiver points extending from 360 m left to the 360 m right of the axis of the TRV model. Figure 2a depicts the *SH*-wave response of the elastic TRV model. The incident *SH* wave and diffracted *SH* wave are very clearly visible on the receiver points in the valley as well as at the free surface. Analysis of this figure shows that the *SH*-wave characteristics are highly affected by the presence of the valley. On an average, a decrease of amplitude of the incident *SH* wave towards the trough of valley can be inferred. The diffracted waves from the top corners of the valley and defocusing caused by the valley topography may be responsible for the amplitude variation along both the inclined and the horizontal flanks of the TRV model. Very large amplitude at the receivers' points very near the top-edges of valley may be due to interference of the incoming and diffracted waves. It means the response very near the top-edge of the valley cannot be used as a reference trace for the computation of topographic amplifications. In order to avoid the aforesaid problem, seismic response without considering the valley in the model was used for the computation of spectral amplification/de-amplification caused by the valley (Boore 1972, Lee *et al.* 2009). A comparison of the *SH*-wave response of a model without valley with the response at the tough of the elastic TRV model is shown in Fig. 2b. This figure clearly depicts the decrease of amplitude of incident *SH* wave (first arrival) due to the valley de-focusing. The diffracted *SH* waves from the top corners of the valley are also visible as the second arrival.

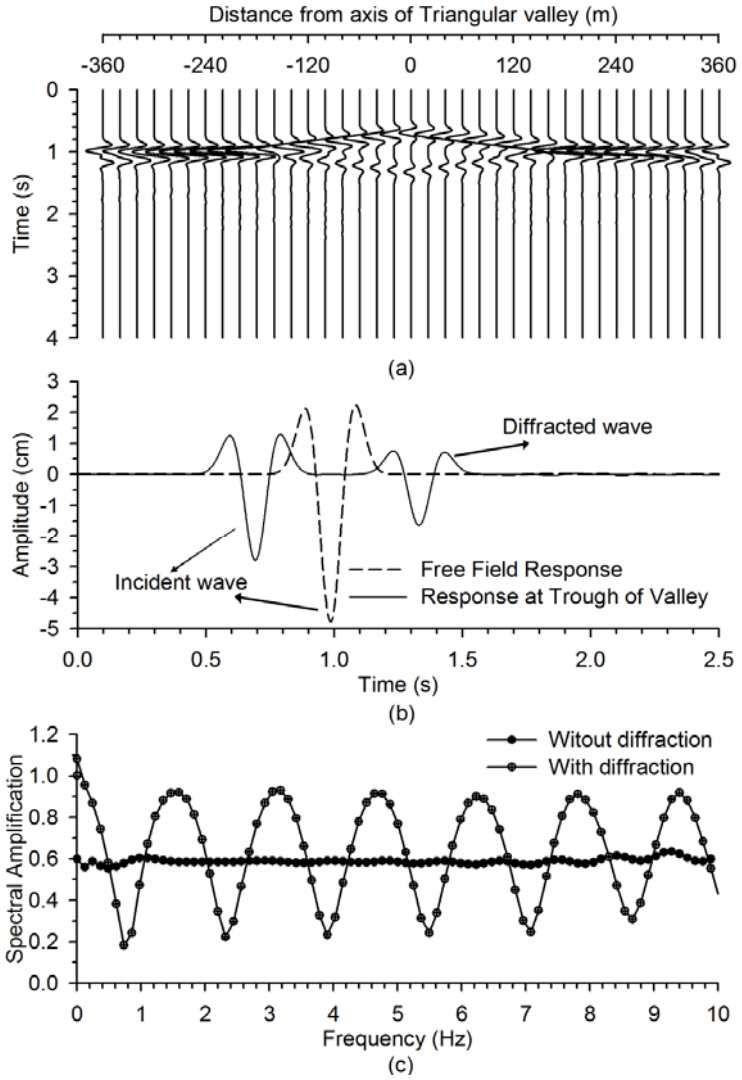


Fig. 2: (a) *SH*-wave response of elastic TRV model, (b) response at the trough of valley and free field, and (c) spectral amplification with and without diffracted waves.

In order to infer whether valley de-focusing is frequency dependent or not, the spectral amplification was computed at the trough of the valley without considering the diffracted *SH* wave. Before computing the FFT of the trace recorded on the trough of valley for computation of spectral amplification, the diffracted waves were removed manually. Figure 2c depicts the spectral amplification at the trough of the valley with and without consider-

ing the diffracted waves. The computed spectral amplification without considering the diffracted wave reveals that the valley de-focusing is frequency-independent in contrast to the ridge-focusing (Geli *et al.* 1988). On an average, the amplification factors for all the considered frequencies are more or less equal to 0.58 at the valley-trough. The analytically computed *SH*-wave

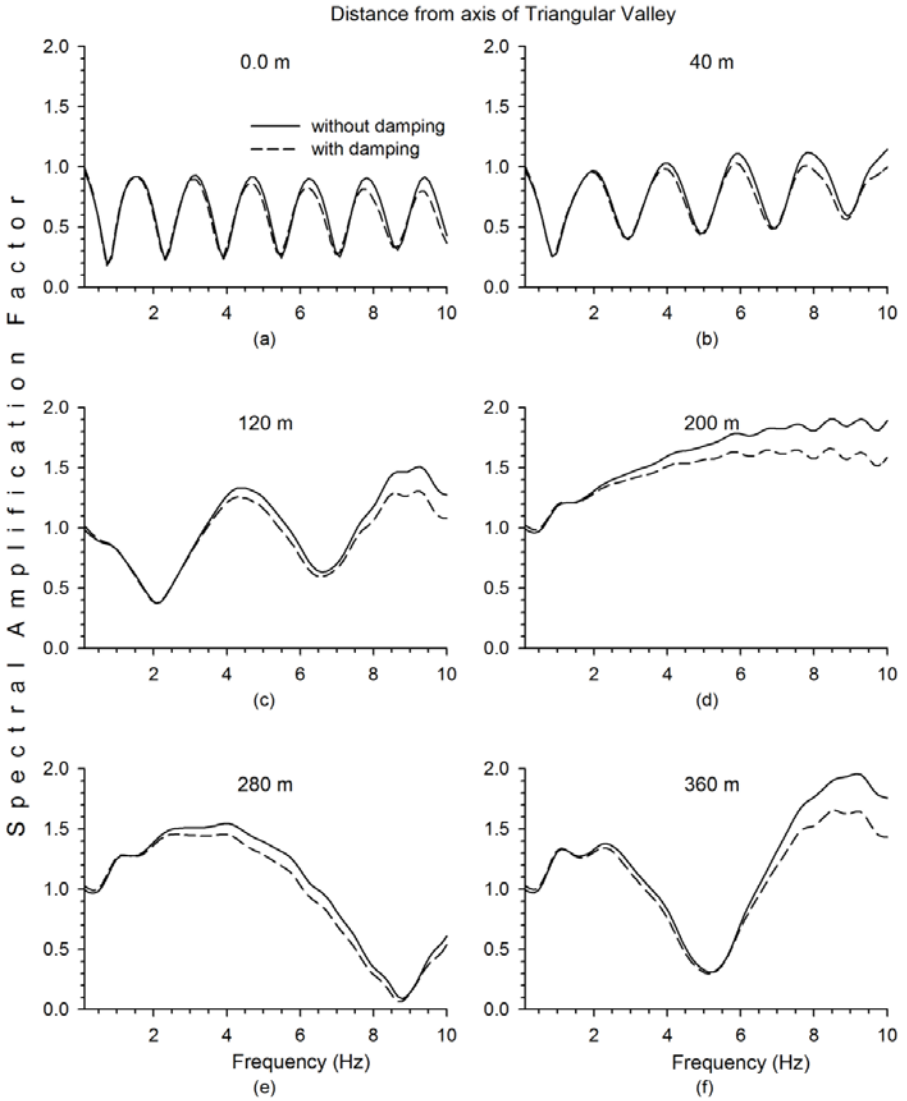


Fig. 3. A comparison of spectral amplification in case of elastic and viscoelastic responses of TRV model.

amplification at the trough of the triangular valley is also equal to 0.58 (Faccioli 1991). An excellent match of numerically computed amplification with the analytical one also reflects the accuracy of the used modified vacuum formulation as free surface boundary condition (Zeng *et al.* 2012). On the other hand, the spectral amplification is highly varying with frequency if diffracted waves are considered (Fig. 2c). The frequency independent defocusing was also reported by Kumar and Narayan (2013) based on the analysis of response of an anticlinal unbounded basement topography. So, it can be concluded that the de-amplification caused by the valley-de-focusing is frequency independent in contrast to the amplification caused by the ridge-focusing (Geli *et al.* 1988).

In order to infer the material damping on the spectral amplification, seismic responses of the viscoelastic valley model were also computed. The phase velocity and quality factor at a reference frequency, density and unrelaxed modulus of rigidity in the viscoelastic rock are given in Table 1. A comparison of spectral amplification at different location along the cross-section of the valley in case of elastic and viscoelastic response is shown in Fig. 3. The analysis of Figs. 2c and 3 reveals that the cause of ups and downs in spectral amplification pattern is the presence of diffracted waves and its effect if decreasing with an increase of distance from the top-edges of the valley. However, it is surprising to note that the spectral amplification is consistently increasing with an increase of frequency in case of elastic response very near the top-edges of the valley (Fig. 3d). The maximum amplification in case of viscoelastic response is 1.5 at 10 Hz.

4. *SH*-WAVE SNAPSHOTS

In order to demonstrate the behaviour of incident *SH* wave along the flanks, bottom and top edges of the elliptical valley, snapshots were computed at different times. Snapshots were computed in a rectangular area extending 20 m above to 480 m below the free surface 540 m south to 540 m north of axis of elliptical valley. The snapshots at times 0.54 and 0.62 s (Fig. 4a, b) depict that the *SH*-wave front has entered in the considered rectangular area and the outline of the valley is being reflected. The propagation of incident *SH* wave towards the free surface can also be inferred. Figure 4c-f shows the splitting of the incident *SH*-wave front, the reflection from the flanks of the valley and the diffraction at the trough. Further, the reflected *SH* waves from the flanks of the valley are not propagating towards the free surface. Figure 4g clearly depicts the amplitude build-up very near the top corners of the valley. The reflected *SH* wave from the free surface and the diffracted *SH* wave from the top edges of elliptical valley and their propagation can be very clearly inferred from Fig. 4h-l.

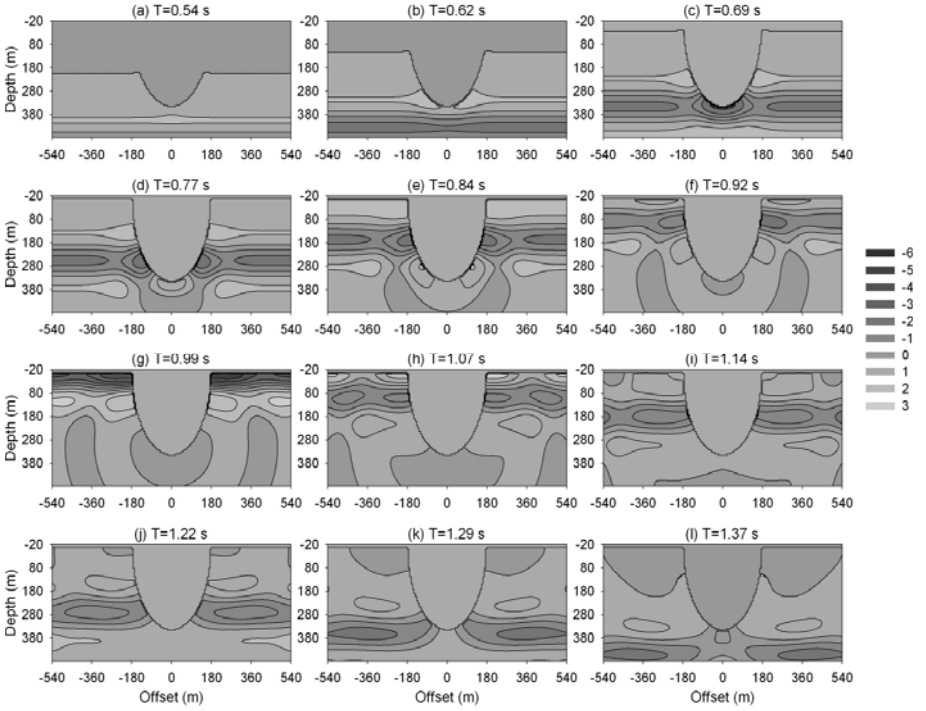


Fig. 4. *SH*-wave snapshots for elliptical valley at different moments (note: amplitude variation in scale is in cm).

5. EFFECTS OF VALLEY SHAPE

To study the effects of shape of the valley on the amplification pattern, *SH*-wave responses of viscoelastic triangular valley (TRV), trapezoidal valley (TPV), and an elliptical valley (ELV) models were computed. The shape-ratio for the viscoelastic TRV, TPV, and ELV models was the same as given in Fig. 1. The shape-ratio is the ratio of depth of valley with the half-width of the valley-mouth. The *S*-wave velocity and quality factor at a reference frequency (1.0 Hz), density and unrelaxed modulus of rigidity are given in Table 1. Further, relaxation frequencies were taken as 0.02, 0.2, 2.0, and 20.0 Hz for the computation of anelastic coefficients and anelastic functions (Narayan and Kumar 2013). The other model parameters and source receiver configuration were the same as in the previous case. Figure 5a-c shows the seismic responses of the TRV, TPV, and ELV models, respectively. Analysis of Fig. 5 reveals that both the de-focusing effect at the valley-troughs and diffractions from the top corners are greater in case of the TRV model and least in case of the ELV model. The seismic responses of the TRV and TPV models seem to be almost similar, except of the amplitude de-amplification

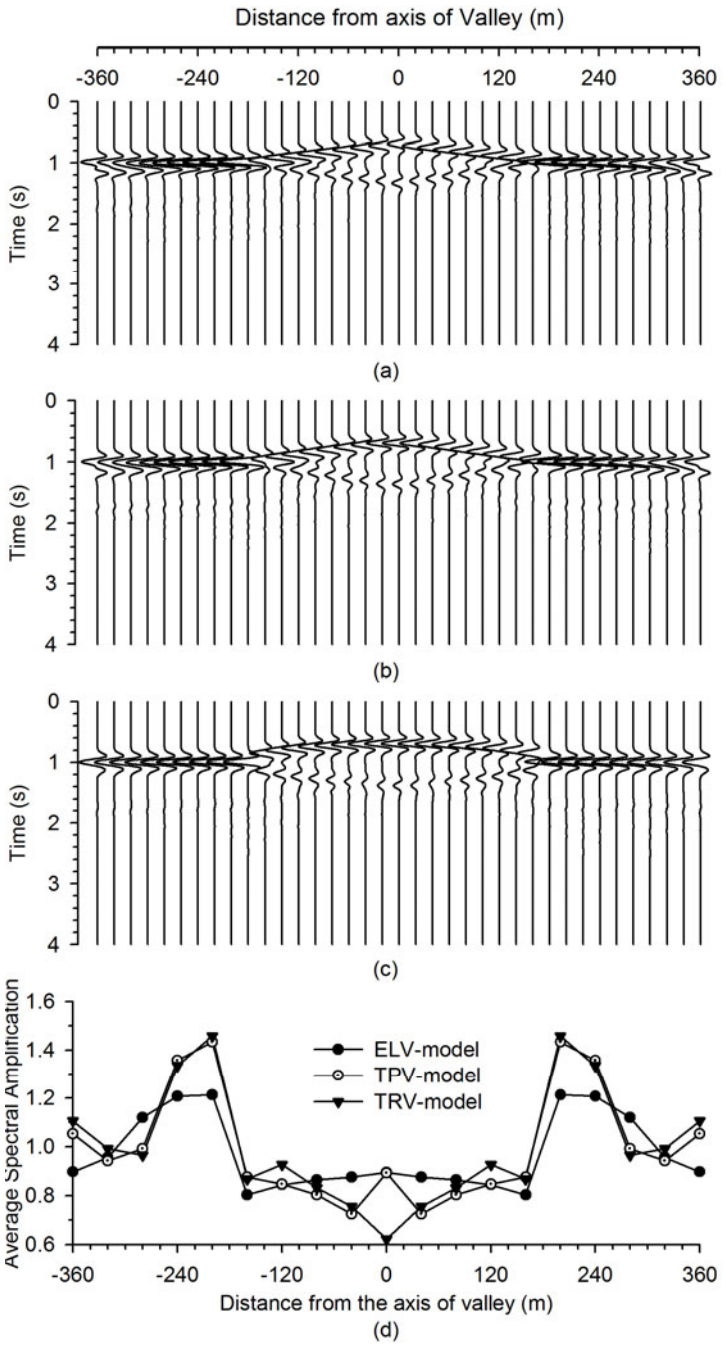


Fig. 5. *SH*-wave response of: triangular (a), trapezoidal (b), and elliptical (c) valleys; and comparison of ASA caused be considered valleys (d).

along the flanks of the valleys. Figure 5d depicts a comparison of average spectral amplification (ASA) caused by the TRV, TPV, and ELV models. The ASA at troughs of the TRV, TPV, and ELV models are 0.62, 0.9 and 0.9, respectively. Similarly, the largest value of ASA very near the top edges (at an offset of 20 m) of the TRV, TPV, and ELV models are 1.45, 1.43, and 1.21, respectively. In case of ELV model, the ASA is larger at the trough of the valley as compared to the ASA along the slant part of the valley. Even, there is minor decrease of ASA from trough towards the top of the valley. The TRV model with a least surface area has the largest amplification at the top-edges and least amplification at the trough. It is also inferred that for an equal shape-ratio, the valley de-amplification pattern along the synclinal part is very much dependent on the shape of the valley. It may be concluded that for an equal shape-ratio, the valley de-amplification at the trough and amplification at the top-corners are larger for the valley whose surface area is less.

6. EFFECTS OF SHAPE-RATIO OF NON-WEATHERED VALLEY

In order to study the effects of valley shape-ratio on the ground motion characteristics, *SH*-wave responses of non-weathered triangular valley STRV1-STRV4 and elliptical valley SELV1-SELV4 models with different shape-ratio are computed. Both triangular and elliptical valleys have a fixed depth of 350 m and varying widths as 234, 180, 140, and 116 m, respectively. So, valley-shape-ratios for the STRV1-STRV4 or SELV1-SELV4 models are 1.49, 1.94, 2.50, and 3.02, respectively. The *SH*-wave responses of the non-weathered STRV1-STRV4 models (left panel) and SELV1-SELV4 models (right panel) are shown in Fig. 6. An increase of valley de-focusing at the trough and diffractions from the top-edges of the valleys can be inferred with an increase of shape-ratio. The amplitude of diffracted *SH* waves is less in the SELV models as compared to the respective STRV models.

6.1 Spatial variation of ASA for the non-weathered valley

Figure 7a, b shows a comparison of spatial variation of ASA with the shape-ratio for the non-weathered triangular and elliptical valleys, respectively. An analysis of this figure very clearly reveals a decrease of ASA (increase of de-amplification) with an increase of shape-ratio in both the valley models for the considered models parameters and the frequency range. The ASA is least at the trough and largest at the top-corners of both the valley models. Furthermore, the obtained largest amplification near the top-corners of both the valley models is increasing with decrease of shape ratio. However, the ASA in the synclinal part of the elliptical valley is more or less same everywhere in contrast to the triangular valleys, where there is an increase of amplifica-

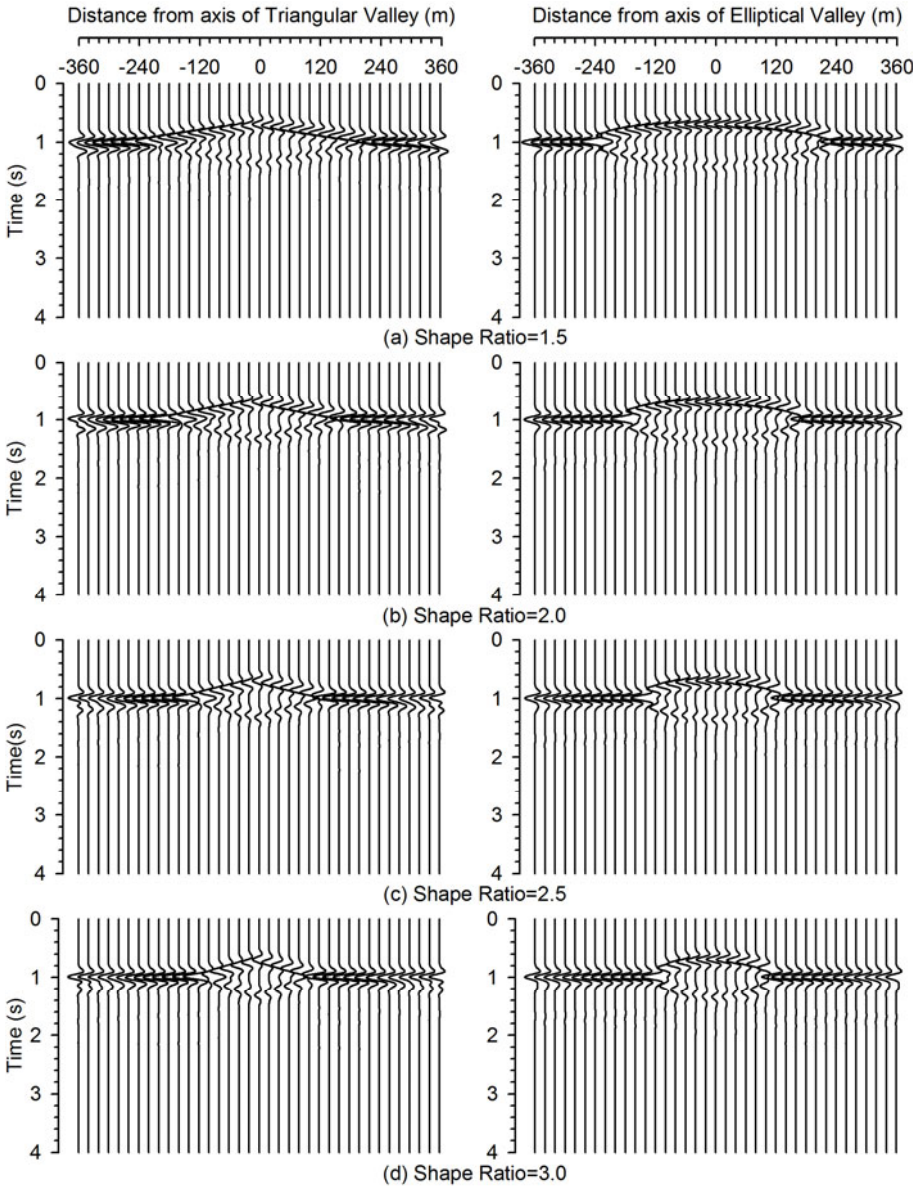


Fig. 6. *SH*-wave response of triangular and elliptical valleys with different shape ratio.

tion towards the top-corners. Finally, it can be concluded that the de-amplification at the trough and amplification at the top-corners of the STRV models are larger than that caused by the SELV models.

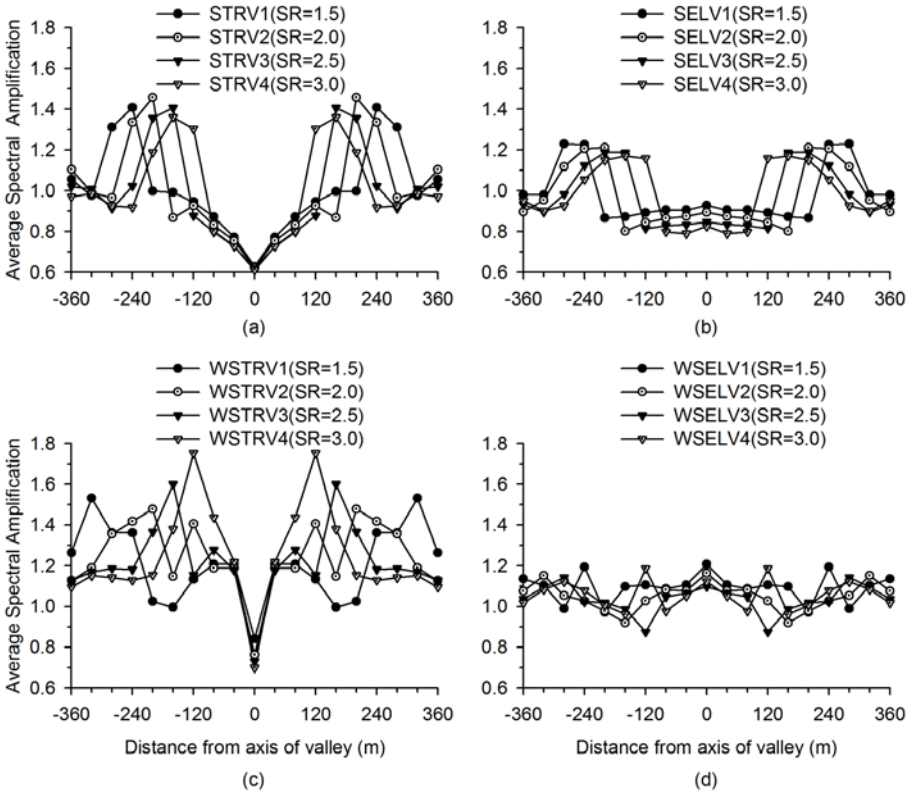


Fig. 7. Spatial variation of average spectral amplification of: STRV (a), SELV (b), WSTRV (c), and WSELV (d) models.

6.2 Spatial variation of DGM for the non-weathered valley

To quantify the effects of the shape-ratio on the DGM developed by the *SH* wave along the non-weathered flanks of the STRV and SELV models, responses were computed on another array with 29 equidistant (10 m apart horizontally) receiver points extending from 0 to 280 m leftwards of the valley-axis. Figure 8 shows the STRV model with the positions of source and receiver array. Further, before computing the DGM, all the traces were normalized with the maximum amplitude in the response of a model with no valley in the model, so that the computed DGM corresponds to the DGM caused by unit amplitude of the incident *SH* wave. The distances between two consecutive receiver points along the valley-flanks were computed considering the shapes of the various considered valley-models. Figure 9a-d shows the computed DGM for the STRV1-STRV4 models (left panel) and SELV1-SELV4 models (right panel), respectively. Near the edges of the val-

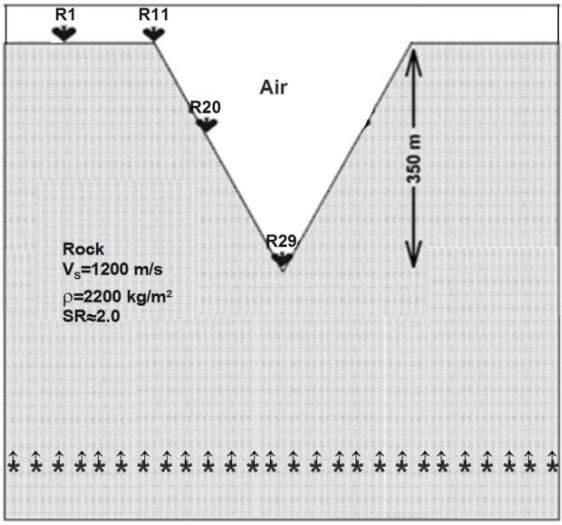


Fig. 8. Triangular valley model with the position of point sources to generate a plane wave front, air above the free surface, and a new array with 29 equidistant (10 m apart) receiver points.

leys there is sudden increase and decrease of DGM. Further, DGM can also be inferred even on the horizontal flanks of the valleys caused by the diffracted waves. Figure 10a, b depicts the spatial variation of the peak DGM developed by the *SH* wave along the flanks of the non-weathered triangular and elliptical valleys. The peak DGM at the trough of the triangular STRV1-STRV4 valley models are of the order of 9.87×10^{-5}, 1.11×10^{-4}, 1.13×10^{-4}, and 1.08×10^{-4}, respectively for 1.0 cm amplitude in the incident *SH* wave. Similarly, the peak DGM at the trough of the elliptical SELV1-SELV4 valley models is of the order of 1.88×10^{-5}, 2.44×10^{-5}, 4.0×10^{-5}, and 4.0×10^{-5}, respectively. Further, the largest DGMs caused by the triangular STRV1-STRV4 valleys are of the order of 2.92×10^{-4}, 2.89×10^{-4}, 2.58×10^{-4}, and 2.56×10^{-4}, respectively, very near to the top of the valleys. Similarly, the largest DGMs caused by the elliptical SELV1-SELV4 valleys are of the order of 2.29×10^{-4}, 1.82×10^{-4}, 1.8×10^{-4}, and 1.48×10^{-4}, respectively, very near to the top of the valley. Analysis of Fig. 10a, b depicts that the DGM is increasing with shape-ratio at the trough and decreasing with shape-ratio at the top-corners of the considered valleys of both types. There may be three possible reasons behind the sudden increase and decrease of DGM very near the edges of the valleys. The first one is the longer offset along the slant part of the valley as compared to the horizontal flanks. The second reason is the sudden increase of ground motion near the top corners of the valley (around

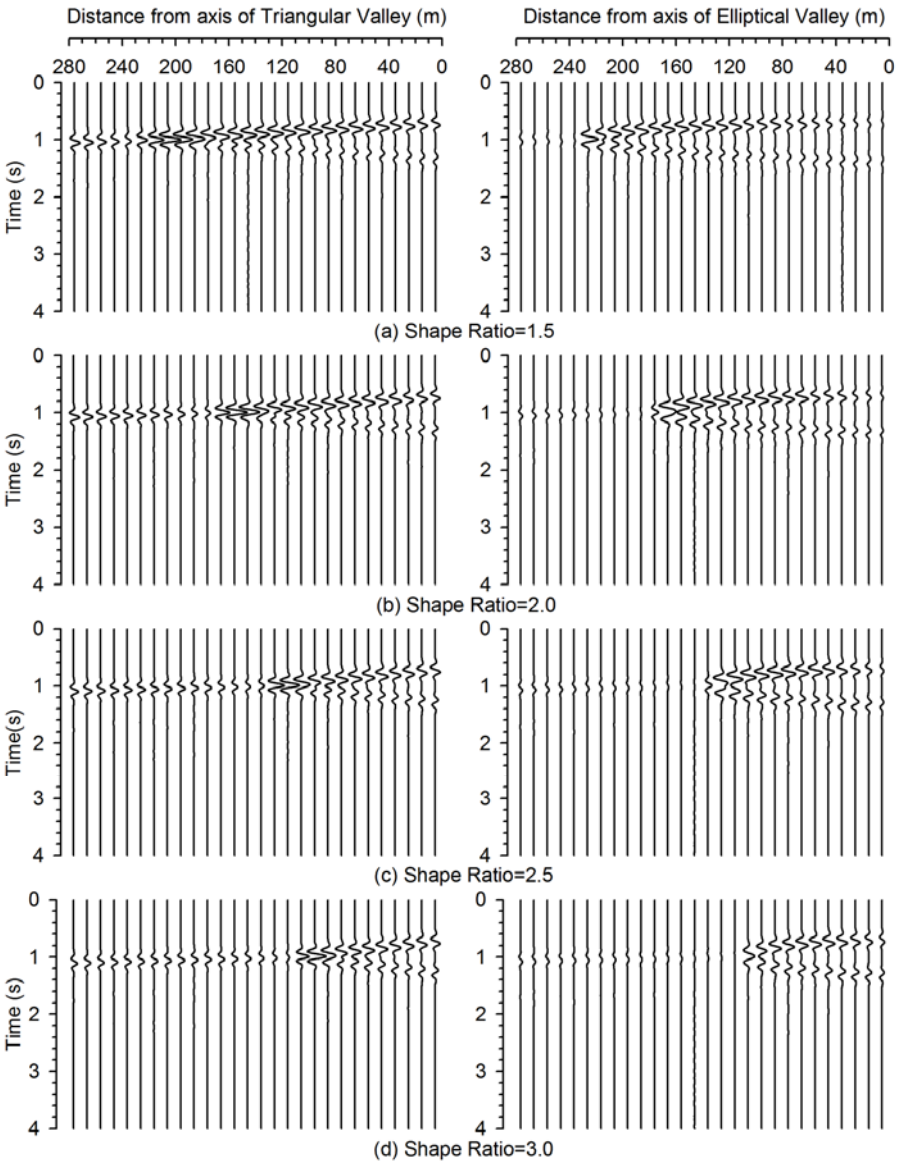


Fig. 9. Differential ground motion (DGM) for different STRV and SELV models.

20 m away from the top corners). The third reason may be the more or less simultaneous arrival of the incident plane wave front on the receiver point on the horizontal flanks and different arrival times on the slant part of the valley. Overall, the DGM developed by the triangular valleys are larger than that developed by the elliptical valleys.

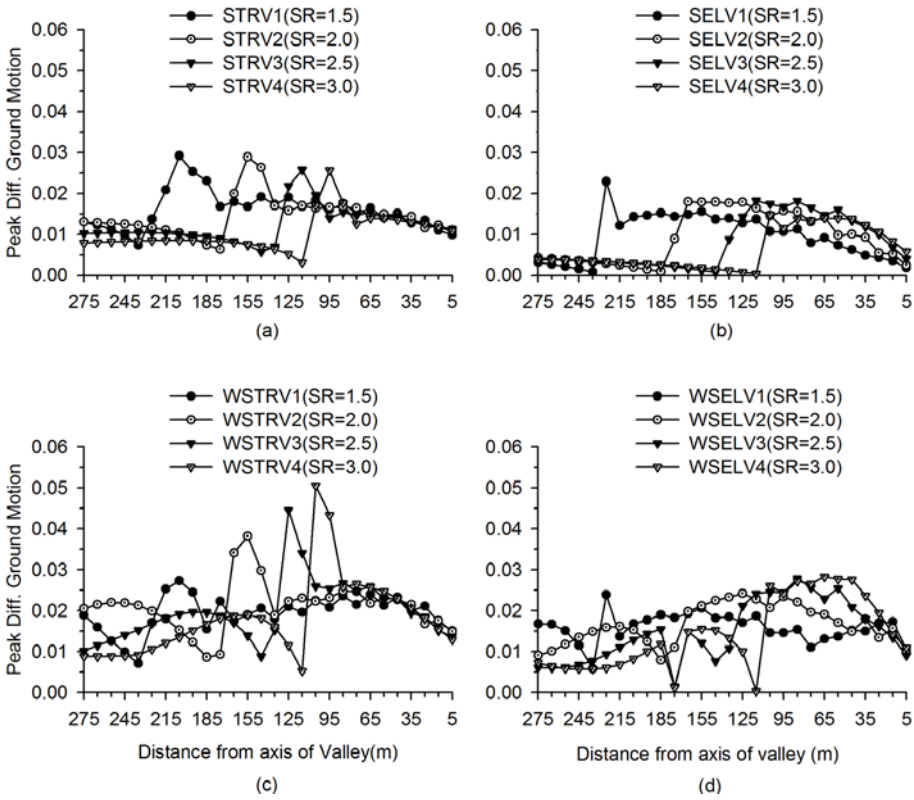


Fig. 10. Spatial variation of peak DGM for: STRV (a), SELV (b), WSTRV (c), and WSELV (d) models.

7. EFFECTS OF THE SHAPE-RATIO OF WEATHERED VALLEY

In order to study the combined effects of valley-weathering and the shape-ratio on the ground motion characteristics, *SH*-wave responses of both the triangular weathered WSTRV1-WSTRV4 models and elliptical weathered WSELV1-WSELV4 valley models were simulated. All the weathered-valley models have weathering thickness of 30 m on both the slant and the horizontal flanks. Further, the shape-ratios of the WSTRV and WSELV models were same as the corresponding STRV and SELV models, respectively. The *S*-wave velocity and quality factor at a reference frequency (1.0 Hz), density, and unrelaxed modulus of rigidity for both the viscoelastic weathering and rock are given in Table 1. In order to quantify the combined effects of shape-ratio and valley-weathering, seismic responses without considering the valley but a weathering of thickness 30 m were also simulated.

7.1 Spatial variation of ASA for the weathered valley

Figure 7c, d shows a comparison of spatial variation of ASA with the shape-ratio for the weathered triangular and elliptical valleys, respectively (please note that the amplification is computed with respect to a model containing a horizontal weathered layer over the half-space). An analysis of these figures reveals a very complex ASA pattern with shape-ratio. On an average, there is amplification of ground motion everywhere in both the weathered triangular and elliptical valleys, except at the trough of the triangular valleys. A comparison of amplification patterns of weathered and non-weathered triangular valleys depicts that the amplification on the horizontal flanks of the weathered valley is greater than the non-weathered valley. An increase of amplification with shape-ratio near the top-corners of the triangular valley can be inferred in contrast to the non-weathered triangular valleys. On the other hand, amplification levels on the horizontal flanks of the weathered and non-weathered elliptical valleys are comparable. So, it may be concluded that weathering has caused amplification of ground motion along the synclinal part of both the triangular and elliptical valleys, except at the trough of triangular valley. Further, a considerable amplification on the horizontal flanks of both valleys was obtained. In case of triangular valleys, ASA was even larger than the non-weathered triangular valleys.

7.2 Spatial variation of DGM for the weathered valley

Figure 11a-d shows the DGM developed by the *SH* wave in case of the weathered WSTRV1-WSTRV4 models (left panel) and weathered WSELV1-WSELV4 models, respectively. Analysis of Fig. 10 reveals an increase of DGM with an increase of shape-ratio on the slant flanks of both the valleys. In contrast, a decrease of DGM on the horizontal flanks of both types of valleys with an increase of shape-ratio can be inferred. It appears that in case of weathered elliptical valleys trapping of waves in the weathering has developed Love waves, which is very clear in case of WSELV1 model. Figure 10c, d depicts a comparison of the spatial variation of the peak DGM for the weathered WSTRV and WSELV models, respectively. Analysis of these figures reflects an increase of peak DGM in both the WSTRV and WSELV valleys with an increase of shape-ratio and is larger than that caused by the corresponding non-weathered valley models (Fig. 10a, b). The large peaks and troughs in peak DGM near the top corners may be due to valley shape-effect. The peak DGM in case of WSELV models is comparable to that in case of WSTRV models for larger shape-ratio, but lesser in case of small shape-ratio.

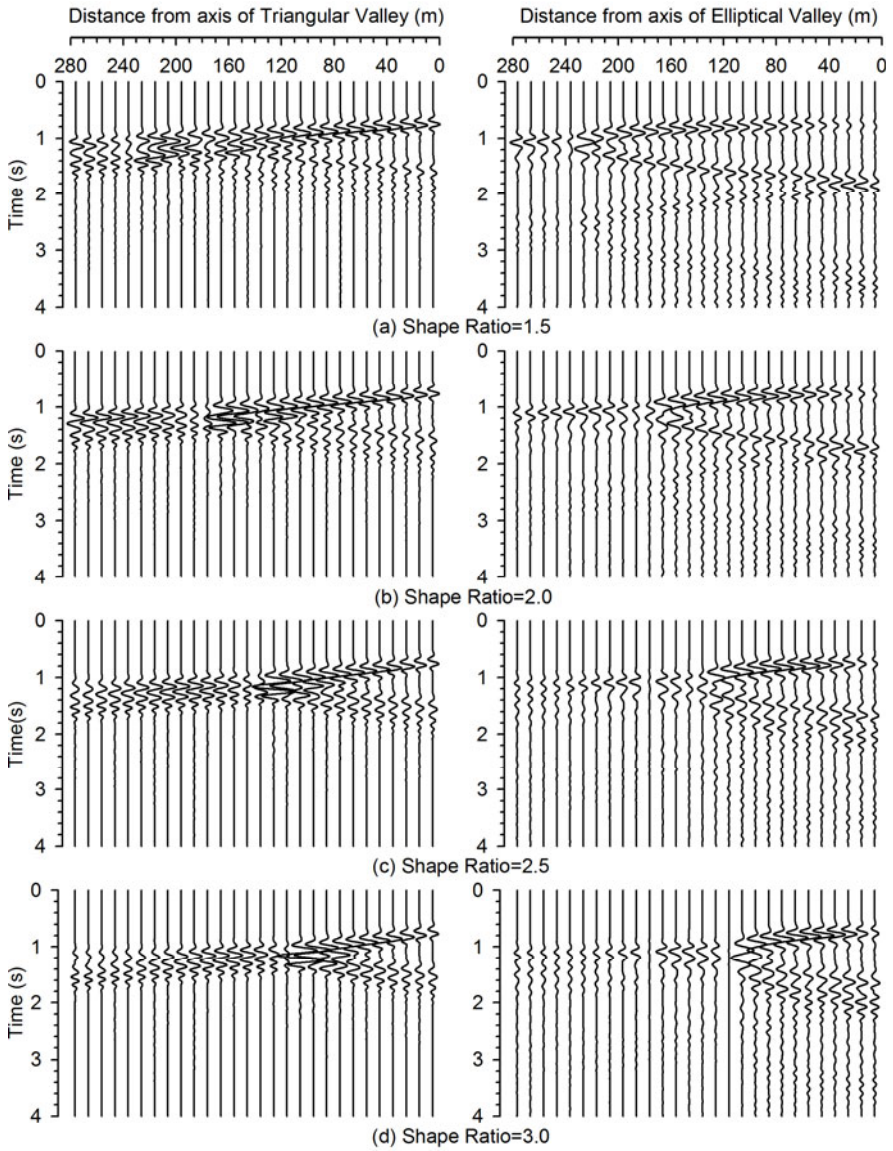


Fig. 11. Differential Ground Motion (DGM) for different WSTRV and WSELV models with 30 m weathering thickness.

8. CONCLUSIONS

Based on the comparison of simulated result with analytical one (Faccioli 1991), it is concluded that the de-focusing caused by the valley is frequency-independent, in contrast to the focusing caused by ridge/hill, which is fre-

quency dependent (Geli *et al.* 1988). Furthermore, the valley causes amplification of all the frequencies very near its top-corners. It is also inferred that for an equal shape-ratio, the valley de-amplification at the trough and amplification at the top-corners are larger for the valley whose surface area is smaller. A decrease of the spatial variations of the average spectral amplification (ASA) for both the non-weathered triangular and elliptical valleys was obtained with an increase of the shape-ratio. There is minor decrease of ASA from trough to the top of the non-weathered elliptical valleys, in contrast to the non-weathered triangular valleys. Both de-amplification at the trough and amplification near the top-corners is larger in case of the non-weathered triangular valleys as compared to the non-weathered elliptical valleys. It may be concluded that weathering has caused amplification of ground motion along the synclinal part of both the triangular and elliptical valleys, except at the trough of triangular valley. Further, a considerable amplification on the horizontal flanks of both the valleys was obtained. In case of triangular valleys, ASA was even larger than in the non-weathered triangular valleys. A weathered triangular valley with large shape-ratio may cause unexpected damage very near its top-corners since both the ASA and DGM are the largest.

Acknowledgement. The first author is grateful to the Council of Scientific and Industrial Research, New Delhi, for the financial assistance through Grant No. CSR-569-EQD.

References

Boore, D.M. (1972), Finite difference methods for seismic wave propagation in heterogeneous materials. **In:** B.A. Bolt, B. Adler, S. Fernbach, and M. Rotenberg (eds.), *Methods in Computational Physics*, Seismology: Surface Waves and Earth Oscillations, Vol. 11, Academic Press, New York, 1-37.

Bouchon, M. (1985), A simple, complete numerical solution to the problem of diffraction of SH-waves by an irregular surface, *J. Acoustic. Soc. Am.* **77**, 1, 1-5, DOI: 10.1121/1.392258.

Clayton, R., and B. Engquist (1977), Absorbing boundary conditions for acoustic and elastic wave equations, *Bull. Seismol. Soc. Am.* **67**, 6, 1529-1540.

Emmerich, H., and M. Korn (1987), Incorporation of attenuation into time-domain computations of seismic wave fields, *Geophysics* **52**, 9, 1252-1264, DOI: 10.1190/1.1442386.

Faccioli, E. (1991), Seismic amplification in the presence of geological and topographic irregularities. **In:** *Proc. 2nd Int. Conf. on Recent Advances in Geo-*

technical Earthquake Engineering and Soil Dynamics, 11-15 March 1991, St. Louis, Missouri, USA, 1779-1797.

Gao, Y., N. Zhang, D. Li, H. Liu, Y. Cai, and Y. Wu (2012), Effects of topographic amplification induced by U-shaped canyon on seismic waves, *Bull. Seismol. Soc. Am.* **102**, 4, 1748-1763, DOI: 10.1785/0120110306.

Geli, L., P.Y. Bard, and B. Jullien (1988), The effect of topography on earthquake ground motion: A review and new results, *Bull. Seismol. Soc. Am.* **78**, 1, 42-63.

Hirai, H. (1988), Analysis of transient response of SH wave scattering in half space by the boundary element method, *Eng. Anal.* **5**, 4, 189-194, DOI: 10.1016/0264-682X(88)90015-9.

Israeli, M., and S.A. Orszag (1981), Approximation of radiation boundary conditions, *J. Comput. Phys.* **41**, 1, 115-135, DOI: 10.1016/0021-9991(81)90082-6.

Kamalian, M., M.K. Jafari, A. Sohrabi-Bidar, A. Razmkhah, and B. Gatmiri (2006), Time-domain two-dimensional site response analysis of non-homogeneous topographic structures by a hybrid BE/FE method, *Soil Dyn. Earthq. Eng.* **26**, 8, 753-765, DOI: 10.1016/j.soildyn.2005.12.008.

Kawase, H., and K. Aki (1990), Topography effect at the critical SV-wave incidence: Possible explanation of damage pattern by the Whittier Narrows, California, earthquake of 1 October 1987, *Bull. Seismol. Soc. Am.* **80**, 1, 1-22.

Kristek, J., and P. Moczo (2003), Seismic-wave propagation in viscoelastic media with material discontinuities: A 3D fourth-order staggered-grid finite-difference modeling, *Bull. Seismol. Soc. Am.* **93**, 5, 2273-2280, DOI: 10.1785/0120030023.

Kumar, S., and J.P. Narayan (2008), Absorbing boundary conditions in a fourth-order accurate SH-wave staggered grid finite difference algorithm, *Acta Geophys.* **56**, 4, 1090-1108, DOI: 10.2478/s11600-008-0043-9.

Kumar, V., and J.P. Narayan (2013), Study of combined effects of sediment rheology and anticlinal basement topography on ground motion characteristics, *Geofizika* **30**, 1, 75-93.

Lee, S.J., D. Komatitsch, B.S. Huang, and J. Tromp (2009), Effects of topography on seismic-wave propagation: An example from northern Taiwan, *Bull. Seismol. Soc. Am.* **99**, 1, 314-325, DOI: 10.1785/0120080020.

Moczo, P., J. Kristek, V. Vavryčuk, R.J. Archuleta, and L. Halada (2002), 3D heterogeneous staggered-grid finite-difference modeling of seismic motion with volume harmonic and arithmetic averaging of elastic moduli and densities, *Bull. Seismol. Soc. Am.* **92**, 8, 3042-3066, DOI: 10.1785/0120010167.

Narayan, J.P. (2003), Simulation of ridge-weathering effects on the ground motion characteristics, *J. Earthq. Eng.* **7**, 3, 447-461, DOI: 10.1080/13632460309350458.

Narayan, J.P., and S. Kumar (2008), A fourth order accurate SH-wave staggered grid finite-difference algorithm with variable grid size and VGR-stress imaging technique, *Pure Appl. Geophys.* **165**, 2, 271-294, DOI: 10.1007/s00024-008-0298-8.

Narayan, J.P., and V. Kumar (2013), A fourth-order accurate finite-difference program for the simulation of SH-wave propagation in heterogeneous viscoelastic medium, *Geofizika* **30**, 2, 173-189.

Narayan, J.P., and V. Kumar (2014), Study of combined effects of sediment rheology and basement focusing in an unbounded viscoelastic medium using P-SV-wave finite-difference modelling, *Acta Geophys.*, DOI: 10.2478/s11600-013-0199-9.

Narayan, J.P., and P.V. Prasad Rao (2003), Two and half dimensional simulation of ridge effects on the ground motion characteristics, *Pure Appl. Geophys.* **160**, 8, 1557-1571, DOI: 10.1007/s00024-003-2360-x.

Narayan, J.P., and D.C. Rai (2001), An observational study of local site effects in Chamoli earthquake. **In:** *Proc. Workshop on Recent Earthquakes of Chamoli and Bhuj, 22-23 May 2001, Roorkee, India*, Indian Society of Earthquake Technology, 273-280.

Nguyen, K.V., and B. Gatmiri (2007), Evaluation of seismic ground motion induced by topographic irregularity, *Soil Dyn. Earthq. Eng.* **27**, 2, 183-188, DOI: 10.1016/j.soildyn.2006.06.005.

Pedersen, H., B. Le Brun, D. Hatzfeld, M. Campillo, and P.Y. Bard (1994), Ground-motion amplitude across ridges, *Bull. Seismol. Soc. Am.* **84**, 6, 1786-1800.

Sánchez-Sesma, F.J., and M. Campillo (1991), Diffraction of P, SV, and Rayleigh waves by topographic features: A boundary integral formulation, *Bull. Seismol. Soc. Am.* **81**, 6, 2234-2253.

Sextos, A.G., A.J. Kappos, and K.D. Pitilakis (2003), Inelastic dynamic analysis of RC bridges accounting for spatial variability of ground motion, site effects and soil-structure interaction phenomena. Part 2: Parametric study, *Earthq. Eng. Struct. Dyn*. **32**, 4, 629-652, DOI: 10.1002/eqe.242.

Spudich, P., M. Hellweg, and W.H.K. Lee (1996), Directional topographic site response at Tarzana observed in aftershocks of the 1994 Northridge, California, earthquake: Implications for mainshock motions, *Bull. Seismol. Soc. Am.* **86**, 1B, 193-208.

Trifunac, M.D. (1972), Scattering of plane SH waves by a semi-cylindrical canyon, *Earthq. Eng. Struct. Dyn.* **1**, 3, 267-281, DOI: 10.1002/eqe.4290010307.

Tsaur, D.H., and K.H. Chang (2008), An analytical approach for the scattering of SH waves by a symmetrical V-shaped canyon: Shallow case, *Geophys. J. Int.* **174**, 1, 255-264, DOI: 10.1111/j.1365-246X.2008.03788.x.

Wong, H.L. (1982), Effect of surface topography on the diffraction of P, SV, and Rayleigh waves, *Bull. Seismol. Soc. Am.* **72**, 4, 1167-1183.

Zeng, C., J. Xia, R. Miller, and G. Tsoflias (2012), An improved vacuum formulation for 2D finite-difference modeling of Rayleigh waves including surface topography and internal discontinuities, *Geophysics* **77**, 1, T1-T9, DOI: 10.1190/geo2011-0067.1.

Zhao, C. (2009), *Dynamic and Transient Infinite Elements: Theory and Geophysical, Geotechnical and Geoenvironmental Applications*, Advances in Geophysical and Environmental Mechanics and Mathematics, Springer, Dordrecht.

Zhao, C. (2010), Coupled method of finite and dynamic infinite elements for simulating wave propagation in elastic solids involving infinite domains, *Sci. China Technol. Sci*. **53**, 6, 1678-1687, DOI: 10.1007/s11431-010-3205-3.

Zhou, G., X. Li., and X. Qi (2010), Seismic response analysis of continuous rigid frame bridge considering canyon topography effects under incident SV waves, *Earthq. Sci.* **23**, 1, 53-61, DOI: 10.1007/s11589-009-0065-7.

Received 7 November 2013
Received in revised form 29 January 2014
Accepted 18 March 2014

DE GRUYTER OPEN

Acta Geophysica
vol. 63, no. 1, Feb. 2015, pp. 176-200
DOI: 10.2478/s11600-014-0249-y

Mining Induced Seismic Event on an Inactive Fault

Grzegorz LIZUREK, Łukasz RUDZIŃSKI and Beata PLESIEWICZ

Institute of Geophysics, Polish Academy of Sciences, Warszawa, Poland
e-mail: lizurek@igf.edu.pl

A b s t r a c t

On 19 March 2013, a tremor shook the surface of Polkowice town where the Rudna Mine is located. This event, of $ML = 4.2$, was the third most powerful seismic event recorded in the Legnica Głogów Copper District (LGCD). Inhabitants of the area reported that the felt tremor was bigger and lasted longer than any other ones felt in the last couple of years. Analysis of spectral parameters of the records from in-mine seismic system and surface LUMINEOS network along with broadband station KSP record were carried out. The location of the event was close to the Rudna Główna Fault zone; the nodal planes orientations determined with two different approaches were almost parallel to the strike of the fault. The mechanism solutions were also obtained as Full Moment Tensor from *P*-wave amplitude pulses of underground records and waveform inversion of surface network seismograms. The results from the seismic analysis along with macroseismic survey and observed effects from the destroyed part of the mining panel indicate that the mechanism of the event was complex rupture initiated as thrust faulting on an inactive tectonic normal fault zone. The results confirm that the fault zones are the areas of higher risk, even in case of carefully taken mining operations.

Key words: focal mechanism, full moment tensor, mining induced seismicity.

1. INTRODUCTION

Fault zones in underground mining operations are considered as potentially dangerous zones, where excavation works may be complicated and can cause various geomechanical issues that may lead to roof instability (*e.g.*, Tajduś *et al.* 1996, Medhurst *et al.* 2008). Therefore, in regions of intense mining production all fault zones are treated with special care. In tectonically active zones, most of the seismic events are located along discontinuities in the rockmass, such as fault zones. In the case of Legnica Głogów Copper District (LGCD), Western Poland, the tectonic activity is not considered. There is only a small subsidence up to 1 mm/year (Zuchiewicz *et al.* 2007). The main fault zones in the vicinity of LGCD such as Sudetic Marginal Fault are considered not active as well as most discontinuities in Western Poland (Zuchiewicz *et al.* 2007), although two unexpected seismic events of magnitudes 2.8 and 3.8 occurred in 2007 and 2012, respectively, and were connected with movement along a small local normal fault near Jarocin (Lizurek *et al.* 2013). The Carpathian Mountains in the southern part of Poland are moderately active tectonically. This tectonic activity is accompanied by seismicity (*e.g.*, Wiejacz and Dębski 2009, Guterch *et al*. 2005, Guterch and Lewandowska-Marciniak 2002). This seismically active area is about 300-400 km away from the LGCD; therefore, all the seismicity reported in the LGCD region is considered as a mining induced seismicity (Gibowicz and Kijko 1994). The anthropogenic seismic activity of LGCD is related to the copper ore underground exploitation in 3 mines: Rudna, Polkowice-Sieroszowice, and Lubin. Those mines produce annually several hundreds of seismic events with a local magnitude in the range of 0.4 to 4.5 (Lasocki 2005). The largest events in this area were $Ml = 4.5$ of 24 March 1977 (Gibowicz *et al.* 1979), and $Ml = 4.3$ of 20 June 1987 (Gibowicz *et al.* 1989) both located within the Lubin Mine. On 19 March 2013, a tremor shook the surface of Polkowice town where the Rudna Mine is located; inhabitants reported that it was one of the most powerful tremors they experienced in the last couple of years. The magnitude of this event, $mb = 4.6$ according to EMSC (2013) and $Ml = 4.2$ according to IGF PAS, was the largest reported event since 1987 in this mining district. The seismic event was followed by a rockburst and several miners were trapped within the mining panel. Fortunately, there were no casualties, except of only two miners who were hurt and hospitalized. The event was reported to be located within G-3/4 mining panel of Rudna Mine, in the vicinity of Rudna Główna Fault. The aim of the study is to find if the mechanism of the event is connected with the existing fault zone nearby. To fulfil this goal we obtained two moment tensor solutions: the first one is obtained from the amplitudes of the first arrivals of waveforms recorded by in-mine seismic monitoring sys-

tem, and the second one which supports the results by inversion of the waveforms from new surface local monitoring network Legnica-Głogów Underground Mining INduced Earthquake Observing System (LUMINEOS). Macroseismic questionnaires were also prepared and incorporated within local community for estimation of the tremors felt range, which was performed to find more information for an event depth determination (Lizurek *et al.* 2013).

2. SITE AND DATA DESCRIPTION

Rudna Mine is located in the Legnica-Głogów Copper District (LGCD), Southwestern Poland (Fig. 1). Mining in the Rudna Mine is carried out in three mining regions: Rudna Central, Rudna West, and Rudna North. The Rudna Mine operates 11 shafts (3 for extraction, 4 for ventilation, and 4 for personnel and materials) ranging from 950 to 1150 m in depth. The mine is operated by KGHM S.A. company (Kombinat Górniczo-Hutniczy Miedzi).

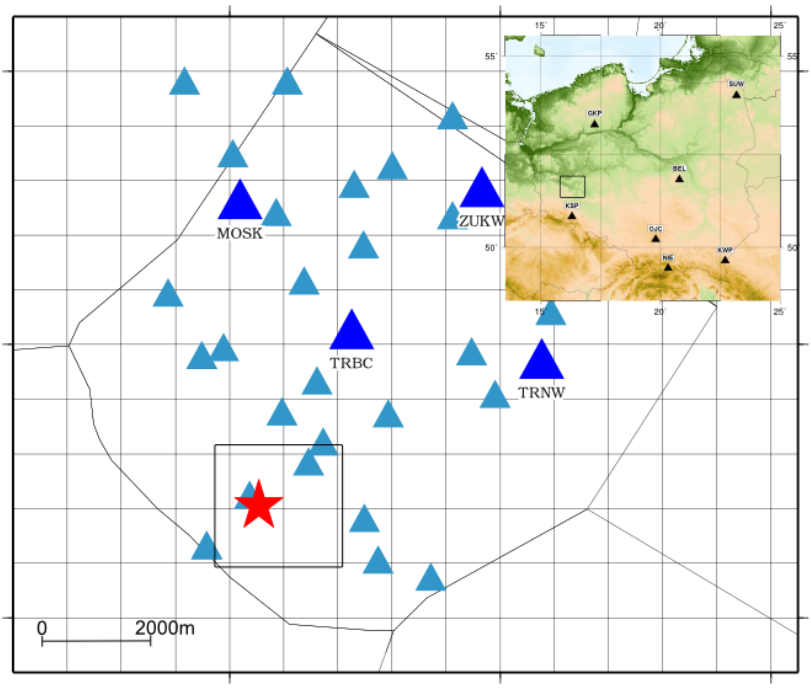


Fig. 1. Location of the 19 March 2013 seismic event and the seismic networks operating in the Polkowice area. Dark blue triangles denote LUMINEOS network stations and light blue triangles denote underground Rudna Mine seismic stations. The location of the considered mining panel is denoted with a rectangle. A topographic map of Poland with Polish Seismological Network broadband stations is in the upper right corner inset. The area of LGCD is denoted with black rectangle.

The ore bearing strata includes Permian series of dolomite, shale and sandstone layers. Above them is the evaporate series composed of anhydrite and salt layers. The approximate depth of ore bearing layers is between 900 and 1100 m. Ore is currently excavated at a depth of over 1150 m, with an increase to a depth greater than 1200 m expected during the next several years. Below the copper ore strata lies the Rotliegend sandstone. Exploitation is performed in several mining panels. A pillar-chamber mining method is employed. In this system, the ore is separated into pillars and passages separated by structural pillars, which serve to support the roof. Blasting is the method used for extracting the copper ore due to its hardness (15-170 MPa) and the diversity of the ore body that comprises 3 types of rock: dolomite and sandstone, separated by a main layer of copper-bearing shale. Those parts of the ore, having a thickness of up to 7 m, are worked out using the room-and-pillar method with an extra roof support; for ore having a thickness of over 7 m the room-and-pillar method is used with elimination of cavities by rock or hydraulic backfilling. The whole ore bearing strata are flat and slightly dipping towards NE (KGHM 2013). G-3/4 mining panel is located in southwestern part of Rudna Mine (Fig. 1). The Rudna Główna normal fault zone is located in north-eastern part of the panel; therefore, the panel lies in the hanging wall of this fault, which has NW-SE strike dipping to NE. Maximum fault throw values in this region are up to 4.5 m; there are smaller faults located within the panel area with NW-SE and NE-SW strikes.

Seismicity induced in the Rudna Mine is quite high: annually there are several hundreds of events reported by the mine with magnitude range from 0.4 up to 4.2. Among them there are couple of high energy events ($Ml > 3.0$) per month, for example in 2012 there were 2236 events of Ml above 0.9, and 1626 events with Ml above 1.2, which was magnitude of completeness for the Rudna Mine seismic catalog in 2012. The maximum magnitude observed of Rudna Mine induced events in 2012 was 3.4. Seismicity in G-3/4 panel started in 2001 and was observed during all mining works carried in this area to April 2013. There were almost 2000 seismic events in this panel and 31 events with energy higher than 10^7 J ($Ml > 3$). Activity varied from 0.12 to 0.56 events per day. The 19 March 2013 tremor has the highest energy released from the whole G-3/4 panel catalog.

Seismicity in the Rudna Mine is monitored continuously with local underground seismic network (Fig. 1), which is composed of 32 vertical seismometers at mining level; except of 5 sensors placed in elevator shafts, seismometers location depth varies from 300 down to 1000 m below surface. Epicenter location accuracy is about 50 m in terms of location uncertainty (Rudziński and Dębski 2011, Leśniak and Pszczoła 2008). The seismometers used in this network are vertical short-period Willmore MkII and MkIII sensors, with the frequency band from 1 to 100 Hz. This system is designed for

fast and precise location of events, but has some dynamic range limitations, which cause saturation of the *S*-wave part of the signal for high energy events at the nearest stations. Therefore, there are 4 stations doubled in the acquisition system and recording with lower gain to increase the ability of the network operator for more precise determination of the seismic energy released during seismic events. Lower gain stations have about 10 times lower gain coefficients: 1.5×10^{-7} m/s per count, while normal gain setting is 1.6×10^{-8} m/s per count. Examples of the saturated and unsaturated signals from the in-mine system are presented in Fig. 2.

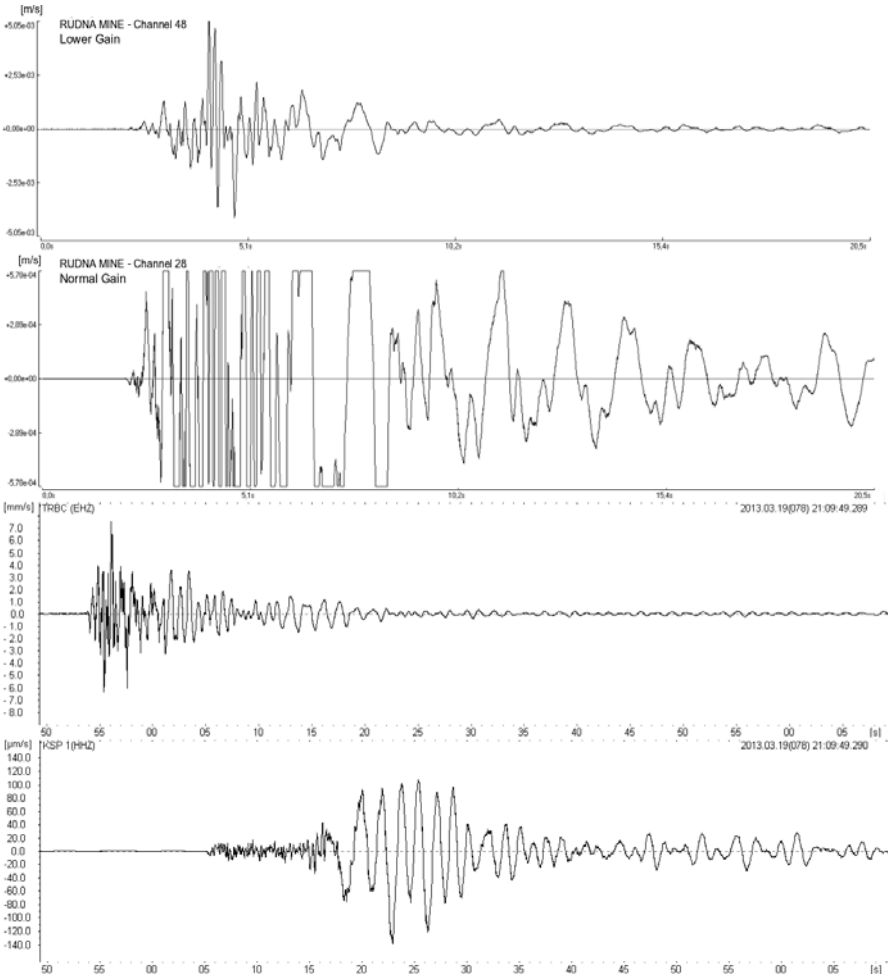


Fig. 2. Comparison of the signals recorded by Rudna Mine underground network (two upper traces, first one with lower gain, second one saturated), LUMINEOS network TRBC station, and broadband station KSP from PLSN.

At the beginning of 2013 the local surface network of Institute of Geophysics, Polish Academy of Sciences (IGF PAS) with acronym LUMINEOS was installed under the agreement with KGHM S.A. and Rudna Mine officials. This network at the moment of the 19 March 2013 event (Fig. 1) was composed of 4 short-period one-second triaxial seismometers LE-3D/1s manufactured by Lenartz Electronics.

Initial locations of the event provided by the mine have been improved using Single Event Relocation Method (Rudziński and Dębski 2011) with the use of *P*-wave arrivals recorded on in-mine seismic system. The event was located in north-eastern part of mining panel in the vicinity of Rudna Główna Fault at the excavation level (Figs. 1 and 3).

Spectral analysis was applied to estimate source parameters such as seismic moment, seismic energy, radius, and slip (Table 1). Seismic moment

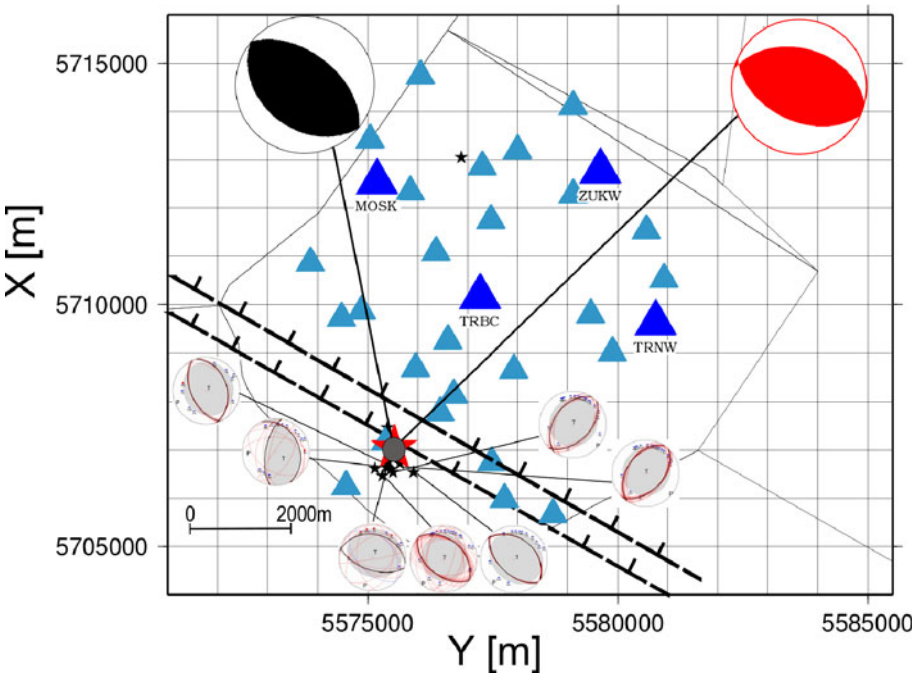


Fig. 3. Focal mechanism solutions presented as beach balls on the background of the Rudna Mine contour with underground seismic network (small light blue triangles) and LUMINEOS network (bigger dark blue triangles). On the left (black color) *P*-wave amplitude inversion fault plane solution is shown, and on the right (red color) waveform inversion fault plane solution is shown. The big star denotes epicentral location of the main event, the grey circle denotes epicenter location uncertainty, small stars present the aftershock locations with beach balls for the strongest of them, and the thick dashed line denotes Rudna Główna Fault zone margins.

Table 1

Source parameters of the 19 March 2013 seismic event calculated from mine network, LUMINEOS network, and KSP broadband station records

Station	F_c [Hz]	*P*-wave energy [J]	*S*-wave energy [J]	Seismic energy [J]	E_s/E_p	M_0 [Nm]	*Ml*	M_w	Radius [m]	Slip [m]
MOSK	3.7	2.4×10^8	4.60×10^9	4.84×10^9	19.2	2.80×10^{14}	4.3	3.6	175	0.11
TRBC	3.2	7.7×10^7	1.84×10^9	1.92×10^9	23.9	1.82×10^{14}	4.1	3.5	198	0.06
TARN	5.1	9.2×10^7	4.80×10^9	5.72×10^9	5.2	1.26×10^{14}	4.4	3.4	125	0.09
ZUKW	4.9	7.5×10^7	1.82×10^9	1.90×10^9	24.3	1.58×10^{14}	4.1	3.4	132	0.11
Rudna Mine network	2.7	1.3×10^8	1.72×10^9	1.85×10^9	13.2	5.58×10^{14}	4.1	3.8	223	0.13
KSP	3.3	7.2×10^8	9.62×10^8	1.68×10^9	1.34	8.34×10^{14}	4.1	3.9	180	0.30
Average	3.7 ± 0.9	2.2×10^8 ± 2.5×10^8	2.62×10^9 ± 1.6×10^9	2.84×10^9 ± 1.7×10^9	11.9 ± 9.7	3.56×10^{14} ± 2.8×10^{14}	4.2 ± 0.2	3.7 ± 0.2	172 ± 38	0.13 ± 0.09

was calculated upon Boore and Boatwright (1984) assumption of low-frequency level of the far-field displacement, seismic energy was calculated with the use of the Boatwright and Fletcher (1984) relation of radiated energy of *P* and *S* waves and the energy flux contained in the *P*- and *S*-wave trains.

Source radius was calculated upon Brune (1970, 1971) formalism with the use of Madariaga (1976) circular fault model, which was also used for the slip estimation from Aki and Richards (1980) seismic moment definition. Moment magnitude was estimated from seismic moment (Hanks and Kanamori 1979), while local magnitude was estimated from empirical relation between energy estimates and local magnitude for LGCD; all the details of the estimation methods of the above-mentioned parameters were described by Gibowicz and Kijko (1994) and Niewiadomski (1997). The analysis was performed using unsaturated records of the seismic system of Rudna Mine, the records from LUMINEOS network and KSP broadband station of PLSN network, which is located about 80 km south from the epicentre. Signal examples from the above-mentioned networks are presented in Fig. 2.

The relevant parts of seismograms of *P* and *S* waves were selected manually, and then transformed by Fast Fourier Transformation (FFT). The resulting amplitude spectra were corrected for attenuation effects with $Q = 400$ and 200, respectively, for *P* and *S* waves. For further calculations we set velocities of *P* and *S* waves in source at $Vp = 5900$ m/s and $Vs = 3400$ m/s.

The free surface correction for surface network and broadband stations were used, while in case of in-mine system the free surface correction can be neglected. Shear modulus was set at 2.7×10^{10} Nm^{-2}; the account for *P*- and *S*-wave radiations were set for 0.52 and 0.63, respectively (Lizurek and Wiejacz 2011, Kwiatek 2013, Gibowicz and Kijko 1994). Far-field displacement spectra were approximated on logarithmic scale. The two asymptotes were calculated using of integrals of squared displacement and velocities of ground motion. This has allowed us to estimate the spectral plateau, corner frequency and energy flux in observation point for every spectrum of *P* and *S* waves.

Further energy of *P* and *S* waves was calculated from an energy flux in far field displacement. The average value of the seismic energy released during the event is 2.84×10^9 J, average seismic moment 3.56×10^{14} Nm, local magnitude 4.2, moment magnitude 3.7, source radius 172 m, and slip 13.3 cm. These values are relatively high compared with the mining induced events in LGCD, where energy release is rarely above 10^8 J. In the past, such high energy release was noted for the two biggest events from LGCD: in 1977 with energy 2.5×10^{10} J, $Ml = 4.5$ (Gibowicz *et al.* 1979) and in 1987 event with $Ml = 4.3$ and energy 5.0×10^9 J (Gibowicz *et al.* 1989). *S*-wave to *P*-wave energy ratio is quite high − almost 12 in average, which is characteristic for tectonic earthquakes with shearing as a main rupture component, rather than for mining induced events, where non-shearing components may play a role in rupture process, especially in the case of LGCD mines (Gibowicz and Kijko 1994, Lizurek and Wiejacz 2011). Source parameters are comparable with the parameters obtained using the same methodology for tectonic seismic event near Jarocin from 6 January 2012 (Lizurek *et al.* 2013), but the mining induced event from 19 March 2013, was shallower, had higher energy release, larger slip value and higher magnitude, which made it more intensely felt on smaller area. The large event was followed by 10 aftershocks within the area of the G-3/4 mining panel during next 16 hours with the highest energy released, 7.1×10^4 J. During the next month, only 4 events were reported for this mining panel, none of them with energy higher than 1×10^4 J. The aftershock distribution does not follow the fault strike, but is dispersed over the vicinity of the studied event, mostly within the G-3/4 panel (Fig. 3).

3. FOCAL MECHANISM

Focal mechanism was calculated using two different approaches: the first one using the moment tensor inversion in the time domain from amplitudes of *P*-wave onset recorded by underground seismic network in Rudna Mine (Wiejacz 1992, Awad and Kwiatek 2005), the other using waveform inver-

sion (Cesca *et al.* 2010) of the signals recorded with LUMINEOS network (Figs. 1 and 3).

3.1 *P*-wave amplitude inversion

Calculations of the moment tensor from the mine underground network records were performed using FOCI software 3.0 (Kwiatek 2013). The inversion of the *P*-wave amplitude was performed in time domain. The registered first onsets may be of different types: direct *P* waves and refracted from the overlying anhydrite layer or the thick strata of sandstone underlying the ore deposit. This is caused by the geological situation. The layers are almost flat, the dip of the layers being about 4 degrees NE. The ore bearing strata are a dolomite layer of 60-90 m thickness. Above the ore bearing strata, there exists the evaporate series mainly composed of the anhydrite of 160 m thickness, and below the ore bearing strata there is a sandstone layer of about 300 m thickness.

The observation of seismic recordings from the in-mine system led to the conclusion that the direct *P* wave is observed when the source-receiver distance is smaller than 1 km, because the direct wave gradually disappears and only the refracted waves are observed at longer distance. The velocities determined for the waves were: 5 km/s for the direct wave, 5.9 km/s for the refracted wave from the overlying anhydrite layer, and 5.6 km/s for the refracted wave from the underlying sandstone layer (Wiejacz 1992). In the present study all types of waves were used with 22 onsets on 22 stations; these were mostly the waves refracted from the overlying anhydrite layer, 2 onsets of wave refracted from the underlying sandstone strata, and only one direct wave onset was recorded at the station nearest to the source.

The input parameters are the amplitude and polarity information on the first *P*-wave displacement pulses. According to Fitch *et al.* (1980), the recorded displacement for the vertical component of the *P*-wave phase is:

$$U_z^P(x,t) = \frac{1}{4\pi\rho\alpha^3 r}\left[\bar{\gamma} M \dot{s}\left(t - \frac{r}{\alpha}\right)\bar{\gamma}\right] l_z \ , \tag{1}$$

where ρ is the average density, r is the source-receiver distance, α is the average velocity of *P* wave, M is the seismic moment, l_z is the cosine of the angle of the incidence, and $\bar{\gamma}$ is the take-off angle. The Source Time Function (STF) was based on the Haskell's source model (Haskell 1953):

$$\dot{s} = \begin{matrix} 1/T, \ 0 < t < T \\ 0, \quad \text{elswhere} \end{matrix} , \tag{2}$$

where T is the rupture time.

Moment tensor is obtained by solution of a set of N equations of type 1. Due to the six independent elements of the moment tensor there must be at least six such equations, but the more, the better. When the condition of the zero trace is imposed on solution, the deviatoric moment tensor may be determined, which excludes the type of mechanism with volumetric change in the source. When the condition of zero trace and zero denominator are set, the solution is limited to the double couple source. The deviatoric, pure shear

Table 2

Source parameters (moment tensor and nodal planes) of the 19 March 2013, seismic event calculated from in-mine network, LUMINEOS network

Moment tensor solutions			Nodal plane 1 orientation	Nodal plane 2 orientation
Method of inversion	Tensor components	[Nm]	Strike/Dip/Rake	Strike/Dip/Rake
Solution A	M_{11}	-1.26×10^{14}	309/47/96	120/43/83
	M_{22}	-7.10×10^{13}		
	M_{33}	4.46×10^{14}		
	M_{12}	-6.76×10^{13}		
	M_{13}	4.49×10^{12}		
	M_{23}	-7.10×10^{13}		
Solution B	M_{11}	-6.11×10^{14}	115/48/97	285/42/83
	M_{22}	-3.99×10^{14}		
	M_{33}	4.51×10^{14}		
	M_{12}	-7.55×10^{13}		
	M_{13}	3.47×10^{13}		
	M_{23}	6.89×10^{13}		
Decomposition components				
Solution A	DC		42%	
	CLVD		33%	
	ISO		25%	
Solution B	DC		41%	
	CLVD		9%	
	ISO		–50%	
Uncertainties of the methods				
Nodal plane orientation uncertainty (Strike/Dip/Rake)				
Solution A	±12/9/5			
Solution B	±24/11/10			

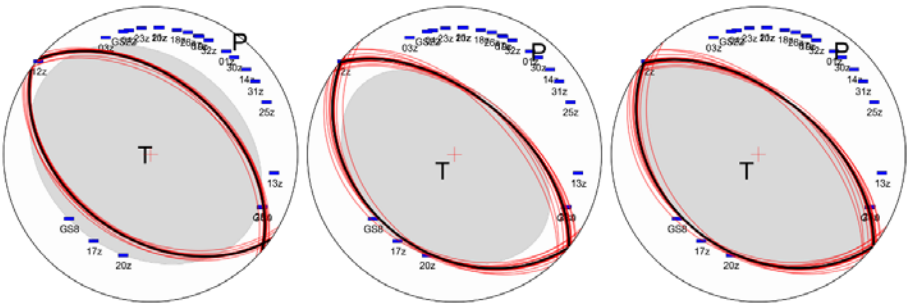


Fig. 4. Mechanism solutions with *P*-wave amplitude inversion (solution A) – from the left: Full Moment Tensor, Trace Null, and Double Couple solution. The numbers and signs denote the station location and polarization on focal sphere. All solutions are presented as a Lambert azimuthal equal-area projection on lower hemisphere Schmidt net. The red lines denote the nodal plane orientation for different station configurations of Jacknife test, and the black lines denote the nodal planes from the obtained solution with all stations used.

and full moment tensors were calculated using the L2 norm as a measure of the misfit (Wiejacz 1992, Awad and Kwiatek 2005). The sensor distance from the hypocenter was mainly above 1500 to 8000 m in case of in-mine system, with one exception of channel 12, which was about 200 m from the source. It is within source dimension radius, but it was the only channel with direct wave recorded. The Jacknife test of station influence on the result has shown (Fig. 3, Table 2) that it has no big influence on tensor components and nodal plane orientation, but the use of this channel allows to decrease the moment tensor uncertainty; therefore, it was included into the moment tensor inversion procedure. Moreover, the strike/dip/rake uncertainty was relatively small – not higher than 12 degrees. The result of the *P*-wave amplitude inversion approach will be named further as solution A. The full moment parameters along with decomposition for double-couple (DC), compensated linear vector dipole (CLVD) and isotropic (ISO) components are presented in Table 2 and Figs. 3 and 4. The maximum estimation error of the moment components calculated with FOCI (Awad and Kwiatek 2005) software was 1.08×10^{13} Nm.

3.2 Waveform inversion

The moment tensor solution using LUMINEOS seismograms inversion was carried out with "kiwi tools" package (Heimann 2011), http://kinherd.org. Because the network consists of 4 three-component short period seismometers, it gave an opportunity to work with 12 traces in frequency range 1-50 Hz. The methodology adopted eikonal source model and multistep in-

version strategy, as described by Cesca *et al.* (2010). It allowed to work with both amplitude spectra and seismograms (time traces) in time domain at different frequency ranges. The eikonal model can be described by 13 parameters which can be grouped into two groups; the first 8 parameters describe a point source while the last 5 are connected with finite rupture (Cesca *et al.* 2010). Parameters needed to describe the point source are: strike, dip, rake, scalar moment and latitude, longitude, depth, and origin time. These groups of parameters explain radiation pattern and locate the event in space and time. The remaining 5 are related to the finite rupture and can be obtained with knowledge of point source. The rupture propagates with relative rupture velocity v_r within the rupture surface, with radius R from nucleation point characterized by parameters nucleation coordinates n_x and n_y. The last parameter is the rise time which defines the time required for the slip occurence at each point. The v_r is defined as percentage of the shear velocity (Cesca *et al.* 2010). Since the LUMINEOS network is very sparse we decided, in this paper, to study only point source inversion. It means the result of the analysis was a focal mechanism described by strike, dip, rake for both nodal planes and scalar moment, centroid latitude, longitude, depth, and origin time. Enough information to obtain point source mechanism can be found in seismograms below the corner frequency (Gibowicz and Kijko 1994). The analysis was performed using the 1D (Fig. 5) very local velocity model for Rudna Mine (Dec *et al.* 2011).

The inversion procedure started with hypocentral location given by the relocated hypocentre with underground in-mine network (Figs. 1 and 3) and average scalar moment obtained by spectral analysis (Table 1). To avoid non-uniqueness caused by nonlinearity of the inversion problem, the inversion was run several times with different starting configuration. In this point, the varied starting parameters were source depth (in a range between 0.5 and 1.5 km with a 0.1 km step) Because the event was located outside the LUMINEOS network (Figs. 1 and 3), in the first step the fitting of amplitude spectra was preferred than seismogram fitting in the time domain. Such approach has some advantages and could give more reliable results in this case (Cesca *et al.* 2006).

All 12 displacement traces were used during the inversion. To stabilize the solution we decided to focus on *P*-waves onset. The chosen bandpass was 2.0-3.0 Hz, which covered frequency range below the corner frequency of the recorded seismic waves. The recorded corner frequencies in LUMINEOS network were from 3.7 to 5.1 Hz (Table 1). The first 5 seconds of the seismograms were used in analysis. We decided to use this time interval to be sure that a long enough *P* train was recorded even on the farthest station ZUKW. It was especially important in case of stability of spectral inversion step. The solutions were obtained with L2 norm. After the spectra

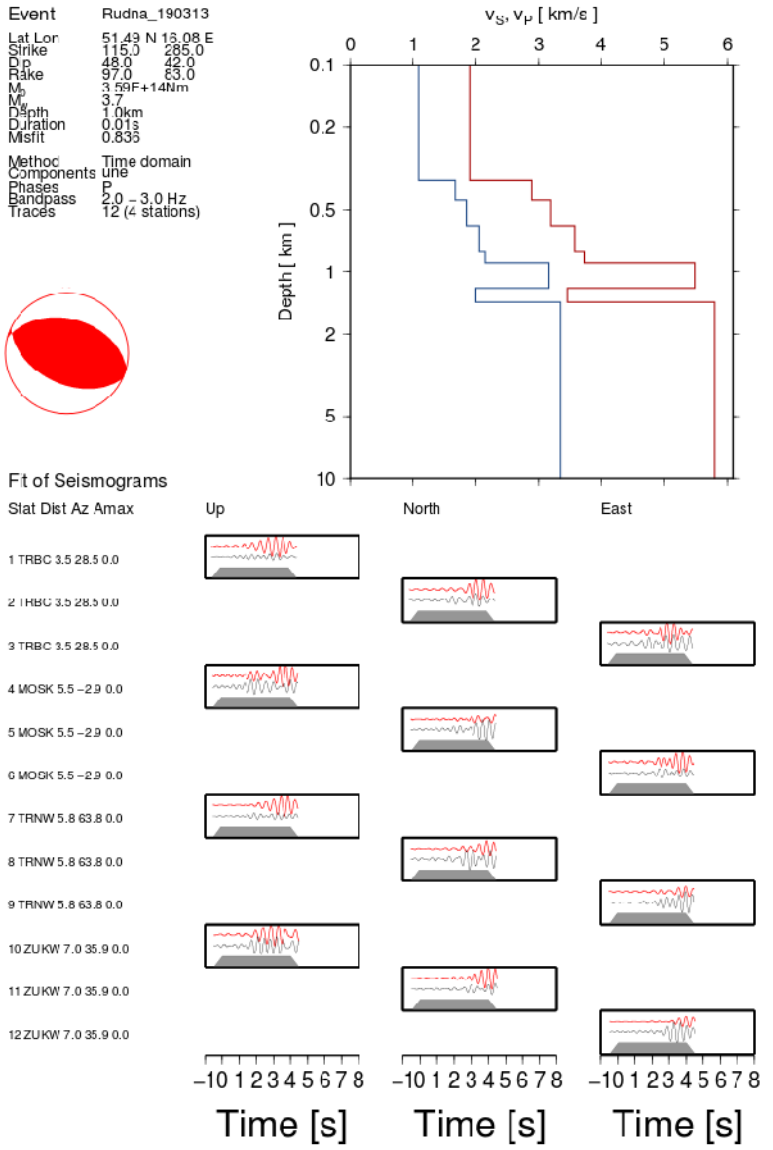


Fig. 5. Details of the inversion of the waveform results with velocity model based on Dec *et al*. (2011) (top-right panel). Inversion results contain point source parameters estimation and focal mechanism. Fit of *P*-waves seismograms for the used stations include observed (red) and synthetics displacements (blue). The gray shaded area is the applied time taper. The point source parameters are: strike: 115, 285; dip: 48, 42; rake: 97, 83; moment magnitude: 3.7, seismic moment: 3.59×10^{14} Nm. The inversion used the *P*-wave trains and three components; u: vertical, n: N-S, e: E-N recorded on 12 traces on 4 stations. The frequency used: 2-3 Hz.

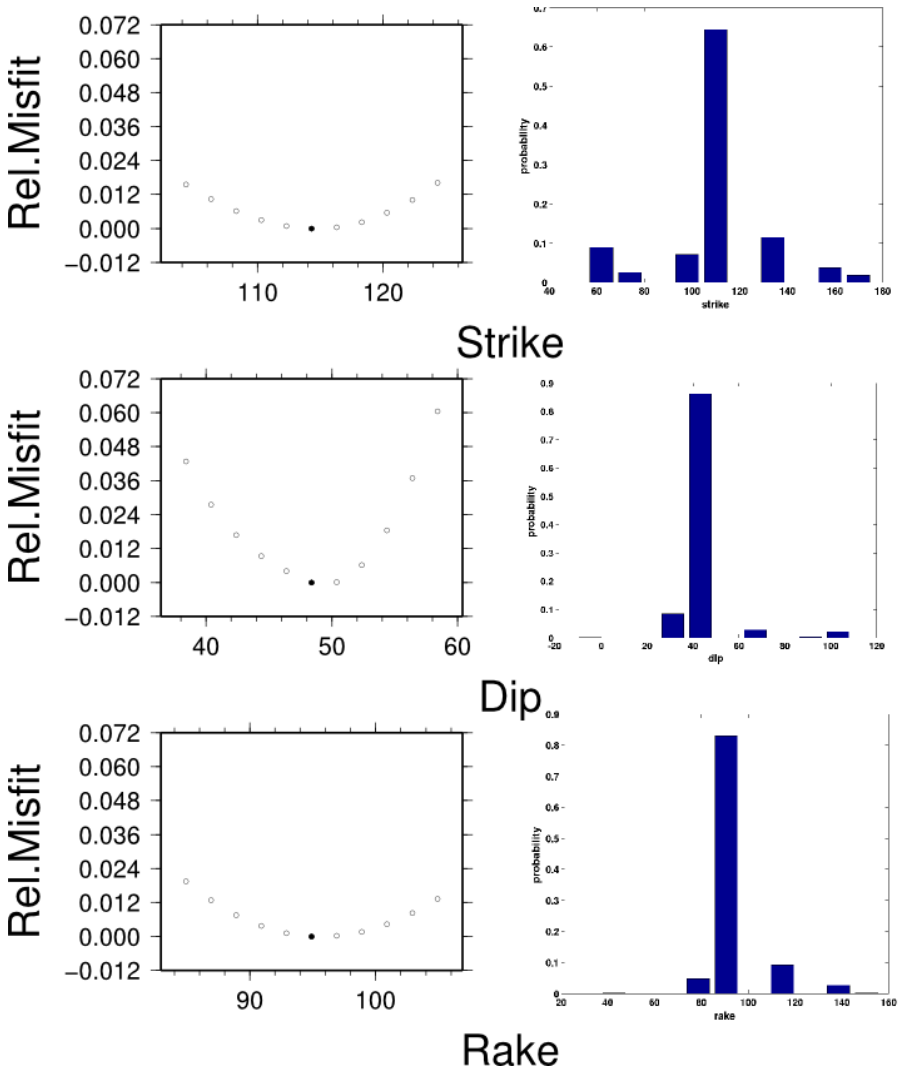


Fig. 6. Optimal solution for strike, dip, and rake (black circles) for one of the nodal planes obtained with spectral inversion step (left panel) and the result of the bootstrap test (right panel). The latter provide probability of the solution B. The histograms are shown for one of the nodal planes.

inversion optimization (Fig. 6) step, the fault parameters: strike, dip, rake for 2 possible fault configurations, scalar moment, and centroid depth were estimated. Since we used just spectra fitting, it was no information about polarities of seismograms. Without this, it is impossible to discriminate compression and dilatation quadrant.

The main task for the second step was to fit the seismograms in time domain and, in practice, to include polarities into the analysis. The same frequency and time windowing like in the previous step were used for fitting seismograms. Synthetic seismograms were calculated for 2 possible fault configurations obtained with the first step and then optimization in L2 norm was used. The results of point source inversion are gathered in Table 2 and the best DC solution is presented in Fig. 5.

The solution B had a relatively high misfit measured in L2 norm as the difference between the amplitude spectra of the observed and synthetic seismograms. It is probably due to the azimuthal coverage and much more complicated velocity model with mostly unknown velocity for laterally variable subsurface sedimentary layer. Both of those issues can influence the obtained results. To be more sure that the solution is stable we provide a bootstrap test which repeatedly solves the inverse problem using a different subset of LUMINEOS seismic stations (Heimann 2011). We used 256 bootstrap samples to provide the probability of the solution B. Results of the test for one of the nodal planes are presented in Fig. 6 and the uncertainty of strike, dip and slip results are included in Table 2. Let note that the stability is very high, but error estimates in the case of strike and rake were two times bigger than in solution A; nevertheless, the studied event was the first high-energy event recorded with LUMINEOS network.

3.3 Interpretation of the results

The focal mechanism of the studied event was a thrust fault of NW-SE orientation similar to the strike of Rudna Główna Fault. Both types of solutions are similar in terms of nodal plane orientations. The nodal plane strikes differ within 24 degrees at most, while dip differs only by 5 degrees, and rake differs up to 14 degrees, which is within the range of error estimates for solution B and slightly higher than the uncertainty of solution A (Table 2). Solution details are listed in Table 2. There was crucial difference in the azimuthal coverage of the stations used in both methods of mechanism estimations (Fig. 3). The azimuthal coverage was very good between NE and SW in the case of solution A, with only two stations on the opposite azimuth angles, while LUMINEOS stations cover about 60 degree angle from N to ESE. In the case of *P*-wave amplitude inversion method there was performed Jacknife test of the solution sensibility for the station distribution (Kwiatek 2013). The obtained result shows very small sensitivity for changes in station distribution performed by excluding single station from the station set and then repeating such action for all the stations (Fig. 4).

Full Moment Tensor decompositions show a similar double-couple (DC) component: 42% in solution A and 41% in solution B. In the latter one, the

dominance of implosive component is observed (about 50%), while solution A shows more complex decomposition with the highest value of DC component, but smaller than the sum of non-DC components (CLVD – 33% and ISO – 25%). The *Es*/*Ep* ratio obtained with spectral analysis was about 12, which may suggest the shearing mechanism. Lizurek and Wiejacz (2011) have shown that in case of Rudna Mine events the high *Es*/*Ep* ratio does not always indicate pure shearing mechanism, which is also the case of the studied event.

Relatively high non-DC components in the solution of focal mechanism in the case of mining induced events in comparison with tectonic earthquakes is not exceptional (*e.g.*, Gibowicz and Kijko 1994, and literature therein). Mechanisms with implosive or explosive component were reported for LGCD mines (Wiejacz 1992). Recently the non-DC component of focal mechanism solution was also considered as an indicator of the induced seismicity (Cesca *et al.* 2013). On the other hand, in tectonic earthquakes non-DC components of mechanism solutions are considered a result of anisotropy of the rocks. Even in case of pure shear faulting, the non-DC values derived from full moment tensor solution can reach up to 30% (Vavryčuk 2005, Davi and Vavryčuk 2012). The solution A indicates that the mechanism was complex: a mix of shearing with one-axial tension and explosive component in total decomposition of full moment tensor, while in solution B the mechanism was rather closing of the excavation chamber with an upward movement of the floor, which is a hanging wall of Rudna Główna Fault. Those significant discrepancies of mechanism solutions needed further explanation, but due to essential differences between the methods used for the inversions and obtained quality measures of the results it is impossible to directly compare the obtained results in terms of misfit and error estimates. To solve these issues we provided the forward modelling of waveforms for both of those solutions and compared the synthetic waveforms obtained for 4 in-mine stations with lower gain. Signals of the studied event recorded on those stations were characterised by not saturated beginning of the signal (*P*-wave train) and additionally those signals were not used for both of the obtained solutions.

Synthetic signals were obtained within frequency range of 2-3 Hz, which corresponds with frequency range used for inversion of waveforms in the solution B and it is a range below the average corner frequency of the observed signals. However, it is worth to note that the first *P*-wave pulses used for solution A were not filtered and characterised by the 4-17 Hz range. A comparison was made with the use of signal correlation coefficient of the synthetic signal with the recorded one. The higher correlation coefficient the better the synthetic signal fits the recorded one. Results of the modelling are shown in Fig. 7.

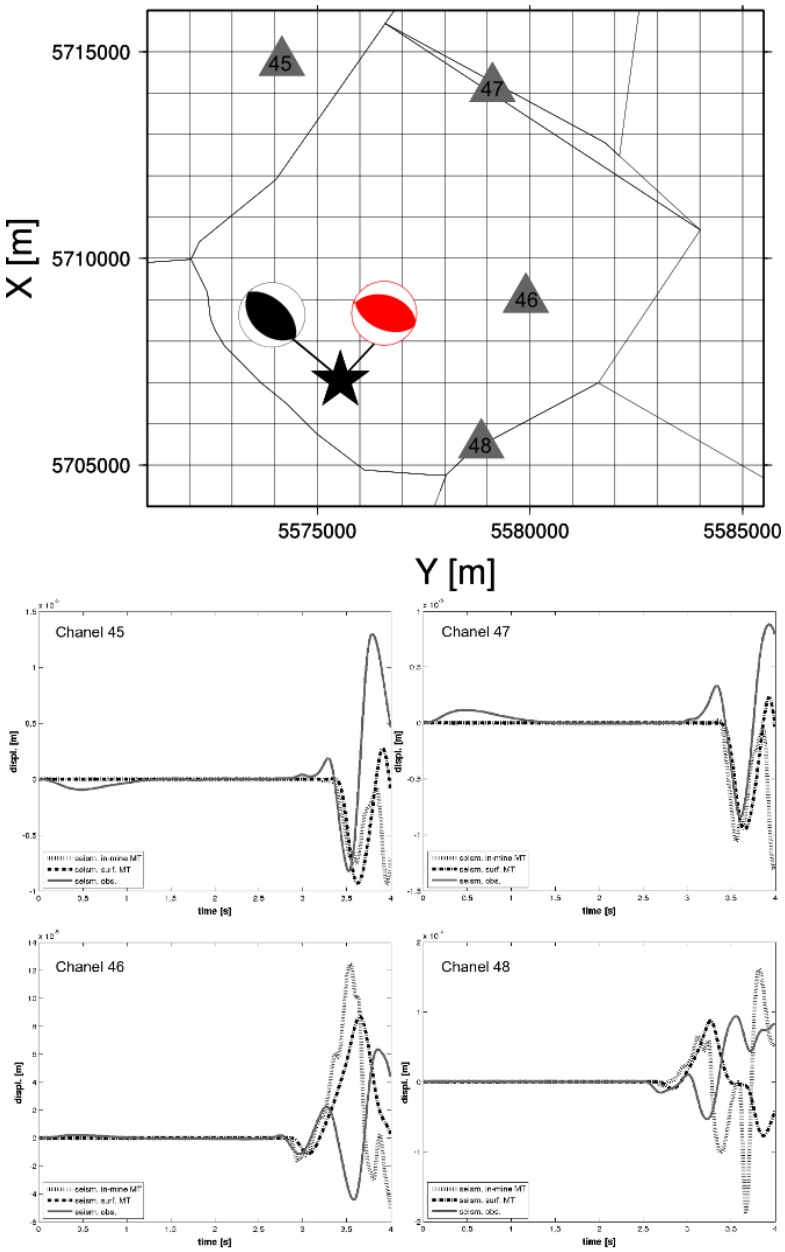


Fig. 7. Forward modeling results for full moment tensor solutions. The observed displacement signal filtered in frequency range 2-3 Hz for 4 in-mine stations is represented by solid line, the solution A and B synthetic displacement waveforms modeled in the same frequency range (2-3 Hz) are represented with thin dashed and thick dot-dashed lines, respectively.

Table 3

Results of cross-correlation coefficients of the modelled waveforms and observed ones from the in-mine network

Chanel of Rudna Mine network	Solution A – Observed waveform correlation	Solution B – Observed waveform correlation	Solution A – Solution B correlation
Channel 45	0.72	0.73	0.98
Channel 46	0.66	0.61	0.92
Channel 47	0.49	0.64	0.95
Channel 48	0.45	0.53	0.42
Average	0.58	0.63	0.81

The values of the correlation coefficients for the modelled waveforms are presented in Table 3. Visual comparison shows that the solution A more accurately resolves the first onset of the *P* waves in terms of the polarity and amplitude, while the solution B is more accurate in resolving the longer *P*-wave train, especially in terms of the amplitude values. Neither of the synthetic displacement waveforms fits the whole observed *P*-wave displacement waveform perfectly well. The correlation of observed signal and synthetic ones are, in average, around 0.6 for both tensors used as source models. In case of two channels of solution A correlation coefficient value is lower than 0.5 while for solution B in every examined channel this values are higher than 0.5. Channel 48 waveform was the worst modeled one by both solutions even in terms of cross-correlation of both synthetic waveforms it has the lowest coefficient value, below 0.5. The azimuth between the event epicentre and the station was almost parallel to the strike of the nodal planes, and that is why the poor signal correlation may be expected due to the radiation pattern of the waves. Only in case of channel 46, the solution A modelling obtained higher cross-correlation coefficient than the solution B. According to the forward modelling approach, the solution B is better in terms of retrieving waveform within range of 2-3 Hz. But both solutions of Moment Tensor allowed to model very similar waveforms – their cross-correlation coefficient of synthetic waveforms is, in average, about 0.8 (Table 3), but in case of 3 out of 4 modelled records it is above 0.9, in agreement with the nodal plane orientations of both solutions, which do not differ much (up to 24 degrees). The main difference in results obtained in forward modelling of waveforms with the use of Moment Tensors comes from the different frequency range covered by the data used in Moment Tensor inversion.

Solution A Moment Tensor was obtained with the use of first *P*-wave pulse (frequency range 4-16 Hz), while the solution B was obtained with

waveform inversion in 2-3 Hz band. In case of surface recordings used for solution B inversion, the high frequency part of the signal is not recorded because of the scattering and attenuation through the ray path and the thick sediment layers working as a natural filter.

The aftershock sequences in LGCD are rare; they were observed only in case of the two biggest events (Gibowicz *et al*. 1979, 1989). In those cases, the location of the aftershock and their mechanisms were reported as consistent with the main event focal mechanism fault plane orientation. Aftershocks of 19 March event are mainly concentrated within the G-3/4 mining panel where the main event occurred. All of them were located on SW wing of the Rudna Główna Fault (Fig. 3). The aftershocks location was in ceiling above the ore bearing strata (Koziarz 2013), but it does not align with the Rudna Główna Fault, while the main event nodal plane strikes are almost parallel to the Rudna Główna Fault strike reported as 285 to 295 degrees. This fault is dipping towards NE (Koziarz 2013), while the resulting solutions had a dip between 43 and 45 degrees, consistent with the fault dip. Moreover, there are only 2 groups of crack orientations observed in rockmass within the G-3/4 mining panel, which are of about 290 and about 25 degrees (Koziarz 2013).

For 7 biggest aftershocks ($E \approx 10^4$ J) it was possible to calculate focal mechanism with the use of *P*-wave amplitude inversion. The obtained focal mechanisms were mostly non-DC with CLVD component range from 59 to 87% with reverse fault regime. The aftershocks (Table 4) can be divided into three groups: first with nodal plane orientations almost parallel to the main

Table 4

Aftershock nodal plane orientation and full moment tensor components

Aftershock		Nodal plane 1 Strike/dip/slip	Nodal plane 2 Strike/dip/slip	Full moment tensor components ISO/CLVD/DC [%]
No.	Time			
1	19 Mar 2013, 22:13:17	155/53/86	342/38/96	21/67/12
2	19 Mar 2013, 22:59:25	122/52/92	298/37/87	23/71/6
3	19 Mar 2013, 23:05:28	108/59/94	280/31/83	21/64/15
4	19 Mar 2013, 23:27:39	130/50/89	311/41/91	11/87/2
5	19 Mar 2013, 23:45:18	134/54/93	309/37/86	21/59/20
6	20 Mar 2013, 00:14:47	46/48/98	214/43/82	13/75/12
7	20 Mar 2013, 08:41:08	44/47/97	214/44/83	14/77/9

crack orientations (aftershocks Nos. 6 and 7), second with nodal plane similar to the obtained orientation of the main event (aftershocks Nos. 2-5), and third not aligned to any of reported discontinuities (aftershock No. 1).

The results of aftershock analysis are in good agreement with main event mechanism. The nodal plane orientations of aftershocks, as well as their locations, indicate the reverse faulting regime of the main event and following roof collapse, which activated the ceiling weakened zones on SW wing of Rudna Główna Fault. It is consistent with Rudna Główna Fault plane orientation dipping toward NE, while the main event reverse faulting caused a relative upward movement of NE wing toward SW.

4. MACROSEISMIC DATA

The macroseismic questionnaire with felt reports, intensity and consequences of the earthquake was placed on the web page of the Institute of Geophysics, Polish Academy of Sciences (IGF PAS). Local authorities were informed about the survey and were very helpful in distribution of the ques-

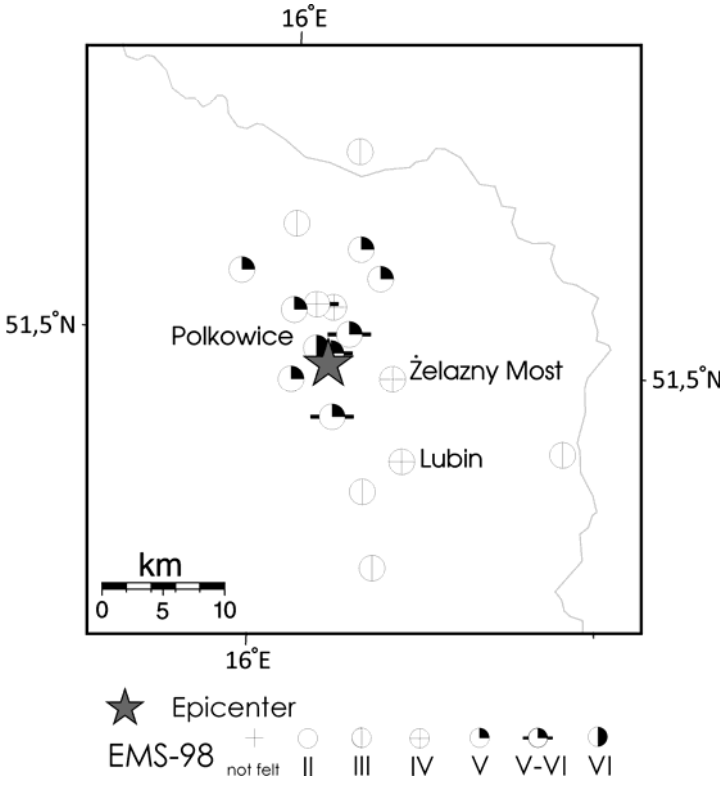


Fig. 8. Map of felt intensity of the tremor on 19 March 2013, from Polkowice and its vicinity.

tionnaire in the local communities. As a result, about 100 questionnaires have been filled in, which is not representative for a full macroseismic analysis including isoseismal maps. Analysis of such a small number of inquiries gave only rough estimation of the scale of intensity of macroseismic effect. People reported oscillations, swinging of the lamps, movement of small objects and chairs, opening/closing of windows or doors, also minor cracks on walls and chimneys. The citizens of Polkowice reported that it was the most powerful tremor they experienced in the last couple of years. The earthquake was felt mainly in Polkowice town and its surroundings, in Głogów and Lubin towns. The felt reports cover an area up to 20 km from the epicenter (Fig. 8). The highest intensity was VI according to EMS-98 scale (Grünthal 1998) it was reported from a small area close to the epicenter, not farther than 2 km; the next reports of V-VI intensity levels were reported from the area up to 8 km to the south from the epicenter. The macroseismic observations confirm that the event was shallow; the area of the highest intensity was small and limited to the closest vicinity of the source epicentre location. Unfortunately, the small number of filled questionnaires did not allow to perform any analysis of directivity of the intensity according to the focal mechanism solutions.

5. DISCUSSION AND CONCLUSIONS

The seismic event of 19 March 2013, was the third most powerful event recorded in LGCD, $Ml = 4.2$. The felt reports described this event as an exceptionally strong and broadly felt in comparison with other induced events from LGCD area in the last couple of years. It was a very dangerous event with rockburst and high degradation of the mining panel installations. On-site examinations brought the following conclusions: there were floor cracks observed after the event, which were not continued in the roof and the rockburst was limited only to the very thin part of the roof (Koziarz 2013). The main rockmass of the roof in the panel was stable and solid. The event's location was within the fault zone at the depth of excavation, but the floor deformation and stable ceiling suggest that the rupture started below the excavation level. Taking into account that the depth determination is not as accurate as epicentral location (Rudziński and Dębski 2011) and the radius of the rupture was about 200 m, those observations are consistent with instrumental data.

The mechanism of the event was complex, but the rupture initially acted as a reverse fault, although the Rudna Główna Fault is a normal one; therefore, a possible scenario of this event is that the mining works changed the stress state in the rockmass near the fault zone. The result was that the fault regime was reverse. When the accumulated energy overcame the friction the fault plane acting as the weakest part of the medium, the rupture and move-

ment followed the old fault plane orientation. It is confirmed by the focal mechanism solutions, that the nodal plane strikes are almost parallel to the Rudna Główna Fault strike reported as 285 to 295 degrees. This fault is dipping towards NE (Koziarz 2013), while the resulting solutions had a dip between 43 and 45 degrees, which is consistent with the fault dip. Moreover, there are only 2 groups of crack orientations observed in rockmass within the G-3/4 mining panel: about 290 and about 25 degrees (Koziarz 2013).

The amplitude inversion results of Full Moment Tensor (solution A) differ from those obtained with the waveform inversion (solution B) most probably due to the uncertainty of the velocity model, poorer azimuthal coverage for the latter method and different frequency ranges of the observed signals used for inversion. Comparison of the results with forward modelling described in Section 3 shows that the in-mine network and LUMINEOS network records are dominated by different frequency ranges, which correspond to the different rupture stages. The high frequency *P*-wave observations of in-mine network used for inversion allowed to retrieve the initial rupture of the studied event, which had a complex character, but its orientation was parallel to the pre-existing cracks of the Rudna Główna Fault zone. The rupture acted on the weak part of the fault zone. The initial rupture was then followed by the rockburst, which dominated the surface records with the lower frequency signal. This was much better retrieved from the waveform inversion (solution B) in terms of the forward modelling test results. The methods used for Moment Tensor solutions are complementary and allow to describe the rupture during the event. Both mechanism solution results support the hypothesis that this event was induced along a pre-existing fault. Some geomechanical works, such as Marcak and Mutke (2013), support the thesis that mining stress changes play a significant role in the reverse faulting mechanism occurrence. This work was focused on an influence of stress changes caused by mining on tectonic stress in Polish coal mines. Similar works in LGCD may in future put some new light onto the mechanism of the stress pattern change. Currently, there is no proof in the available data for reactivation of the tectonic processes in this area, which is considered as inactive.

Acknowledgments. We express our gratitude to Mr. Eugeniusz Koziarz, Head of the Rudna Mine Seismological Station, for help and provision of data necessary for our analysis.

This paper is partially supported by Institute of Geophysics scientific project No. 3d/IGFPAN/2013mł.

References

Aki, K., and P.G. Richards (1980), *Quantative Seismology: Theory and Methods*, W.H. Freeman & Co., San Francisco.

Awad, H., and G. Kwiatek (2005), Focal mechanism of earthquakes from the June 1987 swarm in Aswan, Egypt, calculated by the moment tensor inversion, *Acta Geophys. Pol.* **53**, 3, 275-291.

Boatwright, J., and J.B. Fletcher (1984), The partition of radiated energy between P and S waves, *Bull. Seismol. Soc. Am.* **74**, 2, 361-376.

Boore, D.M., and J. Boatwhright (1984), Average body-wave radiation coefficients, *Bull. Seismol. Soc. Am.* **74**, 5, 1615-1621.

Brune, J.N. (1970), Tectonic stress and the spectra of seismic shear waves from earthquakes, *J. Geophys. Res.* **75**, 26, 4997-5009, DOI: 10.1029/JB075i026p04997.

Brune, J.N. (1971), Correction to "Tectonic stress and the spectra, of seismic shear waves from earthquakes", *J. Geophys. Res.* **76**, 20, 5002, DOI: 10.1029/JB076i020p05002.

Cesca, S., E. Buforn, and T. Dahm (2006), Amplitude spectra moment tensor inversion of shallow earthquakes in Spain, *Geophys. J. Int.* **166**, 2, 839-854, DOI: 10.1111/j.1365-246X.2006.03073.x.

Cesca, S., S. Heimann, K. Stammler, and T. Dahm (2010), Automated procedure for point and kinematic source inversion at regional distances, *J. Geophys. Res*. **115**, B6, B06304, DOI: 10.1029/2009JB006450.

Cesca, S., A. Rohr, and T. Dahm (2013), Discrimination of induced seismicity by full moment tensor inversion and decomposition, *J. Seismol.* **17**, 1, 147-163, DOI: 10.1007/s10950-012-9305-8.

Davi, R., and V. Vavryčuk (2012), Seismic network calibration for retrieving accurate moment tensors, *Bull. Seismol. Soc. Am.* **102**, 6, 2491-2506, DOI: 10.1785/0120110344.

Dec, J., K. Pietsch, and P. Marzec (2011), Application of seismic methods to identify potential gas concentration zones at the Zechstein Limestone level in the "Rudna" mining area, SW Poland, *Ann. Soc. Geol. Pol.* **81**, 1, 63-78.

EMSC (2013), M 4.6 – Poland – 2013-03-19 21:09:52 UTC, Lastquake – the official EMSC earthquakes app, http://www.emsc-csem.org/Earthquake/earthquake.php?id=308921.

Fitch, T.J., D.W. McCowan, and M.W. Shields (1980), Estimation of seismic moment tensor from teleseismic body wave data with applications to intraplate and mantle earthquakes, *J. Geophys. Res.* **85**, B7, 3817-3828, DOI: 10.1029/JB085iB07p03817.

Gibowicz, S.J., and A. Kijko (1994), *An Introduction to Mining Seismology*, Academic Press, San Diego.

Gibowicz, S.J., A. Bober, A. Cichowicz, Z. Droste, Z. Dychtowicz, J. Hordejuk, M. Kazimierczyk, and A. Kijko (1979), Source study of the Lubin, Poland, tremor of 24 March 1977, *Acta Geophys. Pol.* **27**, 1, 3-38.

Gibowicz, S.J., J. Niewiadomski, P. Wiejacz, and B. Domański (1989), Source study of the Lubin, Poland, mine tremor of 20 June 1987, *Acta Geophys. Pol.* **37**, 2, 111-132.

Grünthal, G. (1998), *EMS-98: European Macroseismic Scale 1998*, Cahiers du Centre Européen de Géodynamique et de Séismologie, Vol. 15, Centre Européen de Géodynamique et de Séismologie, Luxembourg.

Guterch, B., and H. Lewandowska-Marciniak (2002), Seismicity and seismic hazard in Poland, *Fol. Quater*. **73**, 85-99.

Guterch, B., H. Lewandowska-Marciniak, and J. Niewiadomski (2005), Earthquakes recorded in Poland along the Pieniny Klippen Belt, Western Carpathians, *Acta Geophys. Pol.* **53**, 1, 27-45.

Hanks, T.C., and H. Kanamori (1979), A moment magnitude scale, *J. Geophys. Res.* **84**, B5, 2348-2350, DOI: 10.1029/JB084iB05p02348.

Haskell, N.A. (1953), The dispersion of surface waves on multilayered media, *Bull. Seismol. Soc. Am.* **43**, 1, 17-34.

Heimann, S. (2011), A robust method to estimate kinematic earthquake source parameters, Ph.D. Thesis, University of Hamburg, Hamburg, Germany.

KGHM (2013), KGHM Polska Miedź S.A., Rudna mine, http://www.kghm.pl/index.dhtml?category_id=265_24.06.2013&laang=en.

Koziarz, E. (2013), Personal communication.

Kwiatek, G. (2013), http://www.induced.pl/projekty/foci (in Polish).

Lasocki, S. (2005), Probabilistic analysis of seismic hazard posed by mining induced events. **In:** Y. Potvin, and M. Hudyma (eds.), *Proc. Sixth Int. Symp. on Rockburst and Seismicity in Mines "Controlling Seismic Risk", 9-11 March 2005, Australian Centre for Geomechanics, Nedlands, Australia*, 151-156.

Leśniak, A., and G. Pszczoła (2008), Combined mine tremors source location and error evaluation in the Lubin Copper Mine (Poland), *Tectonophysics* **456**, 1-2, 16-27, DOI: 10.1016/j.tecto.2007.04.012.

Lizurek, G., and P. Wiejacz (2011), Moment tensor solution and physical parameters of selected recent seismic events at Rudna Copper Mine. **In:** A.F. Idziak and R. Dubiel (eds.), *Geophysics in Mining and Environmental Protection*, Geoplanet: Earth and Planetary Sciences, Springer, Berlin Heidelberg, 11-19, DOI: 10.1007/978-3-642-19097-1_2.

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.

Madariaga, R. (1976), Dynamics of expanding circular fault, *Bull. Seismol. Soc. Am.* **66**, 3, 639-666.

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.

Medhurst, T., M. Bartlett, and R. Sliwa (2008), Effect of Grouting on Longwall Mining Through Faults. **In:** N. Aziz (ed.), *Coal 2008: Coal Operators' Conference, 14-15 February 2008, University of Wollongong and the Australasian Institute of Mining and Metallurgy, Australia*, 44-55.

Niewiadomski, J. (1997), Spectral analysis and seismic source parameters. **In:** A.J. Mendecki (ed.), *Seismic Monitoring in Mines*, Chapman & Hall, London, 144-158.

Rudziński, Ł., and W. Dębski (2011), Extending the double-difference location technique to mining applications. Part I: Numerical study, *Acta Geophys*. **59**, 4, 785-814, DOI: 10.2478/s11600-011-0021-5.

Tajduś, A., T. Majcherczyk, and M. Cała (1996), Effect of faults on rockbursts hazard. **In:** Z. Rakowski (ed.), *Geomechanics'96*, Balkema, Rotterdam.

Vavryčuk, V. (2005), Focal mechanisms in anisotropic media, *Geophys. J. Int.* **161**, 2, 334-346, DOI: 10.1111/j.1365-246X.2005.02585.x.

Wiejacz, P. (1992), Calculation of seismic moment tensor for mine tremors from the Legnica-Głogów Copper Basin, *Acta Geophys. Pol*. **40**, 2, 103-122.

Wiejacz, P., and W. Dębski (2009), Podhale, Poland, earthquake of November 30, 2004, *Acta Geophys.* **57**, 2, 346-366, DOI: 10.2478/s11600-009-0007-8.

Zuchiewicz, W., J. Badura, and M. Jarosiński (2007), Neotectonics of Poland: An overview of active faulting, *Stud. Quater.* **24**, 5-20.

Received 25 July 2013
Received in revised form 17 July 2014
Accepted 22 July2014

DE GRUYTER OPEN

Acta Geophysica
vol. 63, no. 1, Feb. 2015, pp. 201-213
DOI: 10.2478/s11600-014-0262-1

Detection of the Long Period Long Duration (LPLD) Events in Time- and Frequency-Domain

Anna KWIETNIAK

Academy of Science and Technology AGH,
Faculty of Geology, Geophysics and Environmental Protection,
Kraków, Poland; e-mail: kwietnia@agh.edu.pl

A b s t r a c t

Long Period Long Duration (LPLD) signals are unusual seismic events that can be observed during hydraulic fracturing. These events are very similar in appearance to tectonic tremors sequences, which were first observed in subduction zones. Their nature is not well known. LPLD might be related to the productivity of the reservoir. Different methods of the LPLD events' detection recorded during hydraulic fracturing are presented. The author applied two methods for LPLD detection – Butterworth filtering and Continuous Wavelet Transform (CWT). Additionally, a new approach to LPLD events detection – instantaneous seismic attributes – was used, common in a classical seismic interpretation but not in microseismic monitoring.

Key words: Long Period Long Duration (LPLD), microseismicity, microseismic monitoring, hydraulic fracturing, shale gas.

1. INTRODUCTION

The acronym LPLD comes from Long Period Long Duration events, which means that they last more than 10 s (from 10 to 100 s). LPLDs have significantly longer durations than microseismic events (usually from 0.1 to 1 s). Their frequency bandwidth is 10 to 80 Hz, but can be specified as lower than

80 Hz, due to the fact that it is hardly possible to obtain any information from frequencies below 10 Hz (geophones with natural frequency of 10-15 Hz are usually used). LPLD events were registered during hydraulic stimulation in various geological settings, and it is possible that the formation properties have a role to play in the frequency content of these signals (Das and Zoback 2013). According to previous studies dedicated to these unusual seismic signals (Das and Zoback 2011, 2013), they are similar in appearance to tectonic tremor sequences. Tremor sequences were first observed at active volcanoes and reflect the internal dynamics of the volcanic system (Chouet 1996). Long-period signals and tremors have the same temporal and spectral components, differing only in duration of events (Chouet 1996). The possible source of this kind of tremors is flow-inducted oscillation in channels transporting magmatic fluid. Nonvolcanic seismic tremors with long-duration seismic signals with no clear *P* or *S* waves were also observed in subduction zones. Their occurrence associated with subduction in southwest Japan as well as their potential shear slip mechanism have been particularly widely described (Obara 2002, Shelly *et al*. 2006, Nadeau and Guilhem 2009). Low-frequency earthquakes (long-period events) occur not only around active volcanoes, but also in different tectonic contexts compared with subduction zones near active fault systems. Nonvolcanic tremor sequences were observed after two strong earthquakes on the Cholame segment of the San Andreas Fault and near Monarch Peak, Lonoak, California (Nadeau and Guilhem 2009). Considering the long duration and the mobility of tremor activity, the generation of tremors may be related to the movement of fluid in the subduction zones (Obara 2002). Another possibility is that tremor is generated directly by slow shear slip on the plane interface, and under this hypothesis, tremor is the weak seismological signature of slip that is otherwise too slow to generate detectable seismic waves (Shelly *et al.* 2006).

The example of LPLD events, recorded during hydraulic fracturing is shown in Fig. 1 (Das and Zoback 2011); the data come from the fracturing experiment which took place in Barnett Shale reservoir. The events themselves are complex but coherent, which means that their appearance is visible from trace to trace (they are recorded by closely located receivers in monitoring wells), but they also show inner diversity (complexity). The fundamental criterion for classification is that *P* and *S* arrivals cannot be resolved. Sometimes small micro-seismicity (classical microseismic events) occurs within LPLD sequences, but whether they are causal or coincidental is not known. Originally, microtremors or simply tremors were observed in subduction zones and on the slopes of volcanoes. Their source was not defined and many hypotheses were proposed. LPLD events are similar to those signals from subduction zones but their frequency range is different. Micro-

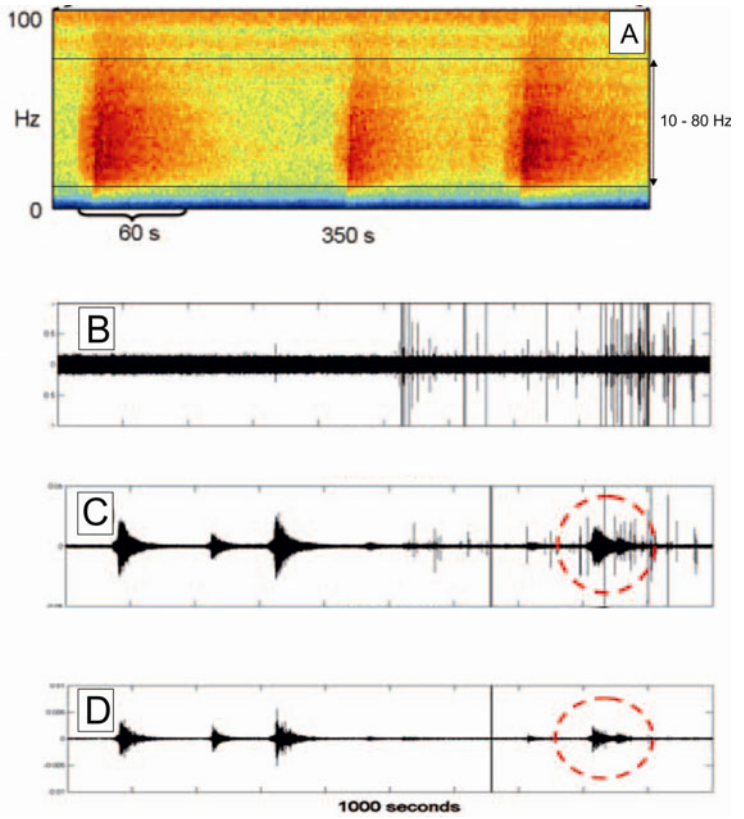


Fig. 1. (A) LPLD events observed below 100 Hz with the extended time scale, (B) original seismogram from stage 8 of simulfrac experiment, (C) signal after band-pass filtering from 10-80 Hz, and (D) signal obtained after wavelet decomposition and reconstruction. Red dotted circles in (C) and (D) show the difference between two methods (Das and Zoback 2011).

tremors are defined for low frequencies – to 10 Hz, whereas a much higher range is characteristic of LPLD events. The amplitudes of microtremors are very similar, displacements are in order of 10^{-4} to 10^{-2} mm (Okada 2003). Although they are very weak, they represent a source of noise for researchers earthquake seismology (Okada 2003). The most significant similarity is that both LPLDs and microtremors occur in swarm-like groups and are coherent (Das and Zoback 2011).

During stimulation of tight hydrocarbons through hydraulic fracturing, it is assumed that the volume of the reservoir corresponds with the volume where microseismicity occurs. Nowadays, microseismic monitoring in order to investigate deformations inside the reservoir records microearthquakes

that are related mostly to the brittle failure (Tary and van der Baan 2012). However, a comparison of orders of magnitude of the energy injected into the system with the energy recorded during microseismic monitoring does not account (Maxwell *et al.* 2009). Therefore, not the whole energy produced during hydraulic fracturing causes microseismic events. We can make an assumption, which is intuitive, that part of this energy is involved in different "activity of the reservoir". LPLD events can be understood as representation of another, very interesting process which is present during the stimulation of the shale or tight gas reservoirs.

One of the possible explanations of the LPLD events source is a reactivation of pre-existing faults (shear events), which have been formed naturally in the shale gas reservoirs. The main aim of hydraulic fracturing is to stimulate the productivity of the reservoir enhancing the permeability and creating additional flow paths for hydrocarbons. We cannot tell whether these flow paths are new tensile fractures created during hydraulic fracturing or reactivated, pre-existing features. It is commonly believed that new fractures are the main channels for reservoir fluids, but another hypothesis which is connected with pre-existing faults is also possible. Supposing that the productivity of the reservoir is at least partially correlated with the shearing on pre-existing features, and if the LPLD microseismic events are linked to this phenomenon, they will require further analysis.

The first step, which is the subject of this article, is to detect LPLD events and separate them from other microseismic signals. Once LPLD events are distinguished, a further analysis will be possible.

2. METHODS

The goal in microseismic data processing, in general, is to increase signal to noise ratio. There are many ways to achieve this, and one of them is analyzing the signal in joint time-frequency domain. The frequency domain is a good environment for signal analysis. The representation of time series in the frequency domain often illustrates many features difficult to visualize in the time domain. The manner in which the time series is mapped into the frequency domain determines the amount of new information that can be obtained (Chakraborty and Okaya 1995). The change of domains is possible through Fourier Transform. Apart from the basic Fourier Transform (performed with Fast Fourier algorithm), the decomposition of signal into its frequency components is possible with the use of different decomposition algorithms (Castagna *et al*. 2002, Castagna and Sun 2006, Kumar and Foufoula-Georgiou 1997, Chakraborty and Okaya 1995). Continuous Wavelet Transform is based on the concept of time-scale analysis and has two advantages over the others. Firstly, it keeps good resolution for high and low frequencies because the length of the window depends on the frequency of the

original signal. Secondly, frequency is proportional to scale (Castagna and Sun 2006) which means it contains the same information and its results can be freely comparable with common spectral analysis.

Downhole recorded microseismic data were pre-processed and a band-pass filter was applied. After filtering, three procedures were applied separately (each on the previously filtered data set). The first was instantaneous frequency (IF) which is related to the centroid of the power spectrum of the seismic wavelet (Taner 1992). The algorithm is based on the classical method presented in 1979 (Taner *et al*. 1979). The program can be found at: www.seismicunix.com/w/Suattribute. IF indicates where the "center of mass" of the spectrum is. IF illustrates lateral continuity of a waveform character and is independent of amplitude (Chung 1994). This property of IF is useful for LPLD detection in microseismic data. The output is a scalar value, so its results can be easily interpreted. Applying IF had mainly experimental value. This particular attribute was used, because instantaneous frequency should reveal changes in frequency, especially its increases. Also, the application of IF is straightforward and is not time-consuming.

Second and third procedures involved using CWT. In SeismicUnix this procedure can be performed in more than one way (Stockwell 2008). Two different approaches were used. CWT is especially useful for time series which are considered to be non-stationary. Non-stationarity means that all their parameters, including frequency, vary with time. Because of that fact, seismograms whose spectral content vary significantly with time required not-standard methods of decomposition. CWT provides such a method (Chakraborty and Okaya 1995, Castagna and Sun 2006). The mathematical formula for CWT is shown in the following equation:

$$C(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t)\varphi\left(\frac{t-b}{a}\right)dt \ .$$

The output C is a wavelet coefficient which is a function of scale (a) and translation (b). Translation can be understood as a position in time, changing from $t = 0$ to the end of the seismogram. $f(t)$ is the original seismic trace and $\varphi(t)$ is the so-called "mother wavelet", which can be understood as a second signal of different parameters (for instance, Ricker wavelet, Haar wavelet, *etc*.). Frequency is reversed to the scale: low scale a ("shrinked" wavelets) means compressed wavelets, hence, able to reveal rapidly changing details (high frequencies); high scale a means stretched wavelets, able to show slowly changing course features (low frequencies).

A choice of wavelet seems to be of paramount importance. Morlet wavelet is very popular for CWT because it is very similar to a seismic trace (Tary and van der Baan 2012), but other types are also used. In SeismicUnix

only three wavelets are defined: 2nd derivative of Gaussian function, 4th and 6th (which are commonly called "Mexican hat", "Witch's hat", and "Wizard's hat"). For computation "Mexican hat wavelet" was applied.

The first approach of applying CWT was based on the following parameters:

- wavelet type: "Mexican hat",
- number of decomposition levels: 201 (changing from –1 to 1 with step 0.01),
- filter base: 10.

The second approach:

- wavelet type: "Mexican hat",
- number of decomposition levels: 20 (changing from –1 to 1 with step 0.1),
- filter base: 2 (so-called "diadic scalling").

3. STUDY AREA

The data set consists of borehole data. Seismic monitoring was conducted in Barnett Shale reservoir in Texas; it was a different stimulation than that analyzed by Das and Zoback (2011, 2013). During the analysis, data from one stage of fracturing were closely studied and the procedure described earlier was applied to the whole record (approximately 5 hours of recording).

The target formation of the reservoir stimulation was Cline formation (the Lower Wolfcamp formation) in the Midland Basin. The Midland Basin is one of three sub-basins in the Permian Basin along with the Delaware and the Central Basin Platform. The lower Wolfcamp is of Pennsylvanian-aged (upper Carboniferous) organic rich shale with interbedded silt and sand. There is no direct information about natural fractures in the Lower Wolfcamp Formation, but we do have some information about fractures in the Spraberry which overlays the Wolfcamp Formation. The fractures of NE-SW directions were found in cores in the Spraberry Formation. The horizontal cores from sandstone-siltstone reservoirs in this formation have documented two systems of dramatically different yet dynamically compatible natural fractures (Lorenz *et al*. 2002).

Northeastward-directed Laramide compressive stress has been suggested to be the source of much, if not all, of the minimal, post-Permian structural deformation and fracturing in the Permian basin. The present-day stresses, significant in terms of fracture conductivity, are still generally aligned with this trend (Lorenz *et al*. 2002). It is possible that these fractures also exist in the Wolfcamp formation, but is not guaranteed. Moreover, the examination of vertical cores suggest that natural fractures in the Spraberry Formation are commonly extended by hydraulic processes associated with drilling and cor-

ing, obscuring the differentiation between natural and inducted fracture and complicating fracture interpretations (Lorenz *et al*. 2002).

4. RESULTS

In the data set from one stage of downhole monitoring, 5 LPLD-like events were interpreted. The author presents two of them. The results are shown in Figs. 2 and 3. On the left hand side at the top (Fig. 2A), there is one file after filtering. The file consists of 43 traces; their numbers are displayed on the horizontal scale. The first trace is the calibration trace, then 42 traces follow. Geophones which were used for the recording were 3 components (3C) with a nominal frequency of 15 Hz. As a result, every three traces is a record from one geophone only (vertical component and two horizontal components). In the middle (Fig. 2B), there is a single trace, number 29, which is vertical. On the right hand side at the top (Fig. 2C) there is corresponding instantaneous frequency applied to the file from the left. In Fig. 2D a decomposed signal is visible. It is a decomposition of 29th trace after applying the first approach. Each trace in Fig. 2D is one level of decomposition (so one frequency from the whole spectrum). The left hand side corresponds to lower scales and keeps information about high frequency content, whereas the right hand side about low frequencies. The area of interest is marked with the purple arrow (Fig. 2A and B). The length of this event is approximately 100 s and is visible after filtering as well as after applying IF and CWT. IF shows stronger pattern across the array, even for areas where filtering alone seems to have weaker results. CWT helps to investigate how amplitudes vary with frequency and time. It is also a good tool for determining coherency. The second approach of CWT (Fig. 2E), with the use of dyadic scaling, shows similar results. Moreover, it helps to check the level of inner complexity and variability of an LPLD signal. This property of CWT is very helpful in determining whether the signal can be named as an LPLD event by checking if *P* and *S* arrivals can or cannot be resolved. Figures 2A-E show that it is impossible to separate *P* and *S* arrivals, so the signal is most likely to be classified as an LPLD event. Comparing this signal to the signals recorded during hydraulic fracturing it has extraordinary duration and is quite unique.

In Figure 3 another LPLD-like event is presented. A typical pattern is visible after filtering (Fig. 3A), IF (Fig. 3C), and CWT (Fig. 3D and E). For decomposition also 29th trace was used. In this case, also two approaches of CWT are presented – one with 201 levels of decomposition (Fig. 3D) and the other with dyadic scaling and 20 levels of decomposition (Fig. 3E). The difference between these two approaches lies in the level of details that can be revealed. The decomposition with 201 levels provides more details, and more features that are present in the signal. The event presented in Fig. 3 lasts for approximately 20-30 s; it is much shorter than the one presented in

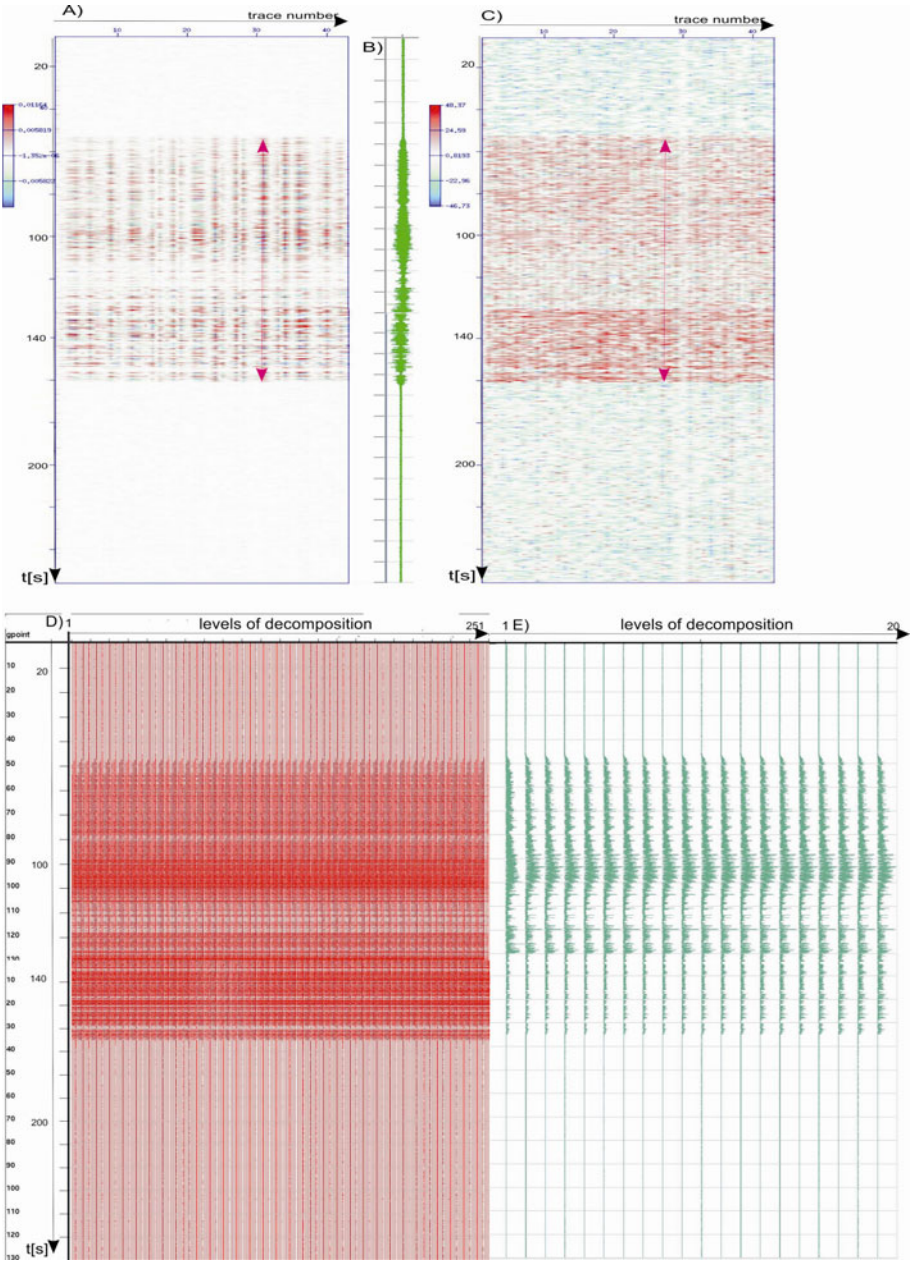


Fig. 2: (A) File after Butterworth band-pass filtering (filter parameters: (11|16|76|79)), (B) 29th trace (vertical component after the same filtering), (C) file after applying instantaneous frequency, (D) 1st approach of CWT, and (E) 2nd approach of CWT. The LPLD event is marked with the purple arrow – (A) and (C).

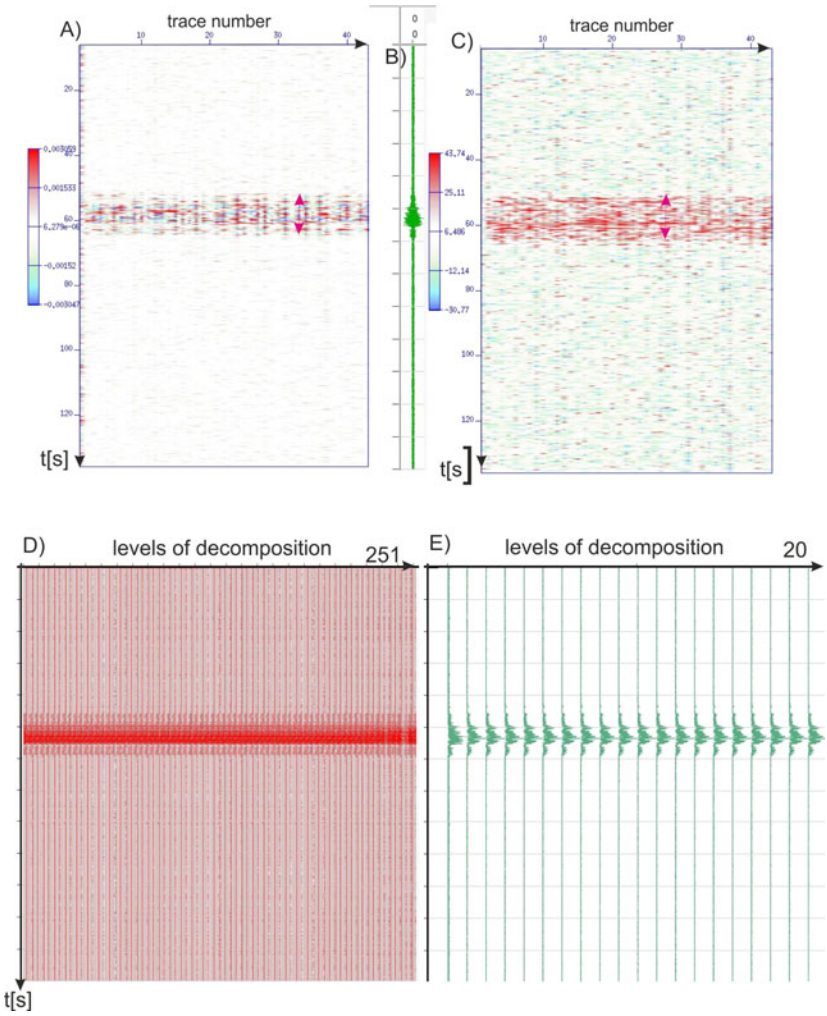


Fig. 3: (A) File after Butterworth bandpass filtering (filter parameters: (11|16|76|79)), (B) 29th trace (vertical component after the same filtering), (C) file after applying instantaneous frequency, (D) 1st approach of CWT, and (E) 2nd approach of CWT. The LPLD event is marked with the purple arrow – (A) and (C).

Fig. 2. It is visible clearly by applying Butterworth filtering; however, its beginning and end can be estimated more precisely for sections after instantaneous frequency and spectral decomposition. The event is coherent across the array and similar pattern is recorded on every trace. CWT is also a good tool for searching microseismic events. After applying CWT, the microseismic events are also visible as straight lines across the array.

5. DISCUSSION

An approach which was originally proposed by Das and Zoback (2011) was applied to the data but with one significant change. Only forward continuous wavelet transform was used, and no reconstruction of the signal was made. This method makes it possible to look carefully on the signal within an LPLD sequence. This approach seems to be valuable, especially if it is not fully understood how the mechanism of LPLD works and what the main source of these events is. CWT reveals many features, which are not necessarily visible in the time domain. Analyzing trace by trace after decomposition makes it possible to determine lateral variability of the event and its coherency. According to the author's opinion, forward CWT gives more information about the character of the event (compare Figs. 1D, 2D-E, and 3D-E). Moreover, by applying CWT it was possible to separate LPLD-like signals from those that at first looked like LPLD but in fact they were not. This was done by indicating *P* and *S* arrivals that were clearly visible after decomposition and not visible on signals after filtering or IF.

Instantaneous frequency is a common and basic attribute that is used almost in every 2D or 3D seismic interpretation, even though it is not popular among seismologists. Microseismic event detection and location are more close to seismology than to seismics, hence seismic attributes are not the "weapon of choice". Applying IF was treated as kind of experiment, which yielded relatively good results. LPLD-like events are characterized by sudden increase of IF. Hence, the source of LPLD events may be connected with the reactivation of preexisting faults, *i.e*., many microseismic events (micro-earthquakes) that take place on a fault plane. Nevertheless, IF results are not sufficient to determine the LPLD event mechanism. Their long durations as well as their broadband spectra give reasons to believe that these events are closer to shear events, like tectonic tremors (Shelly *et al*. 2007) than fluid-driven events, like volcanic tremors (Chouet 1996), even though fluids must play a role in their generation. In the overall tectonic context LPLD events and tremors are closely related and very difficult to distinguish one from the other (Chouet 1996). If a given explanation is relevant to LPLD event mechanism, we shall observe an increase in instantaneous frequency. Instantaneous frequency results for this case are considered to be good, but do not give many details about the nature of the LPLD signals. The only information is that a rapid increase in frequency is observed, so the application of IF is limited.

Time-frequency analyses are particularly well-suited for the study of long-duration, non-stationary phenomena that may take place inside the reservoir (Tary and van der Baan 2012). Applying different transforms which make it possible to analyze signal in a time-frequency domain can reveal

new phenomena. The one that may also be somehow connected to LPLD events occurrence is resonance. Resonance seems to be almost the basic kind of signal which exist naturally and, moreover, can be created during the process of injecting fluid into the reservoir. Resonance frequencies can be generated by source, path, or receiver effects. Each effect has to be considered separately to determine the origin of particular resonance frequency (Tary and van der Baan 2012). Even though geophones which are used for microseismic monitoring are deployed in a boreholes, the design of a borehole itself and the presence of geophones also produce resonance frequencies (Sun and McMehan 1988, St-Onge and Eaton 2011). Before reaching the receivers, resonance frequencies can also be produced along the raypath (Tary and van der Baan 2012). This is connected, *e.g.*, with interference of waves, multiple wave scattering, influence of the near surface, low velocity layer. Additionally, resonance can be created by fractures and flowing fluid themselves. This resonance is connected with source-side effects. Because the source of LPLD events has not been clearly defined yet, it is possible that they are related to the resonance which is present in the reservoir. However, resonance is not very likely to be the direct source of LPLD events. LPLD events are characterized by broad-band spectra, whereas resonance is usually characterized by narrow-band spectra. It probably indicates that the causal mechanism is different in these two cases; nevertheless, the phenomena may still be related somehow. To test this idea, further analyses of their inner variability have to be applied. As can be seen from the examples, CWT helps to study the LPLD events structure, so it might be a good tool for distinguishing the source of these events.

The integration of microseismic monitoring with seismic data seems to be an overriding matter. Such analyses are not widely used because of high costs, and also because it is not common for gas companies to share their seismic data with a monitoring company (usually not the same). The advancement in seismic analyses, and the usefulness of various seismic attributes, might contribute greatly to the analysis of the signal recorded during hydraulic fracturing. In the case of LPLD events such a concept may be extremely useful, especially if their sources are pre-existing faults. Such a research would be of paramount importance.

6. CONCLUSIONS

Instanteous frequency seams to give reasonable results and its interpretation is straightforward and easy. Additionally, the computation of IF is very quick and the algorithm is well-described (Taner *et al*. 1979).

CWT results are dependent on the number of levels of decomposition. More levels reveal much more details and a more accurate analysis is possible. Microseismic events are visible after decomposition. Run times (compu-

tational costs) are extremely time-consuming and computation of CWT requires a lot of memory. The advantage of CWT is also the fact that the robust extraction of the signal of a certain scale (defined frequency) is possible. Moreover, while interpreting results of CWT we can study the whole range of frequencies at the same time.

The results of both, IF and CWT, are consistent and show a similar pattern across the array. Analyzing all traces at the same time gives significant information about the coherency of the signal.

The mechanism of recorded LPLD events is not known. The future work should focus on localizing these signals. Moreover, an attempt to integrate their localizations with seismic data may give valuable information. The author's work was dedicated exclusively to determining the occurrence of LPLD events in a given dataset.

Acknowledgments. The work was part of the study dedicated to LPLDs detection and location in Microseismic.INC. The author wants to express her deepest gratitude to the Development Department of Microseismic.INC, especially to Mike Thornton, Mike Mueller, Julia Kurpan, Chven Mitchell, and Mary Ellison. With their help and scientific support this work benefited a lot. Special thanks go to Dr. Peter Duncan for making this internship possible. The author would like to express appreciation to the anonymous reviewer whose constructive comments helped improved greatly the quality of the paper.

References

Castagna, J.P., and S. Sun (2006), Comparison of spectral decomposition methods, *First Break* **24**, 3, 75-79.

Castagna, J.P., S. Sun, and R.W. Siegfried (2002), The use of spectral decomposition as a hydrocarbon indicator, *GasTIPS*, Summer, 24-27.

Chakraborty, A., and D. Okaya (1995), Frequency-time decomposition of seismic data using wavelet-based methods, *Geophysics* **60**, 6, 1906-1916, DOI: 10.1190/1.1443922.

Chouet, B.A. (1996), Long-period volcano seismicity: its source and use in eruption forecasting, *Nature* **380**, 6572, 309-316, DOI: 10.1038/380309a0.

Chung, H.-M. (1994), Seismic properties of thin beds, Ph.D. Thesis, The University of Calgary, Calgary, Canada, 97-99.

Das, I., and M.D. Zoback (2011), Long-period, long-duration seismic events during hydraulic fracture stimulation of a shale gas reservoir, *The Leading Edge* **30**, 7, 778-786, DOI: 10.1190/1.3609093.

Das, I., and M.D. Zoback (2013), Long-period, long-duration seismic events during hydraulic stimulation of shale and tight-gas reservoirs – Part 1: Waveform characteristics, *Geophysics* **78**, 6, 97-108, DOI: 10.1190/geo2013-0164.1.

Kumar, P., and E. Foufoula-Georgiou (1997), Wavelet analysis for geophysical applications, *Rev. Geophys.* **35**, 4, 385-412, DOI: 10.1029/97RG00427.

Lorenz, J.C., J.L. Sterling, D.S. Schechter, C.L. Whigham, and J.L. Jensen (2002), Natural fractures in the Spraberry Formation, Midland basin, Texas: The effects of mechanical stratigraphy on fracture variability and reservoir behavior, *AAPG Bull.* **86**, 3, 505-524.

Maxwell, S.C., J. Shemeta, E. Campbell, and D. Quirk (2009), Detection of microseismic events using a surface receiver network. **In:** *Second EAGE Passive Seismic Workshop – Exploration and Monitoring Applications 2009*, DOI: 10.3997/2214-4609.20146738.

Nadeau, R.M., and A. Guilhem (2009), Nonvolcanic tremor evolution and the San Simeon and Parkfield, California, earthquakes, *Science* **325**, 5937, 191-193, DOI: 10.1126/science.1174155.

Obara, K. (2002), Nonvolcanic deep tremor associated with subduction in southwest Japan, *Science* **296**, 5573, 1679-1681, DOI: 10.1126/science.1070378.

Okada, H. (2003), *The Microtremor Survey Method*, Geophysical Monograph Series, Vol. 12, Society of Exploration Geophysicists, Tulsa, USA.

Shelly, D.R., G.C. Beroza, S. Ide, and S. Nakamula (2006), Low-frequency earthquakes in Shikoku, Japan, and their relationship to episodic tremor and slip, *Nature* **442**, 7099, 188-191, DOI: 10.1038/ nature04931.

Shelly, D.R., G.C. Beroza, and S. Ide (2007), Non-volcanic tremor and low-frequency earthquake swarns, *Nature* **446**, 305-307, DOI: 10.1038/nature05666.

Stockwell, J. (2008), Complete listing of CWP free program self-documentations, Documentation generated by shell script GENDOCS, Center of Wave Phenomena, Colorado School of Mines, Golden, USA, 147-148.

St-Onge, A., and D.W. Eaton (2011), Noise examples from two microseismic datasets, *CSEG Recorder* **36**, 8, 46-49.

Sun, R., and G.A. McMechan (1988), Finite-difference modeling of borehole resonances, *Energy Sources* **10**, 1, 55-75, DOI: 10.1080/00908318808908916.

Taner, M.T. (1992), Attributes revisited, Rock Solid Images, Houston, Texas, http//wvwv.rocksolidimaqes.com.

Taner, M.T., F. Koehler, and R.E. Sheriff (1979), Complex seismic trace analysis, *Geophysics* **44**, 6, 1041-1063, DOI: 10.1190/1.1440994.

Tary, J.B., and M. van der Baan (2012), Potential use of resonance frequencies in microseismic interpretation, *The Leading Edge* **31**, 11, 1338-1346, DOI: 10.1190/tle31111338.1.

Received 2 March 2014
Received in revised form 30 September 2014
Accepted 3 October 2014

DE GRUYTER OPEN

Acta Geophysica
vol. 63, no. 1, Feb. 2015, pp. 214-231
DOI: 10.2478/s11600-014-0244-3

Pollutant's Horizontal Dispersion Along and Against Sinusoidally Varying Velocity from a Pulse Type Point Source

Mritunjay K. SINGH[1], Nav K. MAHATO[2], and Naveen KUMAR[3]

[1]Department of Applied Mathematics, Indian School of Mines, Dhanbad, India
e-mail: drmks29@rediffmail.com (corresponding author)

[2]Department of Mathematics, C.V. Raman College of Engineering, Bhubaneswar, India; e-mail: kr.nav81@gmail.com

[3]Department of Mathematics, Banaras Hindu University, Varanasi, India
e-mail: nks_1953@yahoo.co.in

A b s t r a c t

An analytical solution of a two-dimensional advection diffusion equation with time dependent coefficients is obtained by using Laplace Integral Transformation Technique. The horizontal medium of solute transport is considered of semi-infinite extent along both the longitudinal and lateral directions. The input concentration is assumed at an intermediate position of the domain. It helps to evaluate concentration level along the flow as well as against the flow through one model only. The source of the input concentration is considered to be of pulse type. In the presence of the source, it is assumed to be decreasing very slowly with time, and just after the elimination of the source it is assumed to be zero. The dispersion coefficient and the advection parameter are considered directly proportional to each other. The analytical solution may be used to predict the solute concentration level with position and time in an open medium as well as in a porous medium. The effect of heterogeneity on the solute transport may also be predicted.

Keywords: solute transport, aquifer, input concentration, heterogeneity, Cartesian system.

1. INTRODUCTION

Pollutant mass originating from a variety of natural and anthropogenic sources (volcano, industries, factories, refineries, sewage systems, garbage disposal sites, mines, *etc.)* is a major cause of degradation of the environment of air, surface water, soil, and groundwater. Mathematical modelers use the advection diffusion equation (ADE) to describe the concentration level of the pollutant mass at different position and time, away from its source, through its analytical and numerical solutions. In one-dimensional Cartesian system of coordinates, ADE is as follows

$$\frac{\partial c}{\partial t} = \frac{\partial}{\partial x}\left(D\frac{\partial c}{\partial x} - uc \right) , \tag{1}$$

where c $[ML^{-3}]$ denotes the solute concentration in the medium at any time $t[T]$, and at a position $x[L]$; $D[L^2T^{-1}]$ and u $[LT^{-1}]$ represent the dispersion coefficient and the advection parameter, respectively, along the longitudinal direction (x axis) of the medium. There are basic theories which relate the dispersion coefficient D and velocity u occurring as the two coefficients in the one-dimensional ADE (Eq. 1): (i) Ebach and White (1958) and Bear (1972) in their one-dimensional analysis suggested that D is proportional to u, and (ii) Taylor (1953) in his one-dimensional analysis obtained D proportional to u^2. Further on, Scheidegger (1957) summarized his analysis on the two possible relationships between D and u according to the role played by molecular diffusion: (i) $D \sim \alpha u^2$, where α, a constant of the porous medium alone (dynamic dispersivity), is derived by a dynamic procedure applicable when there is enough time in each flow channel for appreciable mixing to take place by molecular transverse diffusion; and (ii) $D \sim \beta u^2$, where β, another constant of the porous medium (geometric dispersivity), is derived by a geometric procedure applicable where there is no appreciable molecular transverse diffusion from one streamline into another. Later on, Freeze and Cherry (1979) have modified these dispersion theories by showing that if the dispersion parameter is proportional to the n-th power of the velocity, then n ranges between 1 and 2.

Dispersion in porous media was significantly discussed by Fried and Combarnous (1971) and groundwater pollution was explored by Fried (1975). The occurrence of anomalous diffusion behavior of an upscaled dispersion coefficient, for a transport process driven locally by a normal diffusion was explained by Matheron and De Marsily (1980). Solute transport in heterogeneous porous formations was discussed by Dagan (1984). An anomalous diffusion model of the local dispersion was also considered for porous media with special properties (Suciu 2014). The apparent behavior of the

dispersion at larger scales was discussed by Fried (1975), Matheron and De Marsily (1980), Dagan (1984), Suciu (2010, 2014).

The source of advective-diffusive solute mass transport may be a point source, *e.g.*, garbage disposal sites, mines, *etc*., a line source, *e.g.*, interface of sea water in aquifer or a surface source, *e.g.*, along agriculture field with high doses of chemical fertilizers (Marshal *et al.* 1996). A point source may be of continuous type or pulse type. In either case, the point source may be uniform or of varying nature. In the latter type, the input concentration may be uniform or of increasing nature in the presence of its source. As soon as the source of pollution is eliminated, the input concentration becomes zero or starts decreasing. Solution of an ADE for a pulse type point source is useful for predicting the rehabilitation time period of a polluted domain once its source is eliminated. Smokes coming out of a chimney, wastes from a drainage system reaching a particular location in rivers, lakes, are examples of uniform pulse-type point sources. As soon as the source is eliminated, the input becomes zero. Infiltration from the ground surface point sources (garbage disposal sites), reaches groundwater level or oil reservoirs, degrading their quality (an example of varying pulse-type point source), where as soon as the source is eliminated, the input, starts decreasing, instead of becoming zero. Only in particular cases it is possible to solve ADE analytically. In a more general situation, numerical techniques are required. The literature presenting most of the analytical methods has been reviewed in a recent work by Guerrero *et al.* (2009). Many analytical solutions of Eq. 1 in ideal conditions (the two coefficients are independent of the position and time) with growth and decay terms, subject to various initial and boundary conditions in semi-infinite or finite media have been complied (van Genuchten and Alves 1982, Javandel *et al*. 1984, Domenico and Schwartz 1997). Predicting the fate of pollutants in natural environments, such as rivers and man-made channels, is one of the major concerns. Approximate solutions first began to appear (Banks and Jerasate 1962, Warrick *et al.* 1971), in which a time-dependent dispersion coefficient was used. Usually, the solute transport models assume a constant dispersion coefficient that is calibrated separately for each different downstream sample location, resulting in different dispersion coefficients for the same flow problem. In an attempt to overcome this, the dispersion coefficient was considered as a function of the mean travel distance successfully (Pickens and Grisak 1981). Another approach is to model dispersivity as a time-dependent function. Based on the observation of numerical results (Suresh Kumar *et al.* 2008), dispersivity has been suggested to have a time-dependent behavior which reaches asymptotic values after a long time. Additionally, worth mentioning are some of the works solving ADE with variable coefficients analytically in one-dimension (Sander and Braddock 2005, Singh *et al.* 2008, Chen and Liu 2011, Kumar

et al. 2011, Chen *et al.* 2012a, b) and in two- and three-dimensions (Wilson and Miller 1978, Carnahan and Remer 1984, Goltz and Roberts 1986, Yates 1988, Batu 1989, 1993; Leij *et al.* 1991, Serrano 1995, Aral and Liao 1996, Zou *et al.* 1996, Zoppou and Knight 1997, Tartakovasky 2000, Singh *et al.* 2010, Jaiswal *et al.* 2011, Sudicky *et al.* 2013).

In the tropical regions, like Indian sub-continent, the water table in the rivers and in the aquifers and its flow is the minimum before the rainy season and is the maximum after the rainy season. But the pollution reaching the river water table through the waste water drainage system or reaching aquifer through infiltration from garbage disposal sites, mines, remain uniform throughout the year. The dispersion of pollutants down the stream and against the stream have a lateral component too, though much weaker in comparison of the longitudinal part. The present study formulates this scenario through a two-dimensional ADE, an initial condition, and boundary conditions, one of which being a pulse type input concentration (the concentration of pollutant at the meeting point in the water table). It is solved using Laplace integral transformation technique. A sinusoidal form of temporally dependent velocity is considered. It represents the minimum and maximum velocity in aquifers as well as in river beds, during the months of summer season (June) and peak of the winter season (December), respectively, in successive years. Of the three dispersion theories discussed in the above paragraph, the dispersion coefficient is considered to be directly proportional to the velocity along both directions. It makes the coefficients of the ADE time-dependent. Such dependence may be used to describe the heterogeneity of the medium and its effect on the solute dispersion. It is valid in porous domain like aquifer and on a surface water body like rivers and lakes. The longitudinal and lateral directions are considered to extend up to infinity. In the presence of the source, the input concentration (due to the source) is considered time-dependent of decreasing nature. The source of pollution is meeting the water medium at a point. To study its concentration distribution behavior along the flow (that is, on the right side domain of the point source), and against the flow (that is, in the domain on the left side), its location is not considered at the origin of the medium as taken in most of the works. It is assumed at an intermediate position of the medium nearer to the origin. In this way, concentration distribution in both situations is studied through one model only. The time-dependent coefficients of the ADE are reduced into constant coefficients by modulating the time variable which takes care of the time-dependent sinusoidal expression. Further on, the two space variables are unified into one space variable and it reduces the two-dimensional ADE into one-dimension.

2. MATHEMATICAL FORMULATION AND ANALYTICAL SOLUTION

The two-dimensional ADE with time dependent coefficients may be written as follows:

$$\frac{\partial c}{\partial t} = D_x(t)\frac{\partial^2 c}{\partial x^2} + D_y(t)\frac{\partial^2 c}{\partial y^2} - u(t)\frac{\partial c}{\partial x} - v(t)\frac{\partial c}{\partial y}. \tag{2}$$

The coefficients of second order derivatives are diffusion coefficients and those of first order space derivatives are velocity components along longitudinal and lateral directions, respectively. Let u and v be expressed as

$$u = u_0 V(mt) \qquad \text{and} \qquad v = v_0 V(mt), \tag{3}$$

where $m[T^{-1}]$ may be referred to as unsteady parameter, and u_0 and v_0 may be referred to as initial values of u and v, respectively. We consider the sinusoidal form of velocity

$$V(mt) = 1 - \sin mt. \tag{4}$$

The different time-dependent forms of velocity expression may also be considered similar to Aral and Liao (1996), based on the properties of algebraic sigmoid function which include the error function. It starts a progress from small beginning, accelerates in the rainy season and reaches up to a limit over a period of time. As the dispersion coefficient is considered directly proportional to velocity, we write

$$D_x = D_{x_0} V(mt) \quad \text{and} \quad D_y = D_{y_0} V(mt), \tag{5}$$

where $D_{x_0} = au_0$ and $D_{y_0} = av_0$ are the initial values of D_x and D_y, respectively. Here, a is the dispersivity $[L]$ that depends upon the pore size and geometry. As a result, Eq. 2 may be written as

$$\frac{1}{V(mt)}\frac{\partial c}{\partial t} = D_{x_0}\frac{\partial^2 c}{\partial x^2} + D_{y_0}\frac{\partial^2 c}{\partial y^2} - u_0\frac{\partial c}{\partial x} - v_0\frac{\partial c}{\partial y}. \tag{6}$$

It is assumed that before the introduction of input concentration, the domain is not solute free. It is uniformly polluted though its level is considered very low. The initial condition is

$$c(x, y, t) = c_i; \quad 0 \le x < \infty, \quad 0 \le y < \infty, \quad t = 0. \tag{7}$$

The pulse type input concentration at an intermediate location is defined as

$$c(x, y, t) = \left.\begin{array}{ll} c_0\left[1 + \exp(-qt)\right]; & 0 < t \le t_0 \\ 0; & t > t_0 \end{array}\right\} \qquad \text{at} \quad x = x_0, \quad y = y_0, \tag{8}$$

where $t = t_0$ is the time when the source of the input concentration is removed forever, and $q[T^{-1}]$ is a decay parameter. The second boundary condition at the two extreme ends of horizontal medium is considered of homogeneous flux type, as

$$\frac{\partial c}{\partial x} = 0 \quad \text{and} \quad \frac{\partial c}{\partial y} = 0\,; \quad x \to \infty\,, \quad y \to \infty\,, \quad t \geq 0\,. \tag{9}$$

Using an integral transformation (Crank 1975)

$$T^* = \int_0^t V(mt)dt\,, \tag{10}$$

the left hand side term of Eq. 6 becomes $(\partial c/\partial T^*)$. T^* is a new time variable, as m (of dimension being inverse of that of the time variable) will appear in the denominator of its expression. $V(mt)$ is so chosen that this new time variable becomes zero for $t = 0$; otherwise, the nature of the initial condition in new time domain will alter. It may be verified for its expression in Eq. 4. Further coordinate transformations

$$X = x - x_0\,, \quad Y = y - y_0 \tag{11}$$

convert the location $(x = x_0,\ y = y_0)$ of the input concentration to $(X = 0,\ Y = 0)$. Also Eq. 6 may be obtained in the form

$$\frac{\partial c}{\partial T^*} = D_{x_0}\frac{\partial^2 c}{\partial X^2} + D_{y_0}\frac{\partial^2 c}{\partial Y^2} - u_0\frac{\partial c}{\partial X} - v_0\frac{\partial c}{\partial Y}\,. \tag{12}$$

Similarly, the conditions given by Eqs. 7-9 may be written in new independent variables (X, Y, T^*). Further, using another coordinate transformation

$$z = X + Y\sqrt{\frac{D_{y_0}}{D_{x_0}}} \quad \text{or} \quad X + Y\sqrt{\frac{v_0}{u_0}}\,, \tag{13}$$

the two dimensional Eq. 12 is reduced into a one-dimensional equation

$$\frac{\partial c}{\partial T^*} = D\frac{\partial^2 c}{\partial z^2} - U\frac{\partial c}{\partial z}\,, \tag{14}$$

where $D = D_{x_0}\left(1 + \dfrac{D_{y_0}^2}{D_{x_0}^2}\right)$ and $U = \left(u_0 + v_0\sqrt{\dfrac{v_0}{u_0}}\right)$. Equation 13 is similar to the one used by Carnahan and Remer (1984). The point $(X = 0,\ Y = 0)$ corre-

sponds to ($z = 0$). The initial and boundary conditions, Eqs. 7-9, will assume the respective form as

$$c\left(z,T^*\right) = c_i\,; \quad z \geq 0\,,\ T^* = 0\,, \tag{15}$$

$$c\left(z,T^*\right) = \left.\begin{matrix} c_0(2-qT^*)\,; & 0 < T^* \leq T_0^* \\ 0\,; & T^* > T_0^* \end{matrix}\right\}; \quad z = 0\,, \tag{16}$$

$$\frac{\partial c}{\partial z} = 0\,; \quad z \to \infty\,, \quad T^* \geq 0\,. \tag{17}$$

As the coefficients of ADE in (z, T^*) domain are constants so Laplace integral transformation technique may be used to get the analytical solution of the initial and boundary value problem defined by Eqs. 14-17 (van Genuchten and Alves 1982), as (see the Appendix)

$$c(z,T^*) = c_i + \exp\left(\frac{Uz}{2D} - \frac{U^2T^*}{4D}\right)\left[\left(2c_0 - c_i\right)F\left(z,T^*\right) - qc_0G\left(z,T^*\right)\right];$$
$$0 < T^* \leq T_0^* \tag{18a}$$

$$c(z,T^*) = c_i + \exp\left(\frac{Uz}{2D} - \frac{U^2T^*}{4D}\right)\left[c_0\left(2+qT_0^*\right)\left\{F\left(z,T^*\right) - F\left(z,T^* - T_0^*\right)\right\}\right.$$
$$\left. - c_iF\left(z,T^*\right) - qc_0\left\{G\left(z,T^*\right) - G\left(z,T^* - T_0^*\right)\right\}\right]; \qquad T > T_0^* \tag{18b}$$

where

$$F\left(z,T^*\right) = \frac{1}{2}\exp\left(\frac{U^2T^*}{4D} - \frac{Uz}{2D}\right)\operatorname{erfc}\left(\frac{z}{2\sqrt{DT^*}} - \frac{UT^*}{2\sqrt{DT^*}}\right)$$
$$+ \frac{1}{2}\exp\left(\frac{U^2T^*}{4D} + \frac{Uz}{2D}\right)\operatorname{erfc}\left(\frac{z}{2\sqrt{DT^*}} + \frac{UT^*}{2\sqrt{DT^*}}\right), \tag{18c}$$

$$G\left(z,T^*\right) = \frac{1}{2U}\left(UT^* - z\right)\exp\left(\frac{U^2T^*}{4D} - \frac{Uz}{2D}\right)\operatorname{erfc}\left(\frac{z}{2\sqrt{DT^*}} - \frac{UT^*}{2\sqrt{DT^*}}\right)$$
$$+ \frac{1}{2U}\left(UT^* + z\right)\exp\left(\frac{U^2T^*}{4D} + \frac{Uz}{2D}\right)\operatorname{erfc}\left(\frac{z}{2\sqrt{DT^*}} + \frac{UT^*}{2\sqrt{DT^*}}\right). \tag{18d}$$

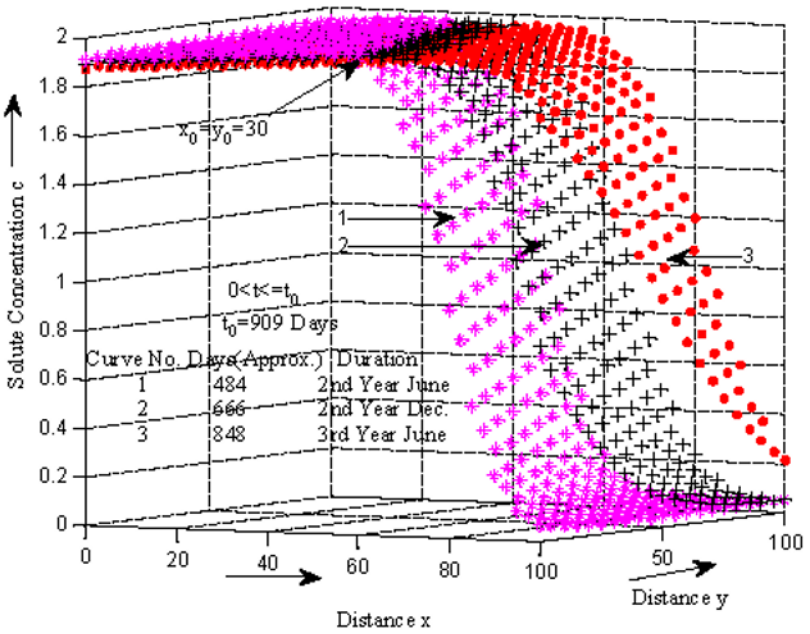
One can express this solution in terms of the original independent variables (x, y, t) by using the appropriate transformation equations introduced at the different stages. By using Eqs. 4 and 10, the new time variable will have the expression

$$T^* = \frac{1}{m}\left[mt - \left(1 - \cos mt\right)\right]. \tag{19}$$

3. RESULTS AND DISCUSSION

To illustrate the analytical solutions (Eq. 18a, b), the values of different parameters are chosen from the literature cited in the Introduction as: initial pollution level parameter $c_i = 0.01$, reference concentration and the decay parameter, both in the input condition in Eq. 8, are chosen as $c_0 = 1.0$, $q = 0.0001(\text{day})^{-1}$, respectively, the velocity components $u_0 = 0.09$ km/day, $v_0 = 0.009$ km/day, and the dispersion coefficient components $D_{x_0} = 0.1$ km^2/day, $D_{y_0} = 0.01$ km^2/day. The concentration values are evaluated in the domain $0 \le z$, y [km] ≤ 100. The location of the point source (input concentration) is assumed at $x_0 = 30$ km, $y_0 = 30$ km. Figure 1a depicts the concentration values obtained from the solution in Eq. 18a in the presence of the source, *i.e.*, in the time domain $t < t_0$ for $mt = 8, 11, 14$. The unsteady parameter is chosen as $m = 0.0165\ (\text{day})^{-1}$. This value yields $t = 484$, 666, and 848 days . From Eq. 4 it may be observed that the velocity is alternatively minimum and maximum at these values of time; hence, they may be regarded as the time periods during the months of June of the second year, December of the same year and again June of the next year, from the introduction of the input concentration at $t = 0$. The source of the input concentration is assumed to be eliminated at $t_0 = 909$ days, which corresponds to $mt = 15$. Beyond this time, the concentration values evaluated from solution in Eq. 18b are depicted in Fig. 1b, at $t = 1030$, 1212, and 1394 days, which corresponds to $mt = 17$, 20, 23, respectively, and represents the months of December of the third year, June and December of the next year, respectively. This scenario of velocity distribution either in aquifers or in rivers occurs in the tropical regions like India. Figure 1a shows that concentration values increase with position and time on either side of the point source (input concentration location), though the increase is much less in the left domain, that is, towards the origin ($x = 0$, $y = 0$) in comparison of that at the same distances on the right domain, that is, away from the origin. Figure 1b shows that, after the removal of the source of pollution, the concentration values at a particular position (x, y) in the already polluted domain decrease with time. The domain nearer to the location of the source gets rehabilitated faster than the positions far away. To have more clarity on the solute mass distribution pattern originating from the point source, the concentration values in both time domains (before and after the elimination of the source) are also provided in Tables 1 and 2, respectively. It may be observed that according to the pulse type condition given in Eq. 8, the input concentration evaluated

(a)



(b)

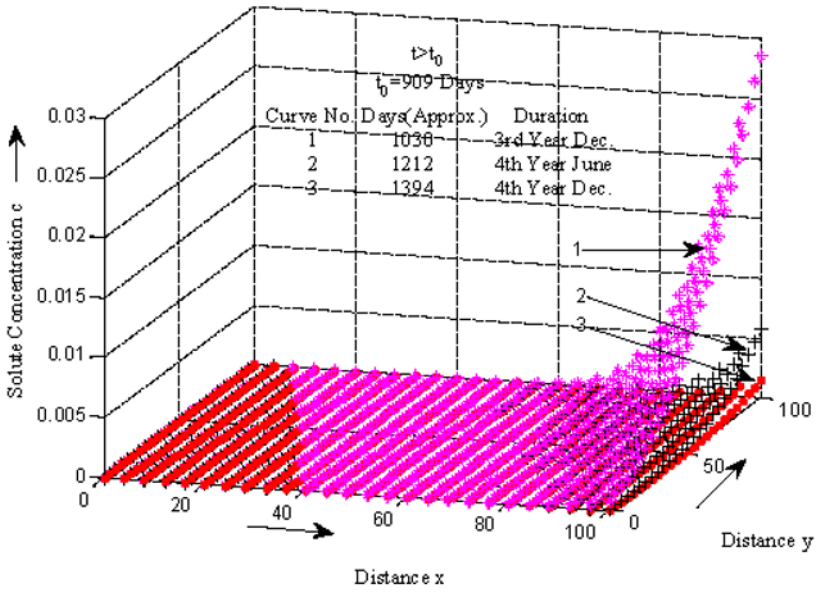


Fig.1. Solute mass horizontal dispersion pattern along and against the sinusoidally varying velocity: (a) in the presence of a point source, and (b) in the absence of a point source.

Table 1

Concentration values for sinusoidal form of velocity in the presence of the source, that is, in the time domain $0 < t < t_0$, $t_0 = 909$ days

t [days]	Distance	$y = 0$	$y = 10$	$y = 20$	$y = 30$	$y = 40$	$y = 50$	$y = 60$	$y = 70$	$y = 80$	$y = 90$	$y = 100$
484	$x = 0$	1.9159	1.9193	1.9227	1.9261	1.9296	1.9330	1.9364	1.9398	1.9432	1.9466	1.9500
	$x = 10$	1.9267	1.9301	1.9335	1.9369	1.9403	1.9437	1.9471	1.9505	1.9539	1.9574	1.9608
	$x = 20$	1.9375	1.9409	1.9443	1.9477	1.9511	1.9545	1.9579	1.9613	1.9646	1.9677	1.9698
	$x = 30$	1.9482	1.9516	1.9551	1.9585	1.9618	1.9652	1.9681	1.9700	1.9686	1.9595	1.9342
	$x = 40$	1.9590	1.9624	1.9657	1.9686	1.9701	1.9679	1.9568	1.9277	1.8675	1.7614	1.5989
	$x = 50$	1.9689	1.9701	1.9669	1.9536	1.9205	1.8538	1.7390	1.5670	1.3405	1.0777	0.8083
	$x = 60$	1.9499	1.9123	1.8388	1.7151	1.5336	1.2996	1.0335	0.7662	0.5285	0.3414	0.2109
	$x = 70$	1.6896	1.4989	1.2579	0.9894	0.7248	0.4943	0.3164	0.1947	0.1210	0.0813	0.0624
	$x = 80$	0.9453	0.6845	0.4616	0.2929	0.1798	0.1126	0.0771	0.0606	0.0537	0.0511	0.0503
	$x = 90$	0.2709	0.1661	0.1050	0.0735	0.0590	0.0531	0.0509	0.0503	0.0501	0.0500	0.0500
	$x = 100$	0.0702	0.0576	0.0526	0.0508	0.0502	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
666	$x = 0$	1.8968	1.9002	1.9036	1.9071	1.9105	1.9139	1.9173	1.9207	1.9241	1.9275	1.9309
	$x = 10$	1.9076	1.9110	1.9144	1.9178	1.9212	1.9246	1.9280	1.9314	1.9349	1.9383	1.9417
	$x = 20$	1.9184	1.9218	1.9252	1.9286	1.9320	1.9354	1.9388	1.9422	1.9456	1.9490	1.9524
	$x = 30$	1.9291	1.9326	1.9360	1.9394	1.9428	1.9462	1.9496	1.9530	1.9563	1.9595	1.9622
	$x = 40$	1.9399	1.9433	1.9467	1.9501	1.9535	1.9568	1.9599	1.9625	1.9637	1.9620	1.9545
	$x = 50$	1.9507	1.9541	1.9573	1.9604	1.9628	1.9637	1.9613	1.9524	1.9318	1.8927	1.8266
	$x = 60$	1.9609	1.9631	1.9636	1.9604	1.9500	1.9270	1.8840	1.8128	1.7060	1.5598	1.3766
	$x = 70$	1.9593	1.9473	1.9216	1.8746	1.7980	1.6850	1.5324	1.3439	1.1300	0.9070	0.6932
	$x = 80$	1.8644	1.7823	1.6629	1.5041	1.3105	1.0940	0.8712	0.6605	0.4772	0.3303	0.2220
	$x = 90$	1.4748	1.2766	1.0578	0.8358	0.6286	0.4507	0.3102	0.2079	0.1394	0.0971	0.0731
	$x = 100$	0.8007	0.5974	0.4253	0.2910	0.1947	0.1309	0.0922	0.0704	0.0592	0.0538	0.0515
848	$x = 0$	1.8779	1.8813	1.8847	1.8881	1.8915	1.8949	1.8983	1.9017	1.9051	1.9085	1.9119
	$x = 10$	1.8886	1.8920	1.8954	1.8988	1.9022	1.9057	1.9091	1.9125	1.9159	1.9193	1.9227
	$x = 20$	1.8994	1.9028	1.9062	1.9096	1.9130	1.9164	1.9198	1.9232	1.9266	1.9300	1.9335
	$x = 30$	1.9102	1.9136	1.9170	1.9204	1.9238	1.9272	1.9306	1.9340	1.9374	1.9408	1.9442
	$x = 40$	1.9209	1.9243	1.9277	1.9312	1.9346	1.9380	1.9414	1.9448	1.9481	1.9514	1.9544
	$x = 50$	1.9317	1.9351	1.9385	1.9419	1.9453	1.9487	1.9519	1.9549	1.9573	1.9586	1.9575
	$x = 60$	1.9425	1.9458	1.9492	1.9524	1.9553	1.9576	1.9586	1.9570	1.9507	1.9366	1.9103
	$x = 70$	1.9529	1.9558	1.9579	1.9586	1.9563	1.9490	1.9333	1.9046	1.8578	1.7873	1.6888
	$x = 80$	1.9584	1.9556	1.9471	1.9296	1.8984	1.8481	1.7733	1.6700	1.5368	1.3760	1.1941
	$x = 90$	1.9256	1.8917	1.8377	1.7585	1.6504	1.5124	1.3476	1.1632	0.9697	0.7791	0.6030
	$x = 100$	1.7430	1.6300	1.4874	1.3188	1.1321	0.9383	0.7493	0.5764	0.4280	0.3084	0.2177

Table 2

Concentration values for sinusoidal form of velocity after the elimination of the source, that is, in the time domain $t > t_0$, $t_0 = 909$ days

t [days]	Distance	$y = 0$	$y = 10$	$y = 20$	$y = 30$	$y = 40$	$y = 50$	$y = 60$	$y = 70$	$y = 80$	$y = 90$	$y = 100$
1030	$x = 0$	0	0	0	0	0	0	0	0	0	0	0
	$x = 10$	0	0	0	0	0	0	0	0	0	0	5.4e-13
	$x = 20$	0	0	0	0	0	0	0	8.56e-13	7.51e-12	4.42e-11	2.1e-10
	$x = 30$	0	0	0	0	1.25e-12	1.02e-11	5.78e-11	2.77e-10	1.18e-09	4.62e-09	1.6e-08
	$x = 40$	6.32e-14	1.8e-12	1.37e-11	7.52e-11	3.53e-10	1.48e-09	5.72e-09	2.04e-08	6.8e-08	2.1e-07	6.2e-07
	$x = 50$	9.75e-11	4.48e-10	1.86e-09	7.07e-09	2.49e-08	8.24e-08	2.54e-07	7.42e-07	2.03e-06	5.29e-06	1.29e-05
	$x = 60$	8.72e-09	3.04e-08	9.94e-08	3.04e-07	8.78e-07	2.38e-06	6.14e-06	1.49e-05	3.45e-05	7.55e-05	1.56e-04
	$x = 70$	3.63e-07	1.03e-06	2.79e-06	7.12e-06	1.71e-05	3.93e-05	8.53e-05	1.75e-04	3.44e-04	6.41e-04	0.0011
	$x = 80$	8.25e-06	1.97e-05	4.47e-05	9.63e-05	1.96e-04	3.82e-04	7.06e-04	0.0012	0.0021	0.0033	0.0051
	$x = 90$	1.08e-04	2.19e-04	4.23e-04	7.77e-04	0.0014	0.0023	0.0036	0.0055	0.0079	0.0110	0.0147
	$x = 100$	8.53e-04	0.0015	0.0024	0.0039	0.0058	0.0084	0.0116	0.0154	0.0196	0.0242	0.0287
1212	$x = 0$	0	0	0	0	0	0	0	0	0	0	0
	$x = 10$	0	0	0	0	0	0	0	0	0	0	3.46e-15
	$x = 20$	0	0	0	0	0	0	0	5.41e-15	4.81e-14	2.18e-13	1.46e-12
	$x = 30$	0	0	0	0	8.13e-15	6.54e-14	3.79e-13	1.87e-12	8.36e-12	3.43e-11	1.31e-10
	$x = 40$	3.95e-16	1.18e-14	8.84e-14	4.96e-13	2.40e-12	1.05e-11	4.29e-11	1.62e-10	5.82e-10	1.96e-09	6.31e-09
	$x = 50$	6.46e-13	3.08e-12	1.33e-11	5.35e-11	2.01e-10	7.12e-10	2.38e-09	7.59e-09	2.29e-08	6.63e-08	1.82e-07
	$x = 60$	6.66e-11	2.47e-10	8.69e-10	2.88e-09	9.11e-09	2.73e-08	7.84e-08	2.14e-07	5.59e-07	1.39e-06	3.34e-06
	$x = 70$	3.49e-09	1.09e-08	3.25e-08	9.26e-08	2.51e-07	6.51e-07	1.61e-06	3.83e-06	8.71e-06	1.89e-05	3.96e-05
	$x = 80$	1.09e-07	2.94e-07	7.57e-07	1.86e-06	4.39e-06	9.92e-06	2.14e-05	4.45e-05	8.86e-05	1.69e-04	3.10e-04
	$x = 90$	2.14e-06	5.02e-06	1.12e-05	2.42e-05	4.99e-05	9.87e-05	1.87e-04	3.41e-04	5.98e-04	0.0010	0.0016
	$x = 100$	2.73e-05	5.59e-05	1.09e-04	2.07e-04	3.75e-04	6.53e-04	0.0011	0.0018	0.0027	0.0041	0.0059
1394	$x = 0$	0	0	0	0	0	0	0	0	0	0	0
	$x = 10$	0	0	0	0	0	0	0	0	0	0	2.05e-16
	$x = 20$	0	0	0	0	0	0	0	3.28e-16	2.92e-15	1.76e-14	9.05e-14
	$x = 30$	0	0	0	0	4.94e-16	3.97e-15	2.32e-14	1.16e-13	5.28e-13	2.21e-12	8.71e-12
	$x = 40$	2.54e-17	7.07e-16	5.37e-15	3.04e-14	1.50e-13	6.70e-13	2.78e-12	1.08e-11	3.97e-11	1.38e-10	4.62e-10
	$x = 50$	3.97e-14	1.92e-13	8.48e-13	3.48e-12	1.34e-11	4.89e-11	1.69e-10	5.59e-10	1.76e-09	5.31e-09	1.53e-08
	$x = 60$	4.34e-12	1.66e-11	6.00e-11	2.06e-10	6.75e-10	2.11e-09	6.33e-09	1.81e-08	4.99e-08	1.31e-07	3.34e-07
	$x = 70$	2.50e-10	8.15e-10	2.53e-09	7.53e-09	2.14e-08	5.86e-08	1.53e-07	3.87e-07	9.38e-07	2.18e-06	4.89e-06
	$x = 80$	8.95e-09	2.53e-08	6.87e-08	1.79e-07	4.48e-07	1.07e-06	2.49e-06	5.55e-06	1.18e-05	2.45e-05	4.87e-05
	$x = 90$	2.08e-07	5.18e-07	1.23e-06	2.84e-06	6.30e-06	1.34e-05	2.74e-05	5.42e-05	1.03e-04	1.89e-04	3.35e-04
	$x = 100$	3.25e-06	7.14e-06	1.51e-05	3.07e-05	6.04e-05	1.14e-04	2.08e-04	3.66e-04	6.21e-04	0.0010	0.0016

from the solutions in Eq. 18a, b decreases very slowly with time in the former time domain and is zero in the latter time domain.

The new time variable T^* has been computed from Eq. 19 for the values of old time variable t considered above, and unsteady parameter $m = 0.0165$ (day)$^{-1}$. We have $T^* = 415$, 606, 796, 802, 953, 1176, 1301 days, for $t = 484$, 666, 848, 909, 1030, 1212, and 1394, respectively. It is evident that a value of new time variable is much less than the respective value of the old time variable. It means the concentration level at a position (x, y) is obtained much earlier in case of sinusoidal flow domain than in case of uniform flow domain. The value of m decides upon the periodicity, as in the present example it is almost half of a year. Similarly, for $m = 0.01$ (day)$^{-1}$, at $mt = 0.05$, and 1.5, the non-dimensional value of the velocity from Eq. 4 will be 0.95 and 0.002, respectively. In other words, the periodicity in this case occurs at the interval of approximately 15 days. The new time variable obtained from Eq. 19 will again be less than the respective old time variable in this set too. In this way Eq. 18a, b may be used for sinusoidal form of velocity of different periodicity.

In real cases, the current through the medium is seldom unidirectional; hence, the solute mass disperses both longitudinally as well as laterally, though that along the lateral direction may be much less than that in the longitudinal direction. In this regard, the values of the lateral components of velocity as well as of dispersion coefficient are considered much less (one-tenth) of the respective longitudinal values but it is evident from the results that the concentration values along the lateral direction are significant. These results are compared with those of a one-dimensional model (Kumar and Kumar 1998) with the same assumptions. The concentration values on the same distance from the origin along the longitudinal direction have been found less in the two-dimensional model than those in the one-dimensional model. The solution given in Eq. 18a, b may also be used for other temporal dependent expressions of $V(mt)$, for example, for exponentially increasing $\exp(mt)$ or decreasing function $\exp(-mt)$. Choosing the value of m appropriately, the pattern of time-dependence may be varied according to the situation and need. The results delivered from the analytical solution (Eq. 18a, b) are validated by the results obtained by solving the same initial and boundary value problem by a two level explicit method under the stability criterion. A very good agreement between the analytical and numerical results has been found.

The present dispersion problem has much deviation from idealistic conditions and incorporates some features nearer to real situations. For example, in the real world nowadays the dispersion of pollutants along different media of air, water, and soil has become a matter of great concern and there is a need to study such a problem through the solutions of advection diffusion

equation. In a real situation, getting numerical solution is an easier way and cost effective but its validation by comparing it with an analytical solution of a similar problem is equally important. Therefore, it demands more and more analytical solutions of real problems and the present work is an effort in that direction. Analytical solutions have another advantage of providing better insight into the factors included in the problem, making it more realistic than any other solution. Most of the analytical solutions are approximate solutions, of which the Laplace transformation technique, as used in the present work, is the most viable method. In comparison to numerical models, analytical models and their solution provide more physical insight into conceptual mathematical behavior of the system. The obtained analytical solution is a benchmark to develop numerical codes and solutions. For practical application in subsurface hydrology, the local dispersion coefficients alone (eventually without molecular diffusion) may be used to a zeroth-order approximation of transport in mildly heterogeneous aquifers, consisting of a diffusion in the mean velocity field (Suciu 2010). A perfectly homogeneous aquifer is a highly idealized approximation of little practical relevance. The fitted diffusion coefficient larger than the local dispersion coefficient is only due to the scale effects induced by the spatial variability of the velocity field (Suciu 2014).

4. CONCLUSIONS

The main motive of the present work is to assess the concentration level in the vicinity as well as at far distances on both sides of the point source of a pollution mass, from one analytical solution. Another one was to use temporal dependence of dispersion coefficients in a two dimensional ADE while obtaining its analytical solution through Laplace Integral Transformation Technique. Such dependence may take care of the heterogeneity of the medium too, in addition to representing the unsteadiness behavior of the diffusing pollutant. It is always better to assess the pollution level more realistically by a two dimensional model than a one dimensional model. The solution obtained is based on the dispersion theory (Scheidegger 1957) valid in the water medium on the surface and in the aquifer.

Acknowledgment. The authors are thankful to the editors and reviewers for their constructive comments and valuable suggestions, which have helped improve the quality of the paper.

Appendix

Using the following transformation

$$c(z,T^*) = K(z,T^*)\exp\left(\frac{Uz}{2D} - \frac{U^2T^*}{4D}\right), \tag{A1}$$

in Eq. 14 and the conditions given in Eqs. 15-17, we get a diffusive problem in new dependent variable as

$$\frac{\partial K}{\partial T^*} = D\frac{\partial^2 K}{\partial z^2}, \tag{A2}$$

$$K\left(z,T^*\right) = c_i \exp\left(-\frac{Uz}{2D}\right); \quad z \geq 0, \quad T^* = 0, \tag{A3}$$

$$K\left(z,T^*\right) = \left.\begin{array}{ll} c_0\left(2 - qT^*\right)\exp\left(\dfrac{U^2T^*}{4D}\right); & 0 < T^* \leq T_0^* \\ 0; & T^* > T_0^* \end{array}\right\}; \quad z = 0, \tag{A4}$$

$$\frac{\partial K}{\partial z} + \frac{UK}{2D} = 0; \quad z \to \infty, \quad T^* \geq 0. \tag{A5}$$

Applying Laplace integral transformation in Eqs. A2-A5, we may get the solution $\bar{K}(z,p)$ as follows:

$$\begin{aligned}\bar{K}\left(z,p\right) = {} & c_0\left[2\bar{K}_1\left(z,p\right) + qT_0^*\bar{K}_2\left(z,p\right) - q\bar{K}_3\left(z,p\right)\right] \\ & -c_i\bar{K}_4\left(z,p\right) + \frac{c_i}{\left(p - \dfrac{U^2}{4D}\right)}\exp\left(-\frac{Uz}{2D}\right),\end{aligned} \tag{A6}$$

where

$$\bar{K}_1\left(z,p\right) = \frac{1}{\left(p - \dfrac{U^2}{4D}\right)}\left[1 - \exp\left\{-\left(p - \frac{U^2}{4D}\right)T_0^*\right\}\right]\exp\left(-\sqrt{\frac{p}{D}}\,z\right), \tag{A7}$$

$$\bar{K}_2\left(z,p\right) = \frac{1}{\left(p - \dfrac{U^2}{4D}\right)}\exp\left\{-\left(p - \frac{U^2}{4D}\right)T_0^*\right\}\exp\left(-\sqrt{\frac{p}{D}}\,z\right), \tag{A8}$$

$$\bar{K}_3(z,p)=\frac{1}{\left(p-\frac{U^2}{4D}\right)^2}\left[1-\exp\left\{-\left(p-\frac{U^2}{4D}\right)T_0^*\right\}\right]\exp\left(-\sqrt{\frac{p}{D}}z\right), \tag{A9}$$

and

$$\bar{K}_4(z,p)=\frac{1}{\left(p-\frac{U^2}{4D}\right)}\exp\left(-\sqrt{\frac{p}{D}}z\right). \tag{A10}$$

Now taking the inverse Laplace transform of Eq. A6, the solution $K(z, T^*)$can be obtained as follows:

$$\begin{aligned}K(z,T^*)=c_0\left[2K_1(z,T^*)+qT_0^*K_2(z,T^*)-qK_3(z,T^*)\right]\\ -c_iK_4(z,T^*)+\frac{c_i}{\left(p-\frac{U^2}{4D}\right)}\exp\left(-\frac{Uz}{2D}\right),\end{aligned} \tag{A11}$$

where

$$K_1(z,T^*)=\begin{cases}F(z,T^*); & 0<T^*\le T_0^*\\ F(z,T^*)-F(z,T^*-T_0^*); & T^*>T_0^*\end{cases}, \tag{A12}$$

$$K_2(z,T^*)=\begin{cases}0; & 0<T^*\le T_0^*\\ F(z,T^*)-F(z,T^*-T_0^*); & T^*>T_0^*\end{cases}, \tag{A13}$$

$$K_3(z,T^*)=\begin{cases}G(z,T^*); & 0<T^*\le T_0^*\\ G(z,T^*)-G(z,T^*-T_0^*); & T^*>T_0^*\end{cases}, \tag{A14}$$

and

$$K_4(z,T^*)=F(z,T^*). \tag{A15}$$

Applying the transformation A1 on this solution, the desired solution given in Eq. 18 may be obtained.

References

Aral, M.M., and B. Liao (1996), Analytical solutions for two-dimensional transport equation with time-dependent dispersion coefficients, *J. Hydrol. Eng.* **1**, 1, 20-32, DOI: 10.1061/(ASCE)1084-0699(1996)1:1(20).

Banks, R.B., and S.J. Jerasate (1962), Dispersion in unsteady porous media flow, *J. Hydraul. Div*. **88**, 1-21.

Batu, V. (1989), A generalized two-dimensional analytical solution for hydrodynamic dispersion in bounded media with the first-type boundary condition at the source, *Water Resour. Res.* **25**, 6, 1125-1132, DOI: 10.1029/WR025i006p01125.

Batu, V. (1993), A generalized two-dimensional analytical solute transport model in bounded media for flux-type finite multiple sources, *Water Resour. Res.* **29**, 8, 2881-2892, DOI: 10.1029/93WR00977.

Bear, J. (1972), *Dynamics of Fluids in Porous Media*, Elsevier, New York.

Carnahan, C.L., and J.S. Remer (1984), Non-equilibrium and equilibrium sorption with a linear sorption isotherm during mass transport through an infinite porous media: Some analytical solutions, *J. Hydrol.* **73**, 3-4, 227-258, DOI: 10.1016/0022-1694(84)90002-7.

Chen, J.-S., and C.-W. Liu (2011), Generalized analytical solution for advection-dispersion equation in finite spatial domain with arbitrary time-dependent inlet boundary condition, *Hydrol. Earth Syst. Sci.* **15**, 8, 2471-2479, DOI: 10.5194/hess-15-2471-2011.

Chen, J.-S., K.-H. Lai, C.-W. Liu, and C.-F. Ni (2012a), A novel method for analytically solving multi-species advective-dispersive transport equations sequentially coupled with first-order decay reactions, *J. Hydrol.* **420-421**, 191-204, DOI: 10.1016/j.jhydrol.2011.12.001.

Chen, J.-S., C.-W. Liu, C.-P. Liang, and K.-H Lai (2012b), Generalized analytical solutions to sequentially coupled multi-species advective-dispersive transport equations in a finite domain subject to an arbitrary time-dependent source boundary condition, *J. Hydrol.* **456-457**, 101-109, DOI: 10.1016/j.jhydrol.2012.06.017.

Crank, J. (1975), *The Mathematics of Diffusion*, Oxford Univ. Press, Oxford.

Dagan, G. (1984), Solute transport in heterogeneous porous formations, *J. Fluid Mech.* **145**, 151-177, DOI: 10.1017/S0022112084002858.

Domenico, P.A., and F.W. Schwartz (1997), *Physical and Chemical Hydrogeology*, John Wiley & Sons, New York.

Ebach, E.A., and R.R. White (1958), Mixing of fluids flowing through beds of packed solids, *AIChE J.* **4**, 2, 161-169, DOI: 10.1002/aic.690040209.

Freeze, R.A., and J.A. Cherry (1979), *Groundwater*, Prentice-Hall, New Jersey.

Fried, J.J. (1975), *Groundwater Pollution*, Developments in Water Science, Vol. 4, Elsevier, Amsterdam.

Fried, J.J., and M.A. Combarnous (1971), Dispersion in porous media. **In:** V.T. Chow (ed.), *Advances in Hydroscience*, Vol. 7, 169-282, Academic Press, New York.

Goltz, M.N., and P.V. Roberts (1986), Three-dimensional solutions for solute transport in an infinite medium with mobile and immobile zones, *Water Resour. Res.* **22**, 7, 1139-1148, DOI: 10.1029/WR022i007p01139.

Guerrero, J.S.P., L.C.G. Pimentel, T.H. Skaggs, and M.Th. van Genuchten (2009), Analytical solution of the advection-diffusion transport equation using a change-of-variable and integral transform technique, *Int. J. Heat. Mass. Transfer* **52**, 13-14, 3297-3304, DOI: 10.1016/j.ijheatmasstransfer. 2009.02.002.

Jaiswal, D.K., A. Kumar, N. Kumar, and M.K. Singh (2011), Solute transport along temporally and spatially dependent flows through horizontal semi-infinite media: Dispersion proportional to square of velocity, *J. Hydrol. Eng.* **16**, 3, 228-238, DOI: 10.1061/(ASCE)HE.1943-5584.0000312.

Javandel, I., C. Doughty, and C.F. Tasang (1984), *Groundwater Transport: Handbook of Mathematical Models,* Water Resources Monogr., Vol. 10, AGU, Washington DC.

Kumar, A., D.K. Jaiswal, and R.R. Yadav (2011), One-dimensional solute transport for uniform and varying pulse type input point source with temporally dependent coefficients in longitudinal semi-infinite homogeneous porous domain, *Int. J. Math. Sci. Comput.* **1**, 2, 56-66.

Kumar, N., and M. Kumar (1998), Solute dispersion along unsteady groundwater flow in a semi-infinite aquifer, *Hydrol. Earth Syst. Sci.* **2**, 1, 93-100, DOI: 10.5194/hess-2-93-1998.

Leij, F.J., T.H. Skaggs, and M.Th. van Genuchten (1991), Analytical solutions for solute transport in three-dimensional semi-infinite porous media, *Water Resour. Res.* **27**, 10, 2719-2733, DOI: 10.1029/91WR01912.

Marshal, T.J., J.W. Holmes, and C.W. Rose (1996), *Soil Physics*, 3rd ed., Cambridge University Press, Cambridge.

Matheron, G., and G. De Marsily (1980), Is transport in porous media always diffusive? A counterexample, *Water Resour. Res.* **16**, 5, 901-917, DOI: 10.1029/ WR016i005p00901.

Pickens, J.F., and G.E. Grisak (1981), Scale-dependent dispersion in a stratified gramular aquifer, *Water Resour. Res.* **17**, 4, 1191-1211, DOI: 10.1029/ WR017i004p01191.

Sander, G.C., and R.D. Braddock (2005), Analytical solutions to the transient, unsaturated transport of water and contaminants through horizontal porous media, *Adv. Water Resour.* **28**, 10, 1102-1111, DOI: 10.1016/j.advwatres. 2004.10.010.

Scheidegger, A. (1957), *The Physics of Flow Through Porous Medi*a, Univ. of Toronto Press, Toronto.

Serrano, S.E. (1995), Forecasting scale-dependent dispersion from spills in heterogeneous aquifers, *J. Hydrol.* **169**, 1-4, 151-169, DOI: 10.1016/0022-1694 (94)02663-V.

Singh, M.K., N.K. Mahato, and P. Singh (2008), Longitudinal dispersion with time-dependent source concentration in semi-infinite aquifer, *J. Earth Syst. Sci.* **117**, 6, 945-949, DOI: 10.1007/s12040-008-0079-x.

Singh, M.K., P. Singh, and V.P. Singh (2010), Analytical solution for two-dimensional solute transport in finite aquifer with time-dependent source concentration, *J. Eng. Mech.* **136**, 10, 1309-1315, DOI: 10.1061/(ASCE)EM.1943-7889.0000177.

Suciu, N. (2010), Spatially inhomogeneous transition probabilities as memory effects for diffusion in statistically homogeneous random velocity fields, *Phys. Rev. E* **81**, 5, 056301, DOI: 10.1103/PhysRevE.81.056301.

Suciu, N. (2014), Diffusion in random velocity fields with applications to contaminant transport in groundwater, *Adv. Water Resour.* **69**, 114-133, DOI: 10.1016/j.advwatres.2014.04.002.

Sudicky, E.A., H.-T. Hwang, W.A. Illman, Y.-S. Wu, J.B. Kool, and P. Huyakorn (2013), A semi-analytical solution for simulating contaminant transport subject to chain-decay reactions, *J. Contam. Hydrol*. **144**, 1, 20-45, DOI: 10.1016/j.jconhyd.2012.10.001.

Suresh Kumar, G., M. Sekhar, and D. Misra (2008), Time-dependent dispersivity of linearly sorbing solutes in a single fracture with matrix diffusion, *J. Hydrol. Eng.* **13**, 4, 250-257, DOI: 10.1061/(ASCE)1084-0699(2008)13:4(250).

Tartakovsky, D.M. (2000), An analytical solution for two-dimensional contaminant transport during groundwater extraction, *J. Contam. Hydrol.* **42**, 2-4, 273-283, DOI: 10.1016/S0169-7722(99)00086-8.

Taylor, G. (1953), Dispersion of soluble matter in solvent flowing slowly through a tube, *Proc. R. Soc. London A* **219**, 1137, 186-203, DOI: 10.1098/rspa.1953.0139.

van Genuchten, M.Th., and W.J. Alves (1982), Analytical solutions of the one-dimensional convective-dispersion solute transport equation, Tech. Bull. No. 1661, US Department of Agriculture, Washington D.C.

Warrick, A.W., J.W. Biggar, and D.R. Nielsen (1971), Simultaneous solute and water transfer for an unsaturated soil, *Water Resour. Res.* **7**, 5, 1216-1225, DOI: 10.1029/WR007i005p01216.

Wilson, J.L., and P.J. Miller (1978), Two-dimensional plume in uniform groundwater flow, *J. Hydraul. Div.* **104**, 4, 503-514.

Yates, S.R. (1988), Three-dimensional radial dispersion in a variable velocity flow field, *Water Resour. Res.* **24**, 7, 1083-1090, DOI: 10.1029/WR024i007p01083.

Zoppou, C., and J.H. Knight (1997), Analytical solutions for advection and advection-diffusion equations with spatially variable coefficients, *ASCE J. Hydraul. Eng*. **123**, 2, 144-148, DOI: 10.1061/(ASCE)0733-9429(1997) 123:2(144).

Zoua, S., J. Ma, and A.D. Koussis (1996), Analytical solutions to non-Fickian subsurface dispersion in uniform groundwater flow, *J. Hydrol.* **179**, 1-4, 237-258, DOI: 10.1016/0022-1694(95)02830-7.

Received 12 February 2014
Received in revised form 23 May 2014
Accepted 9 June 2014

DE GRUYTER OPEN

Acta Geophysica
vol. 63, no. 1, Feb. 2015, pp. 232-274
DOI: 10.2478/s11600-014-0242-5

Development of Local IDF-formula Using Controlled Random Search Method for Global Optimization

Katarzyna WEINEROWSKA-BORDS

Faculty of Civil and Environmental Engineering, Gdańsk University of Technology, Gdańsk, Poland; e-mail: kwein@pg.gda.pl

A b s t r a c t

The aim of the study is to present the effective and relatively simple empirical approach to rainfall intensity-duration-frequency-formulas development, based on Controlled Random Search (CRS) for global optimization. The approach is mainly dedicated to the cases in which the commonly used IDF-relationships do not provide satisfactory fit between simulations and observations, and more complex formulas with higher number of parameters are advisable. Precipitation data from Gdańsk gauge station were analyzed as the example, with use of peak-over-threshold method and Chomicz scale for rainfall intensity. General forms of the IDF-function were chosen and the parameter calibration with use of CRS algorithm was developed. The compliance of the obtained IDF-formulas with precipitation data and the efficiency of the algorithm were analyzed. The study confirmed the proposed empirical approach may be an interesting alternative for probabilistic ones, especially when IDF-relationship has more complex form and precipitation data do not match "typical" hydrological distributions.

Key words: precipitation; data analysis; IDF formulas; optimization.

1. INTRODUCTION

Intensity-duration-frequency (IDF) analysis of extreme precipitation is an important part of hydrological study for many engineering problems. The re-

sults of such analysis, usually expressed by IDF-relations (IDF-formulas) and/or their graphical form – IDF-curves, enable to determine the so-called "design storms" and thus provide essential "input" information for many practical problems, such as: designing of drainage and irrigation systems, designing of sewer conduits, engineering device dimensioning (*e.g.*, dikes, culverts, *etc.*), flood protection, urban impact analysis, water quality management, and many others (*e.g.*, Arnell 1982, Koutsoyiannis *et al.* 1998, Akan and Houghtalen 2003, Elsebaie 2012, Walesh 1989, and others). Despite the fact that, thanks to the significant development of measurement and computing technology, much more advanced hydrological analyses of more sophisticated nature are available (*e.g.*, Molnar and Burlando 2005, Licznar *et al.* 2011a, b; Pui *et al.* 2012), the IDF-formulas, although "simple" and "basic", still pay an important role in hydrological applications (*e.g.*, Willems 2000, Vaes *et al.* 2001, Svensson *et al.* 2007, Venkata Ramana *et al.* 2008, Endreny and Imbeah 2009, El-Sayed 2011, Elsebaie 2012, Ariff *et al.* 2012, Hailegeorgis *et al.* 2013, *etc.*). They become particularly useful in analyzes concerning urban areas. The specificity of such basins expresses itself mainly in relatively small watershed areas (Schilling 1991) with diversified catchment characteristics and factors increasing rapidness and intensity of direct runoff (such as: low roughness, high share of impermeable surfaces, high slopes, existence of artificial conduits and open channel regulation, *etc.*; McCuen 2005 and others) and in local urban phenomena influencing rainfall characteristics (*e.g.*, urban heat island effect; Lutgens and Tarbuck 2004, Delleur 2003). These latter factors cause relatively large temporal and spatial variability of precipitation, which – in addition to other characteristic features of urban basins – entails a need of its more detailed analysis with higher spatial resolution (Schilling 1991, Mutzner 1991, and others).

The most common purpose of IDF-formulas development is to create the "source" for obtaining representative design rainfall hyetographs, which constitutes essential input information for many engineering problems. In many practical cases, such representative hyetographs take form of the "block rainfalls", very useful for the simplest hydrological runoff estimations, *e.g.*, based on rational method or its modifications (Chow 1964, Nagy 1991, Willems 2000, Akan and Houghtalen 2003, Ben-Zvi 2009, Elsebaie 2012). Furthermore, the IDF-formulas may serve also as a basis for the development of synthetic time-variant design storms (storm patterns), that are essential for more complicated hydrological rainfall-runoff models (of conceptual or hydrodynamic character) and for other studies (Licznar and Łomotowski 2007). Such approach is applied, *i.e.*, to determine synthetic hyetographs with use of Huff's, Yen and Chow's, Chicago, SCS, and other methods (Arnell 1982, Vaes *et al.* 2001, Akan and Houghtalen 2003, Grimaldi and Serinaldi 2006). The example of more sophisticated applica-

tions of IDF-relations is testing and calibrating of rainfall generators (Willems 2000). IDF-formulas may be also helpful in synoptic meteorological studies and hydraulic design.

2. THEORETICAL BACKGROUND

IDF-formulas express the relationship between mean precipitation intensity I [mm/min] (or q [$dm^3/(s \cdot ha)$]) and its frequency of occurrence c [1/year] (or return period T [years], or exceedence probability p [%]) for different time intervals of rainfall duration t_d [min] (also called "aggregation-levels"). Due to the fact that in most of engineering problems the "extreme" values of rainfall intensity are essential (and thus computationally and designingly authoritative), the development of IDF-formulas requires the analysis of the largest extreme rainfall episodes. The basis of such analysis consists of the events of tempestuous and torrential rainfalls recorded during continuous observations in sufficiently long period. The length of historical rainfall data series must be vast enough to enable statistical recognition of temporal variability of precipitation, both annual and seasonal. The longer the return period T is considered (and thus – the smaller the frequency of rainfall occurrence c), the longer the period of rainfall observations is necessary. Thus, the needed length of the historical rainfall data series is strongly dependent on the aim of calculations (type of engineering problem), but usually not shorter than 10 years. The rainfalls observed over this time should constitute the set of data of independent and homogeneous character. In practice such conditions are very often impossible to fulfill.

2.1 Homogeneity and independence of the sample

Full homogeneity and independence of the sample is very difficult to verify and demonstrate, although there are mathematical tools developed to recognize these features of rainfall data sample, *e.g.*, autocorrelation analysis (Grace and Eagleson 1967, Wenzel and Voorhees 1978, Arnell 1982). Independence of the rainfall episodes is connected to their duration, the duration of rainless intervals and their succession in time. Although it is possible to assess the independence with the use of "impartial criterion" (*e.g.*, correlation factor related to different time-lags), it is a very laborious and time-consuming procedure. Moreover, the choice of this criterion is subjective and if not adequately determined it may lead to negligence of some important extreme rainfall episodes, subconsequently influencing the interpretation of return period and finally leading to mistaken IDF-formulas (*e.g.*, Arnell 1982, Willems 2000). Thus, many authors choose the independency criterion without any special mathematical analysis, by *a priori* establishment of the minimal rainless interval length enabling to consider two "neighboring" episodes as independent (*e.g.*, Willems 2000, Marsalek 1978, Johansen 1979).

The values of such criterion are usually equal to a few hours; however one can find analysis with 1 hour or even 0.5 hour criterion (Arnell 1982).

Homogeneity of the sample is connected to stable (over time) measurement conditions (*e.g.*, localization of gauge stations, measuring equipment, *etc*.), sufficiently long time of observations and the lack of breaks in observations. Schilling (1991) specified the "desirable" (ideal, optimal) characteristics of the set of rainfall data, *i.e*., the length of the observation period equal to 20 years, temporal resolution of 1 min, spatial resolution of 1 km^2, and lack of breaks in observations. It is obvious that for many practical cases such conditions are impossible to fulfill. Moreover, the probability of totally uninterrupted work of the measuring instruments over several decades is very low. Thus, the above-mentioned author also described the "recommended" features of data set, depending on the purpose of the analysis. For example, for preliminary design of storm water conduits and storm water reservoirs, the 10-year observation period is sufficient and breaks in observations are acceptable. Temporal resolution of 10 min is usually considered as minimal for sewer system design (Vaes *et al.* 2001). For rivers the time-step in rainfall observations may be longer. However, the quality of the analysis is strongly influenced by the quality of data.

2.2 Storm episodes selection

For the purpose of IDF-formula determination, the largest episodes are selected from the whole "population" of recorded precipitation data on the basis of a predetermined criterion separating storms from the "typical" (negligible from an engineering point of view) rainfall episodes. Depending on the assumed method of analyses, two basic approaches can be considered: the "annual maxima series" approach and "partial-duration series" approach (Chow 1964, Arnell 1982, Vukmirović and Petrović 1991, Svensson *et al*. 2007, Reiss and Thomas 2007, Ben-Zvi 2009, and others). For the first one (also called "block maxima method") – a single largest extreme for each calendar year of the observation period is selected, while for the second one – the number of rainfall episodes taken into considerations may exceed the number of years in the period of observation, as it is allowed that several extreme episodes may occur in the same year. This last approach is also called "peak-over-threshold" (POT), as the criterion of selecting the storms from the whole "population" of rainfall events is the defined value of threshold in precipitation depth (corresponding with the considered interval duration t_d), over which the analyzed phase of the episode (of the considered duration equal t_d) is classified as a storm and thus taken into further consideration. The "peak-over-threshold*"* technique is more laborious and thus less often used. However, many authors consider this approach more suitable (*e.g.*, Kisiel *et al.* 1971, Todorovic 1978, Madsen *et al.* 1997a, b, Willems 2000,

Licznar and Łomotowski 2005), as it does not limit the number of storms to the number of years in observation period. This way, years which are "wetter" and with several intense storm episodes are not treated on the same level of importance as the years when no substantial storm occurred. Other advantages of the POT approach include higher accuracy of rainfall intensity prediction for lower return periods and possibility of predicting maximal intensities for return periods shorter than one year (Willems 2000). However, the POT technique requires special attention to rainfall events independency and adoption of the specific criterion of determining the value of threshold. As there are no strict and unified procedures concerning the latter of the mentioned aspects, this stage of data preparing is often connected to relatively high degree of subjectivity. The threshold can be set on the basis of a thorough analysis of the data (*e.g.*, Tanaka and Takara 2002, Svensson *et al.* 2007) – *e.g.*, from mean residual life plots (Coles 2001, Svensson *et al.* 2007), or assumed *a priori* on the basis of previous experiences and local specificity of the region. In some of the approaches the threshold value is determined on the basis of the arbitrarily assumed preferred average number of events per year (usually from 2 to 5) (*e.g.*, Tavares and Da Silva 1983, Trefry *et al.* 2005). Other authors suggest that the size of a "partial-duration series" sample should be at least 1.65 times greater than the number of years on record (*e.g.*, Cunnane 1973). More detailed list of different approaches to this question is presented by Ben-Zvi (2009).

In general, both approaches – "annual maxima series" and "partial-duration series" – lead to different IDF-formulas for the same gauge station and the same length of record period, thus it is important to take into account and state explicitly the method of estimation, and use recalculation formulas if necessary (Chow 1964).

2.3 IDF-formulas determination

After extracting the extreme rainfalls from the whole period of observations, the most intense rainfall phases of considered duration, t_d, are selected. Thus, from each selected rainfall episode, the extreme value of intensity (referred to the considered t_d) is derived by moving-average technique. For example, if the duration t_d equal to 30 min is considered, from each episode the worst (as to the rainfall intensity) phase lasting 30 min is selected, and the average intensity during this phase is recognized as the extreme value. This way, the set of extreme intensities is developed for each considered value of duration t_d. The statistical analysis is then carried out to obtain the formulas for the intensity-duration-frequency relationships.

In order to obtain the final result of IDF-formulas, two approaches may be applied – "empirical" (also called "physical") or "theoretical" (also called "probabilistic"). In the first one, the selected and structured storm data are

described by a chosen mathematical formula (set of formulas) on the basis of matching the general form of the formula and calibrating parameters, usually with use of optimization techniques. In the second one – theoretical distribution functions are applied and the parameters are estimated on the basis of statistical approach. The distribution functions most commonly used in IDF-formula development are Gumbel, Log-Pearson type III, and Generalized Extreme Value (GEV); however, other functions, such as Gamma, Log-normal, Pearson III, Fisher-Tippett, Generalized Pareto or exponential distributions have also been applied (*e.g.*, Sevruk and Geiger 1987, Katz *et al.* 2002, Koutsoyiannis 2004a, b, De Michele and Salvadori 2005, Venkata Ramana *et al.* 2008, Ben-Zvi 2009, Endreny and Imbeah 2009, Elsebaie 2012). Polish experiences indicate, either empirical approach or theoretical Pareto, Pearson III, Fisher-Tippett, and Log-normal distributions were usually adopted (*e.g.*, Wołoszyn 1991, Suligowski 2004 and others). More detailed mathematical background of IDF-formulas development was presented by Koutsoyiannis *et al.* (1998). Although the second (theoretical) way of IDF-formulas estimation may be considered as more justifiable from hydrological point of view, the first approach is also successfully applied and in some cases may even demonstrate better efficiency, especially if rainfall data do not "match" the theoretical distributions (*e.g.*, Langousis and Veneziano 2007). In such cases, the optimized empirical IDF-relation enables better fitting to the observations than "typical" hydrological approaches.

A generalized IDF relationship between rainfall maximal intensity I and duration of the rainfall interval t_d corresponding to specified return period T is usually presented in the literature in the form (Koutsoyiannis *et al.* 1998, Venkata Ramana *et al.* 2008, Endreny and Imbeah 2009):

$$I = \frac{A(T)}{B(t_d)} \,, \tag{1}$$

where $A(T)$ and $B(t_d)$ are functions defining separable dependence of I on T and t_d. The function $B(t_d)$ in Eq. 1 is usually expressed as:

$$B\left(t_d\right) = \left(t_d + \theta\right)^{\eta} \,, \tag{2}$$

where θ and η are parameters ($\theta \geq 0$, $0 < \eta < 1$).

Function $A(T)$ in Eq. 1 is usually presented in the literature in the form of two alternative relations of empirical type:

$$A(T) = A_1 T^k \,, \tag{3a}$$

$$A(T) = A_1 \ln T + c \,, \tag{3b}$$

where A_1, k, and c are empirical parameters. However, it can be also derived for typical distributions of maxima on the basis of hydrologic distribution functions (Koutsoyiannis *et al.* 1998), *e.g.*:

for Gumbel distribution: $$A(T) = \lambda\left\{\psi - \ln\left[-\ln\left(1-\frac{1}{T}\right)\right]\right\}, \tag{4a}$$

for GEV distribution: $$A(T) = \lambda\left\{\psi + \frac{\left[-\ln\left(1-\frac{1}{T}\right)\right]^{-\kappa} - 1}{\kappa}\right\}, \tag{4b}$$

for exponential distribution: $$A(T) = \lambda\left(\psi + \ln T\right), \tag{4c}$$

for Pareto distribution: $$A(T) = \lambda\left(\psi + \frac{T^{\kappa} - 1}{\kappa}\right), \tag{4d}$$

and others, where λ, ψ, and κ are the parameters of distribution functions.

Final formulas of IDF-relationships for each considered case are obtained after identification of all the parameters, the number of which is usually not greater than five. In practice, instead of the simultaneous consideration of maximal rainfall intensities for different values of rainfall durations – leading to combined formulas of a general type (Eq. 1) – a very common approach is to separate development of $I(T)$ formulas for selected specified rainfall durations. This makes the process of parameter identification much easier because of the simplicity of the $I(T)$ formula and the significant decrease of the number of parameters to identify. However, the obtained result has less universal character and can be applied in practice only for the specified rainfall durations. Therefore, a more general form of IDF-formula, expressing the dependence of I on both, T and t_d, is preferable.

Indeed, most of widely applied IDF-formulas can be expressed as combination of Eqs. 1 and 2 with one of the alternative relations for $A(T)$ presented above. Commonly known formulas of Lindley, Talbot, Sherman or Reinhold are in fact particular forms of general relationship (Eq. 1). However, some authors (*e.g.*, Arnell 1982) presented the studies in which the most satisfactory results were obtained for the general form of IDF relation different from Eq. 1. Similar conclusions are also provided by Polish experiences, *e.g.*, published by Lambor, Wołoszyn, Licznar, and Łomotowski (all in Polish), focused on development of local IDF-formulas for selected localizations in the country. Taking into account most of these works, more general form of IDF-relationship may be expressed as:

$$I = \frac{A(T)}{B(t_d)} + C(T) \ , \tag{5}$$

(with the range of $A(T)$ formulas wider than Eqs. 3-4). However, in some particular cases presented by the above-mentioned authors, the best results were obtained for even more complex forms of IDF-formulas, which are not a simple combination of functions defining separable dependence of I on T and t_d as in Eqs. 1 or 5. In such cases, the problem of IDF-formula estimation is more complicated.

The parameters appearing in a general form of IDF-relationship must be estimated in each particular case in order to obtain the final IDF-formula for the specified localization. Estimation of the parameters is often considered to be an easy and well recognized subject. It can be developed in different ways, ether with use of statistical estimators for known probability distribution functions or with use of any optimization technique, suitable for the case analyzed. The choice of the method of parameter estimation is certainly very much connected to the assumed form of IDF-relationship and the number of parameters. In most practical applications, commonly used methods of parameter estimation are: *L*-moments (LM), Maximum Likelihood (ML), Probability Weighted Moments (PWM) (*e.g.*, Koutsoyiannis *et al.* 1998, Endreny and Imbeah 2009, Ben-Zvi 2009, Hailegeorgis *et al.* 2013, and many others) or optimization techniques, *e.g.*, Powell method (Koutsoyiannis *et al.* 1998). Although in many cases (especially when the simple form of IDF-formula is chosen with a small number of parameters) the parameter identification is actually a very easy procedure from a "technical" point of view. However, there are also experiences showing that this "standard" approach can lead to significant errors in final solution (*e.g.*, Katz *et al.* 2002). In some cases the recommended optimization techniques may not lead to satisfactory results due to the relatively high number of parameters being estimated with use of local optimization techniques. For example, Koutsoyiannis *et al.* (1998) present the method of simultaneous estimation of the parameters of both – the distribution function $A(T)$ and duration function $B(t_d)$ – with use of Powell method. Taking into account that in such optimization problem, depending on the chosen form of IDF-relationship, even five or more parameters are searched, it seems that the Powell method (and other local minimum search methods) may lead to "not optimal" solution, as it is difficult to verify whether the objective function is unimodal. Some authors recommend performing the optimization using the solver tools in common spreadsheet packages (usually without pointing the optimization method used in such approach) or present other methods of solving the problem, *e.g.*, robust estimation proposed by Koutsoyiannis *et al.* (1998). However, appli-

cation of global optimization methods to the problem of IDF-formula development is not often considered in bibliography.

This paper proposes the effective and relatively simple empirical approach to IDF-formula development based on global optimization technique of Controlled Random Search (CRS). The approach may be an interesting alternative to other methods, as it leads to satisfactory solution obtained in relatively quick and simple way, especially in those cases in which IDF-formula has more complex form and the optimal values of higher number of parameters are searched simultaneously. In such cases the properties of the objective function (mainly its unimodal/multimodal character and the approximate position of the global optimum) are usually difficult to prior recognition, and thus the choice of global search method enables to achieve a satisfactory solution. CRS method proved to be effective in other applications in engineering (*e.g.*, Ali *et al.* 1997a, Dysarz and Napiórkowski 2002, Manzanares-Filho *et al.* 2005, Manzanares-Filho and Albuquerque 2008). It can be also successfully applied to hydrological problem of IDF-relationship determination.

3. CONTROLLED RANDOM SEARCH METHOD AS A GLOBAL OPTIMIZATION TECHNIQUE

The procedure of the controlled random search (CRS) for global optimization was first presented by Price (1977, 1978, 1983). The author proposed a preliminary version of the method (further referred as CRS1; Price 1977, 1978), which was then modified to CRS2 (Price 1983). The algorithms proposed by Price were later used, modified and improved by many researchers (*e.g.*, Ali *et al.* 1997a, b, Manzanares-Filho *et al.* 2005, Manzanares-Filho and Albuquerque 2008, Tsoulos and Lagaris 2006). In this study the scheme CRS2 in version adopted by Dysarz and Napiórkowski (2002) was applied to IDF-formula development.

The CRS technique combines the "random" algorithms of typical global optimization methods with "deterministic"-"controlled" – procedures of local optimization, increasing the efficiency of the process. The method represents the group of "direct" procedures, where no gradient analysis is involved. It is applicable for both unconstrained and constrained optimization.

Let n be a number of variables (parameters) that are searched, and thus n defines the size of the space of possible solutions (the search domain) V. The search domain V is constituted by specifying the limits to each variable. In optimization problem, the "optimal" set of parameter values is searched, meaning such for which the assumed criterion (objective function) takes the extreme (usually minimal) value.

In the first step of the CRS1 procedure, predetermined number N of trial points are randomly chosen over n-dimensional space V, with respect to additional constraints, providing there are any. In the original version of the procedure (CRS1), Price (1977, 1983) suggests (based on empirical experience) applying the relation $N = 25\,n$. Each of the N points represents one possible combination of the values of parameters that are searched. Let us denote this set of points as S. For each of the chosen trial points, the objective function F is evaluated, and the points are then arranged in the array **A**, in which the position is dependent on the value of F (from 1 – "the best" to N – "the worst"). Let M denote the worst and L – the best point in **A**. In each iteration, the set of $n+1$ points R_i ($i = 1, 2, \ldots, n+1$) is randomly chosen from S, constituting the so-called "simplex" in n-dimensional space. The point R_{n+1} is arbitrarily considered as the pole of the simplex and the next trial point P is defined as the image point of the pole R_{n+1} with respect to centroid G of the remaining n points, according to the formula:

$$\boldsymbol{P} = 2\boldsymbol{G} - \boldsymbol{R}_{n+1} \ , \tag{6}$$

where $\boldsymbol{P}$, $\boldsymbol{G}$, and $\boldsymbol{R}_{n+1}$ represent the position vectors of corresponding points in n-dimensional space. Point P is called preliminary trial point. For the obtained point P the value of objective function F_P is calculated and compared with F_M – the value for the worst point M. If $F_P < F_M$, and if P satisfies the additional constraints (if there are any), the point M is replaced by P, the array **A** is rearranged and a new simplex is chosen. If P fails to improve the value of M, the secondary trial point $\boldsymbol{Q}$ is chosen, according to the formula:

$$\boldsymbol{Q} = \left(\boldsymbol{G} + \boldsymbol{R}_{n+1}\right)/2 \tag{7}$$

and the value of F_Q is calculated. If Q occurs better than M – the points are replaced; otherwise – the trial is discarded. The new simplex is then randomly chosen and so on. The "best" parameter values at each step of the CRS procedure are defined by the coordinates of point L, and are improved in successive iterations. The next iteration is performed as long as the stopping criterion is satisfied.

The modified procedure CRS2 is similar to CRS1 in general approach, however it differs in details. In CRS2 only primary trial points (Eq. 6) are chosen, no secondary trial points are generated. The second important difference is that the point R_1 of a simplex is not randomly searched, but it is always defined by the "best" point L from the array **A** of N points. Thus the simplex is constituted by choosing n points from $N - 1$, because L is arbitrarily chosen. In this way the "local part" of the searching procedure is much more effective. In order to minimize the danger of premature convergence to local optimum (what may take place in the case of multimodal problems),

the sufficiently high number of trial points N should be randomly chosen and the stopping criterion should be carefully considered. For this version of the method, Price (1983) recommends determination of the optimal number of N according to the formula:

$$N = 10(n+1) . \tag{8}$$

Thus, if 4 parameters are searched, the initial set of 50 points is randomly chosen and each simplex is constituted of 5 points in 4-dimensional space. For 8 parameters, the size of the set of initial points increases to 90 and each simplex is built of 9 points.

CRS2 was successfully adapted to different applications, including complex multimodal problems (Ali *et al*. 1997a). It also has been modified many times by different authors (*e.g.*, Rinaudo *et al.* 1998, Manzanares-Filho *et al.* 2005, Tsoulos and Lagaris 2006, Manzanares-Filho and Albuquerque 2008). The modifications, leading to subsequent versions CRS3, CRS4, CRS5, and CRS6 (Ali *et al.* 1997a, b), usually concern the initial procedure of random choice of N points, the choice of $n+1$ points of the simplex, combining CRS with different local search algorithms and modifications of the procedure of controlled random search. Price (1977, 1983), Ali *et al.* (1997b), Tsoulos and Lagaris (2006), and other authors present several "mathematical" examples of test function minimization. One can also find several examples of CRS successful application to the practical problems, *e.g.*, inverse airfoil design (Manzanares-Filho *et al.* 2005, Manzanares-Filho and Albuquerque 2008), macromolecular modeling (finding least energy structure for a given molecular system, Ali *et al.* 1997a), chemical engineering (tank reactor optimization, Ali *et al.* 1997a), applied statistics (pig liver problem; Ali *et al.* 1997a), and optimal flood control in multi-reservoir river system (Dysarz and Napiórkowski 2003). The CRS method and its modifications were compared to other approaches of global optimization by several authors. The interesting study of such a comparison was presented by Ali (1994) and Ali *et al.* (1997a). The authors compared different versions of CRS algorithms with the methods of multilevel single linkage (MSL) (Rinnooy Kan and Timmer 1987), topographical multilevel single linkage (TMSL) (Ali and Storey 1994), simulated annealing (SA) (Dekkers and Aarts 1991) and aspiration-based simulated annealing (ABSA) (Ali and Storey 1997), all applied to selected practical problems. The efficiency of the algorithms was compared taking into account the number of function evaluations, the CPU time required to solve the problem and the value of obtained optimum in each analyzed case. Numerical tests showed that the algorithms of CRS-type were preferable to the others for problems with many optima, especially if the number of local optima was large, even if the optima were close to each other.

The CRS method can be also successfully applied to IDF-curves determination. The problem of IDF-relationship development, in comparison with the ones from the fields of material science or chemical engineering, is relatively easy. However, if the general form of IDF-relationship is more complex than the functions typically applied in such cases (*e.g.*, Eqs. 1-4), and if the number of parameters to be identified is higher (four, five or more) and finally – if the objective function is multimodal, the use of effective but relatively simple global optimization procedures is reasonable. As an example, the application of CRS2 algorithm to IDF-formula determination for the rainfall data collected in one of the gauge stations in Gdańsk (Poland) – in the Gdańsk University of Technology (GUT) station – is presented. The results are compared with the ones obtained with the use of successive search, Powell method (Powell 1964, Pierre 1969, Press *et al.* 2007), modified version of Powell method, Newton search (*e.g.*, Bonnans *et al.* 2006) and evolutionary algorithm (Goldberg 1989). In the two latter approaches, the embedded solver tools of commonly used spreadsheet packages (Excel) were applied. In the first case, Newton's search algorithm implemented in classical optimization tool proposed by Fylstra *et al.* (1998) was used. In the second one, The Evolutionary Solver (in Excel 2010), based on the combination of genetic algorithm and other search methods, was applied.

4. GDAŃSK STUDY CASE CHARACTERISTICS

The city of Gdańsk is located in the northern Poland, at the bank of Baltic Sea, in the region of Vistula river estuary (Fig. 1). Average yearly precipitation depth for Gdańsk (for the period 1951-2008) is 550 mm, thus it is lower than the corresponding value for the whole territory of Poland (*c.a.* 600 mm). The value of total annual precipitation depth for Gdańsk over the period of nearly 60 years of observations is highly variable; however, as a general trend, the well pronounced increase can be observed (Wołoszyn 2009). Precipitation in cold period (November-April) is on average 200 mm, thus it is nearly twice smaller than in warm season (May-October). Despite the abovementioned proportion in precipitation depths, the number of days with measurable precipitation in winter is higher than in summer. Precipitation in cold period (especially in December, January, and February) has a more permanent character and usually takes the form of snow (in average 40-50 days per year). However, daily sums of precipitation in winter are low and snow cover lasts for a relatively short time. Thus, for the purposes of urban hydrology, precipitation in the form of rainfall, (occurring mainly in warm period, often as short but heavy storms) is much more important from an engineering point of view.

Fig. 1. Distribution of averaged total yearly precipitation in Poland.

Nowadays, on the territory of Gdańsk, there are three meteorological and climatological observation posts in charge of the national research institute (Institute of Meteorology and Water Management) and many private stations. Important part of the latter group was created by Gdańskie Melioracje (Gdańsk Meliorations Inc.), thanks to which the whole network of rainfall gauges, designed particularly for urban hydrology needs, was installed. Although the lengths of the observed precipitation records in these stations are still too short to support sufficient data for statistical analysis nowadays, the observations will be the important source of rainfall data in the closest future, both for temporal and spatial variability analysis.

The gauge station at the Gdańsk University of Technology (GUT), remaining in charge of a Faculty of Civil and Environmental Engineering, is located in the central district of the city. The GUT station was established in 1991, when Hellman's pluviometer and two pluviographs were installed. The observation post was additionally equipped with the automatic system in 2008, and since then both analog and automatic measurements of rainfall are carried out.

The observations in Gdańsk indicate that the mean monthly value of precipitation is the highest for July (70.0 mm) and August (63.3 mm) and the lowest for February (25.2 mm) and March (26.2 mm). The same tendency can be observed in the context of the mean daily maximum values – the

highest for July is about 24.1 mm, while the lowest for February is 6.5 mm (Wołoszyn 2009). However, the historical observations have shown that the region is not free from extreme phenomena. The most spectacular one, causing catastrophic flood in Gdańsk, took place in 2001, when during 24 hours (9 July 8.00 a.m. – 10 July 8.00 a.m.) the sum of 123.5 mm in GUT station was observed (*c.a.* 20% of average yearly precipitation in Poland). During the most intense phase of the rain, the amount of 65 mm (during 1 hour 40 min) was measured (Wołoszyn 2003).

The above-presented characteristics of precipitation in Gdańsk are additionally influenced by the factors resulting from local aspects. Observations indicate both: spatial and temporal variability of precipitation, caused by varied rate of urbanization and city development in different parts of the city. This confirms the strong need not only to determine the formulas characterizing "averaged" precipitation conditions in Gdańsk, but to develop the separate local IDF-formulas for several districts of the city as well.

5. PRECIPITATION DATA COLLECTION AND ANALYSIS

Rainfall data were collected in GUT climatological station during the period 1991-2010. Till 2008 daily and weekly pluviograph charts recorded by two Hellman's siphon pluviographs (float-type) were available. Since 2008 additionally the recordings of the automatic tipping-bucket rain gauge (with the time resolution of 10 min and the measurement accuracy of 0.2 mm) have been at disposal.

Data selection was carried out in several steps:

- identification and rough selection of rainfall episodes,
- analysis of chronological course of each selected episode,
- analysis of extreme rainfall phases during each episode,
- selection (for each assumed value of rainfall duration t_d) of the episodes relevant for IDF analysis, according to the pre-selected criterion.

For each day of the considered period of 20 years of observations, precipitation recorded at a resolution of 10-minutes was analyzed. In order to establish the criterion for storm selection, the Chomicz rainfall intensity scale, popular in Poland, was adopted. According to this scale, each rainfall episode (or each selected phase of the episode) can be classified in the context of its intensity to one of thirteen categories on the basis of the relation between rain/phase duration and the corresponding precipitation depth. Each category in this scale is given a grade number (from 0 to 12). The reference values of precipitation depth, distinguishing different types of rainfall, are defined as:

$$U_k = \alpha_k \sqrt{t} \ , \tag{9a}$$

where

$$\alpha_k = \sqrt{2^k} \ , \tag{9b}$$

k is a grade of the Chomicz scale ($k = 0, 1, \ldots, 12$), U_k is the upper limit value of precipitation depth [mm] for the rainfall of a kth grade and t [min] is duration time of a rainfall (or its selected phase). First nine lines describing the relation 9 for grades $k = 1 \ldots 9$ are presented in Fig. 2. According to the Chomicz scale, rainfalls are classified as normal (grade 0), heavy (grade 1), storms (grades 2 and 3), heavy storms (grades 4 and 5), and torrential storms (grades 6-12). Knowing the duration of the analyzed episode/phase and total precipitation depth accumulated in this period, one can find the corresponding point in the plot and therefore determine the category of the rain intensity.

The approach applied in the Chomicz scale differs from the most popular method of rainfall classification, in which the category of rainfall intensity is determined on the basis of precipitation depth accumulated in predefined period (usually 1 hour or 24 hours) or average rainfall intensity. Differences between such approaches in different countries usually concern demarcation values of precipitation depth or rainfall intensities. For example, according to National Meteorological Institute of Spain, rainfall can be classified as light for intensities I not greater than 2 mm/hour, moderate – for $2 < I \leq 15$ mm/hour, heavy – for $15 < I \leq 30$ mm/hour, very heavy – for $30 < I \leq 60$ mm/hour and torrential for $I > 60$ mm/hour (Llasat 2001). In a similar Polish approach proposed by the Institute of Meteorology and Water Man-

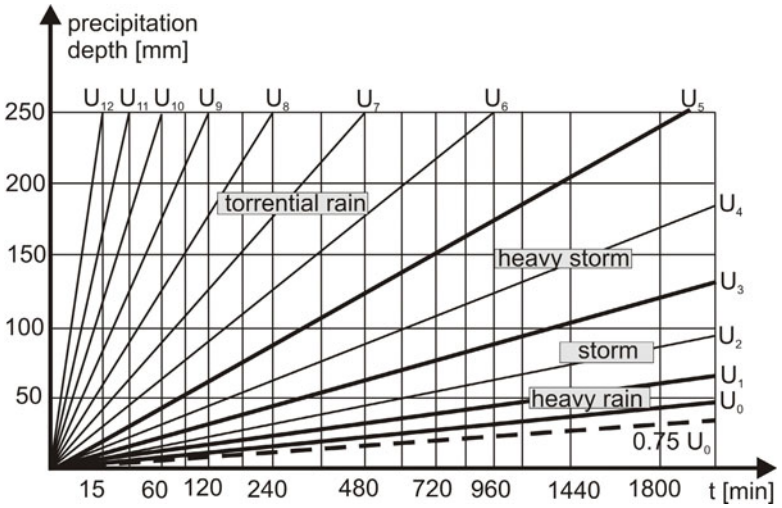


Fig. 2. Illustration of Chomicz scale for precipitation intensity.

agement, heavy rains refer to intensities $0.2 < I \leq 1.0$ mm/min ($12 < I \leq 60$ mm/hour), storms – $1.1 < I \leq 1.9$ mm/min and torrential rainfalls – $2.0 < I$ mm/min (with several grades in each of the mentioned categories). In all these approaches based on the predefined values of intensities, the ratio of precipitation depth to its duration, describing the limits of each rainfall category, is constant, while in the Chomicz scale not. The nonlinear character of formula 9 stresses not only the importance of intensity itself, but also duration of the period in which such intensity was observed.

Besides these simple approaches to rainfall classification, one can find more complex analyses. For example, Llasat (2001) presented the classification of rainfall events on the basis of their convective features described by the parameter $\beta^*_{L,\Delta T}$. This parameter, calculated on the basis of rainfall hyetograph (with the interval ΔT), expresses the ratio of the precipitation that exceeds the predefined intensity threshold L to the total rainfall in the episode. Based on the value of $\beta^*_{L,\Delta T}$ the episode is classified into one of four categories. The author discussed the relations between $\beta^*_{L,\Delta T}$, intensity and rainfall duration for the analyzed region. Maraun *et al.* (2008) presented the method for the classification of rainfall intensity based on 10 categories, each referred to 10% of the total rainfall amount (for the analyzed period of observations), from the 1st category containing the weakest precipitation events (in the number sufficient to make up the lowest 10% of the total rainfall amount) to the 10th category – the highest one. In this approach the values of thresholds depend on the number and intensity of the measured rainfall episodes, and not on the objective criterion, independent of the results of observations.

Different scales of rainfall intensities lead to different classifications of the same set of episodes therefore the choice of the objective criterion is difficult. In this study, Polish experiences in IDF-formula development were taken into account. The Chomicz scale, revised by recent Polish studies (Licznar and Łomotowski 2005), was finally chosen as the basis of the threshold selection. The experiences of Licznar and Łomotowski (2005) showed that "heavy rain" criterion referred to $k > 0$ of the Chomicz scale is too stringent, eliminating a large number of episodes of a lower precipitation depth which in practice should be taken into account. Consequently, the mentioned authors established the criterion:

$$P \geq 0.75 \cdot U_0 \ , \tag{10}$$

where P is the observed precipitation depth [mm], and U_0 is the reference value of Chomicz scale for $k = 0$. Thanks to such approach, each year in the analysis is represented by at least a few rainfall episodes. Following the sug-

gestions of the above-mentioned authors, the criterion 10 was also implemented in this study.

During the preliminary (rough) selection of the episodes (step 1), only the smallest, evidently below criterion episodes were neglected and 122 episodes observed during the period of 20 years were taken into account (Table 1). Average number of episodes per year was 6. The extreme frequency of storm episodes in a year was 13 and it was observed in 2008. The longest continuous rainfall episode was 1092 min long and was observed on 9/10 July 2001, when the already mentioned catastrophic flood in Gdańsk took place. However, the case of such a long rainfall was incidental, as most measured rainfall episodes were not longer than 100 min. Many of them were very short storms, lasting one hour or less, separated from other episodes by rainless period of at least several hours. In most cases, a single episode was observed per day; days with two or three rainfalls were incidental and the duration of rainless interval between the storms in such cases was usually a few hours.

Table 1

Number of extreme episodes from period 1991-2010 taken into considerations

Year	Number of extreme rainfall episodes								
	March	April	May	June	July	August	September	October	Total
1991	–	–	–	5	–	–	–	–	5
1992	–	–	–	1	3	–	–	–	4
1993	–	–	–	–	2	2	2	–	6
1994	–	–	1	–	–	1	2	–	4
1995	–	–	–	2	2	–	–	–	4
1996	–	–	1	1	3	–	–	–	5
1997	–	–	3	–	1	–	–	–	4
1998	–	–	2	–	1	–	–	1	4
1999	–	1	2	1	–	–	–	–	4
2000	–	–	1	2	–	1	2	–	6
2001	–	1	1	–	7	3	–	–	12
2002	–	–	3	1	–	2	–	–	6
2003	–	–	1	1	3	–	2	–	7
2004	–	–	–	1	2	–	1	–	4
2005	–	–	3	–	1	–	–	–	4
2006	–	–	–	3	–	–	1	–	4
2007	–	–	–	1	1	–	5	2	9
2008	1	–	2	3	4	2	–	1	13
2009	–	–	2	3	2	–	1	–	8
2010	–	–	4	1	1	2	1	–	9
Total	1	2	26	26	33	13	17	4	122

The selected episodes were analyzed in their chronological course in terms of the precipitation depth P [mm], intensity I [mm/min] and q [dm^3/(s·ha)] and cumulative precipitation depth ΣP [mm] in each 10-minute step of the episode duration. The example of such analysis (for the extreme episode recorded on 11 June 2010) is presented in Table 2. The analysis let observe that for very short rainfalls (20-30 min) maximal intensity usually fell on first impulse of rain. For short rainfalls (40-90 min), the maximal intensity is usually observed in the first half of the episode. Longer rainfalls usually have several peaks of intensity, at least one in the first and in the second half of the episode.

Table 2

Rainfall characteristics – the episode recorded on 11 June 2010

Clock time hours (p.m.)	Duration [min]	Precipitation depth [mm]	Cumulative precipitation depth [mm]	Intensity i [mm/min]	Intensity q [dm^3/(s·ha)]
4.20 – 4.30	10	0.3	0.3	0.03	5.00
4.30 – 4.40	10	8.6	8.9	0.86	143.36
4.40 – 4.50	10	1.8	10.7	0.18	30.01
4.50 – 5.00	10	0.5	11.2	0.05	8.34
5.00 – 5.10	10	0.3	11.5	0.03	5.00
5.10 – 5.20	10	0.5	12.0	0.05	8.34

Next stage of the study concluded of extreme rainfall phase analysis. In the study, the aggregation levels t_d equal to 10, 20, 30, 60, 90, 120, 150, 180, 210, and 240 min were considered. For each rainfall episode, the most intensive phases of pre-selected durations t_d (not exceeding the total duration of the episode) were found and analyzed as to their intensities. The assumed value of 240 min of the upper limit of t_d results from the specificity of the observed rainfall episodes, but it also has its practical justification. As it is known, the design rainfall duration, specified for the purposes of urban hydrology, is strongly correlated with the time of runoff concentration in the considered catchment. Most practical cases of urban drainage and rainwater system design concern relatively small basins with significant share of impervious areas, where time of concentration is short and runoff processes are very dynamic. Moreover, it has been proved that the vast majority of urban floods is caused by extremely intensive torrential rainfalls of short duration. Thus, the value of 240 min as the upper limit of rainfall duration in the presented considerations enables proper recognition of the rainfalls most important from the practical point of view and corresponding with local conditions in Gdańsk. The example of the analysis of the maximal intensity phases (episode of 11 June 2010) is presented in Table 3 and Fig. 3.

Table 3

Rainfall maximal intensities – the episode recorded on 11 June 2010

Clock time hours (p.m.)	Duration time t_d [min]	Maximal precipitation depth [mm]	Maximal rainfall intensity q [dm^3/(s·ha)]
4.30 – 4.40	10	8.6	143.36
4.30 – 4.50	20	10.4	86.68
4.30 – 5.00	30	10.9	60.57
4.20 – 5.20	60	12.0	33.34

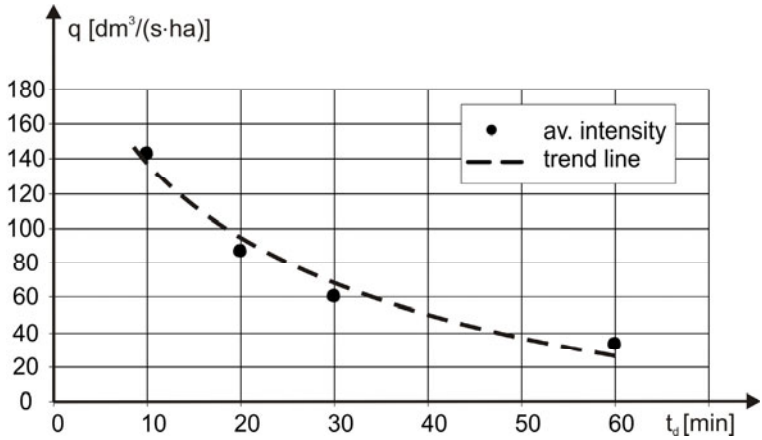


Fig. 3. Maximal intensity *versus* duration time –the episode recorded on 11 June 2010.

The intensities averaged for the analyzed rainfall phase durations were then segregated in tables with each duration t_d separately, and data in each table were arranged in the non-increasing sequence of string-values. For each table the condition 10 was rechecked and the data that did not match the criterion were removed. Such set of data was the starting point for IDF analysis.

6. PRELIMINARY IDF-FORMULA DEVELOPMENT

6.1 The $q(p)$ relationship

The observed rainfall extremes for each duration time t_d, arranged in the statistical order of non-increasing series q_i, $i = 1, 2, \ldots, N$, where $q_1 \geq q_2 \geq \ldots \geq q_N$, were assigned to the empirical quantiles corresponding to empirical probability of exceedence p_i according to the formula:

$$p_i = \frac{i}{N+s} \cdot 100 \ [\%] \ , \tag{11}$$

where N was the total number of extremes in the series and $0 \le s \le 1$. In this study $s = 1$, which corresponds with the so-called Weibull plotting position of a quantile plot. The formula obtained in this way is considered to be the most practical one, which fully satisfies Gumbel's conditions for plotting position (Yevjevich 1972). The example for t_d equal to 120 min is presented in Table 4. The comparison of the obtained relations for selected values of t_d is shown in Fig. 4.

Table 4

Sequence of distribution of extreme rainfall intensity for duration time t_d equal to 120 min

No. in series i	Episode number	Total precipitation depth [mm]	Average maximal rainfall intensity q [dm^3/s·ha]	Probability of exceedence p [%]
1	124	57.000	79.170	4.00
2	5	30.800	42.778	8.00
3	22	28.350	39.375	12.00
4	66	22.400	31.111	16.00
5	123	18.100	25.139	20.00
6	6	15.200	21.111	24.00
7	37	14.100	19.583	28.00
8	39	13.700	19.028	32.00
9	91	13.600	18.889	36.00
10	34	12.900	17.917	40.00
11	58	12.500	17.361	44.00
12	76	12.500	17.361	48.00
13	73	12.200	16.944	52.00
14	15	11.700	16.250	56.00
15	86	11.600	16.111	60.00
16	61	11.000	15.278	64.00
17	25	10.400	14.444	68.00
18	67	10.400	14.444	72.00
19	115	10.000	13.889	76.00
20	117	9.800	13.611	80.00
21	100	9.700	13.472	84.00
22	45	9.600	13.333	88.00
23	94	9.250	12.847	92.00
24	90	9.050	12.569	96.00

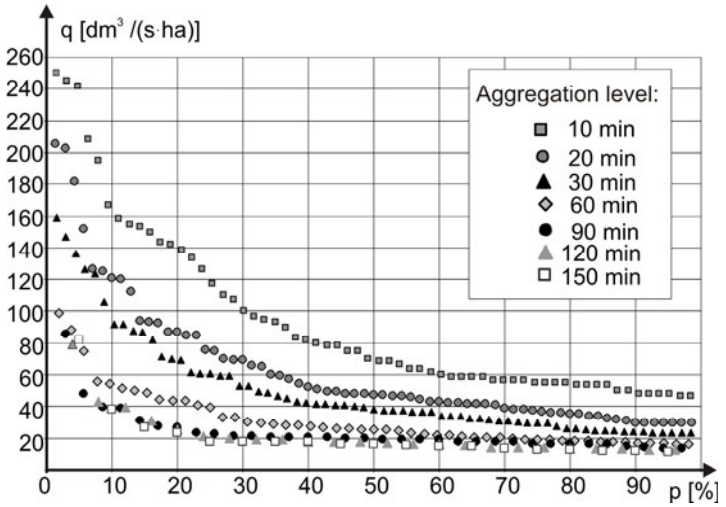


Fig. 4. Comparison of relationship $q(p)$ for different values of duration time.

Table 5

Results of preliminary approximation of $q(p)$ relationships

t_d [min]	$q = a \cdot \ln p + b$			$q = a \cdot p^{-b}$		
	a	b	R^2	a	b	R^2
10	57.931	303.43	0.9768	542.67	0.5194	0.9489
15	47.767	245.07	0.9706	435.57	0.5351	0.9675
20	43.401	220.92	0.9709	398.81	0.5464	0.9698
30	34.933	177.56	0.9778	333.35	0.5597	0.9609
60	19.870	103.66	0.9650	176.74	0.5076	0.9798
90	14.025	74.484	0.8100	105.34	0.4388	0.9605
120	15.945	80.253	0.8252	134.54	0.5303	0.9664
150	17.218	84.021	0.7366	144.88	0.5693	0.9293
180	20.092	95.855	0.7686	214.30	0.6675	0.9466
210	20.044	96.263	0.8159	247.06	0.6963	0.9577
240	22.490	105.150	0.8618	413.55	0.8480	0.9620

The relationships $q(p)$ obtained for each considered aggregation time t_d were analyzed in terms of the best fitting to rainfall data. Different forms of polynomial, logarithmic, and exponential functions were considered and least square method of approximation was applied. Preliminary analysis for each duration t_d showed that there was no type of function that would fit the rainfall data over the whole range of considered values of aggregation time.

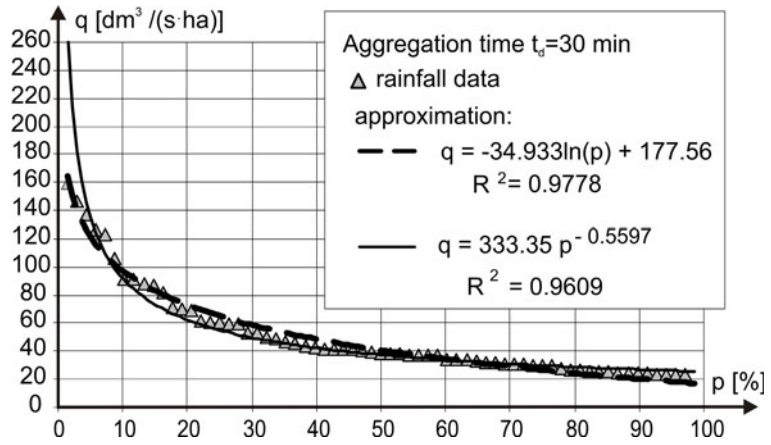


Fig. 5. Approximated $q(p)$ relationship for aggregation time equal to 30 min.

The best fit was obtained for logarithmic and exponential functions (equivalent to Eqs. 3a and b for q and p), however the first one proved to be the best for $t_d < 60$ min, while the second one – for $t_d \geq 60$ min (Table 5). In this last case, logarithmic function did not represent the shape of the relationship $q(p)$ in satisfactory way, while exponential function did not reproduce the $q(p)$ relation properly for the least values of p ($p < 2\%$), causing significant overestimation. The best (in the means of the least square error criterion) approximated functions were characterized by correlation coefficient $R^2 < 0.929,\ 0.979 >$. The example of approximated relationship $q(p)$ for t_d equal to 30 min is presented in Fig. 5.

6.2 The $q(t_d)$ relationship

Based on approximated relationships $q(p)$ for different values of duration time t_d, the $q(t_d)$ relations for selected values of probability p were analyzed. The obtained relationships were approximated with use of least square method. Exponential functions proved to provide the best matching to data points; however, the correlation between the data and approximated function was much poorer (with R^2 even less than 0.7 for some values of p). An example of approximation for $p = 20\%$ is presented in Fig. 6.

The results of the first stage of the analyses showed that the rainfall data required more complicated form of a function representing $q(t_d, p)$ relationship than the easiest forms based on Eq. 1, as the latter did not enable satisfactorily good fitting with observations in this case. Therefore, in order to consider more complex functions describing IDF-relations, more elaborate calibration methods will be desired.

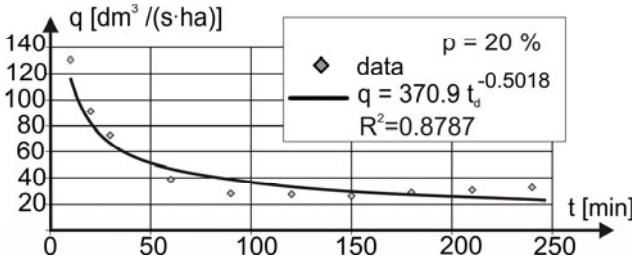


Fig. 6. Approximation of $q(t_d)$ relationship for $p = 20\%$.

6.3 Preliminary calibration of $q(t_d, p)$ relationship with the use of optimization

As mentioned before, the preliminary stage of the analysis showed the need for more complex functions describing IDF-relationship than basic relations expressed by Eqs. 1-3. In order to find the most suitable function describing the relation of q and both – aggregation time and exceedence probability, several forms of functions were taken into considerations:

a) $$P_{\max} = \frac{a + b[-\ln(-\ln(1-p))]}{t_d^n} \,, \tag{12}$$

b) $$P_{\max} = \left[a\left(\frac{1}{p}\right)^n + b \right] t_d^{-n} \,, \tag{13}$$

c) $$P_{\max} = \left[a\left(\frac{1}{p}\right)^m + b \right] t_d^{-n} \,, \tag{14}$$

d) $$P_{\max} = a\left(\frac{1}{p}\right)^m t_d^{-n} + b \,, \tag{15}$$

e) $$P_{\max} = a\left(\frac{1}{p}\right)^m t_d^{-n} + b\, t_d \,, \tag{16}$$

f) $$q = -a \ln p + b \,, \tag{17a}$$

where

$$a = a_1 t_d^2 + a_2 t_d + a_3 \,, \tag{17b}$$

$$b = b_1 t_d^2 + b_2 t_d + b_3 \,, \tag{17c}$$

and

g) $$q = a\, p^{-b} \,, \tag{18}$$

where a and b are defined as in Eqs. 17b, c.

Functions a-e were selected on the basis of previous Polish experiences (mentioned in Section 2.3) and the obtained IDF-formulas for other localizations in the country. Formulas f and g were established on the basis of preliminary analysis. Parameters a and b of logarithmic and exponential relations $q(p)$ (Eqs. 17a and 18), estimated with use of least square method for each selected aggregation level t_d (Table 5), were then analyzed in terms of their dependence on t_d. As a result, it was concluded that the best fit was obtained when each parameter in Eqs. 17a and 18 (a and b) was a quadratic function of t_d. Therefore, formulas 17b, c were proposed. Finally, seven forms of IDF-relationship were taken into account.

In order to find the optimal values of parameters in formulas 12-18, the optimization was carried out. In the first stage, the successive search technique was used. Although time-consuming, it enabled recognition of the general features of the objective function selected as the optimization criterion in terms of its general shape and unimodal or polymodal character. This analysis was then supported by the calculations using Newton algorithm for local optimization. As the optimization criterion, the least square error objective function was assumed, which took the form:

$$F = \sum_{j=1}^{N_t} \left(P_{\mathrm{ob}} - P_{\mathrm{c}} \right)^2 \tag{19}$$

or

$$F = \sum_{j=1}^{N_t} \left(q_{\mathrm{ob}} - q_{\mathrm{c}} \right)^2 \ , \tag{20}$$

where N_t is the total size of a sample, index "ob" denotes "observations" and "c" – "calculations" of precipitation depth P or rainfall intensity q, respectively. The optimal set of parameters (for each formula a-g separately) was the one that minimized the value of F. To compare the obtained results and indicate the goodness-of-fit, several criteria were applied: the total square error (defined by the value of objective function F for the optimized set of parameters), the mean error ε defined as:

$$\varepsilon = \frac{\sqrt{F}}{N_t} \tag{21}$$

and an "optimization cost" C equal to total number of objective function evaluations during the process of optimization. Additionally, the criteria of average relative errors:

$$E_1 = \frac{1}{N_t}\sum_{j=1}^{N_t}\left|\frac{P_{\mathrm{ob}} - P_{\mathrm{c}}}{P_{\mathrm{ob}}}\right| , \tag{22}$$

$$E_2 = \frac{1}{N_t}\sum_{j=1}^{N_t}\left|\frac{P_{\mathrm{ob}} - P_{\mathrm{c}}}{P_{\mathrm{c}}}\right| \tag{23}$$

were also analyzed. An interesting study of practical aspects concerning optimization, including objective functions and identifiability issues, was presented by Romanowicz *et al.* (2013).

After optimizing the IDF-formula parameters, the quality of the optimization (in the sense of the above-mentioned criteria) was checked (Table 6, "all data" tests). The study proved that even when the more complex form of IDF-relationship was applied, it was impossible to find one global formula matching to rainfall data in the whole range of considered values of aggregation time and probability. In all cases (a-g) the matching was not satisfactory, much worse than when $q(p)$ or $q(t_d)$ relations were searched separately. The analysis showed that the general form of relationship $q(t_d, p)$ describing short and very intensive rainfalls should be different from the corresponding formula for the rainfalls of longer duration. Better results than presented in Table 6 ("all data" tests) were obtained when two different formulas, Eq. 17a for $t_d < 60$ min and Eq. 18 for $t_d \geq 60$ min, were taken into account, even when the parameters a_i and b_i ($i = 1, 2, 3$) were estimated not *via* optimization but with use of simple quadratic approximation of $a(t_d)$ and $b(t_d)$ relations on the basis of the data presented in Table 5 only. For the values of a_i and b_i ($i = 1, 2, 3$) obtained in the latter case, the corresponding values of F defined by Eqs. 19 and 20, were 4609 and 24 661, respectively, and thus they were significantly lower than the values in Table 6 ("all data" tests). Therefore, a more complex mathematical description of IDF-relationship than one "global" equation should be applied in this case.

In the last step of the preliminary analysis, all the formulas 12-18 were applied to each of the ranges $t_d \leq 60$ min and $t_d \geq 60$ min separately. This time rainfall data for $t_d = 60$ min were taken into account twice (for both of the ranges considered), in order to obtain better coincidence in the surrounding of 60 min. The optimization was performed in the same manner as in the case of the whole range of t_d. The only difference was the lower number of rainfall data taken into account. The test confirmed that the best results were obtained when Eq. 17a for $t_d \leq 60$ min and Eq. 18 for $t_d \geq 60$ min were applied (Table 6). Thus, finally the Eqs. 17a and 18 were chosen as the best for the case considered.

Once two different formulas to obtain better matching for both, short and longer rainfall durations, are chosen, the main, dispositive stage of optimiza-

Table 6

Effectiveness of different IDF-formulas application

IDF formula (Eq. no.)	Optimal values of parameters	F Eq. 19 [mm^2]	F Eq. 20 [dm^3/(s·ha)]^2	E_1 Eq. 22 [–]	E_2 Eq. 23 [–]	ε^* Eq. 21 [mm]	ε^{**} Eq. 21 [dm^3/(s·ha)]
all data ($t_d \leq 240$ min)							
12	$a = 1.052$, $b = 0.662$, $n = -0.550$	23113	81 693	0.278	0.321	0.321	0.604
13	$a = -0.379$, $b = 4.130$, $n = -0.499$	20859	107 006	0.471	0.337	0.305	0.691
14	$a = 14.407$, $b = -9.476$, $n = -0.555$, $m = 0.078$	7451	42 005	0.126	0.110	0.182	0.433
15	$a = 8.382$, $b = -1.899$, $n = -0.443$, $m = 0.341$	10120	70 231	0.189	0.160	0.213	0.560
16	$a = 4.324$, $b = -0.246$, $n = -0.716$, $m = 0.309$	11328	56 018	0.171	0.526	0.225	0.500
17a-c	$a_1 = 0.003$, $a_2 = -0.783$, $a_3 = 60.127$ $b_1 = 0.015$, $b_2 = -4.045$, $b_3 = 312.846$	23215	83 144	0.237	0.429	0.322	0.610
18	$a_1 = 0.046$, $a_2 = -6.936$, $a_3 = 418.709$ $b_1 = 0.000$, $b_2 = 0.003$, $b_3 = 0.356$	14567	78 407	0.200	0.183	0.255	0.592
$t_d \leq 60$ min							
12	$a = 0.992$, $b = 0.623$, $n = -0.577$	3445	68 692	0.261	0.465	0.162	0.722
13	$a = -0.350$, $b = 4.000$, $n = -0.509$	1386	95 278	0.166	0.165	0.103	0.850
14	$a = 16.169$, $b = -11.547$, $n = -0.577$, $m = 0.063$	789	31 933	0.098	0.091	0.077	0.492
15	$a = 8.464$, $b = -2.122$, $n = -0.438$, $m = 0.331$	861	60 393	0.163	0.137	0.081	0.677
16	$a = 3.955$, $b = -0.566$, $n = -0.794$, $m = 0.231$	1179	40 241	0.107	0.146	0.095	0.553
17a-c	$a_1 = 0.015$, $a_2 = -1.747$, $a_3 = 72.597$ $b_1 = 0.088$, $b_2 = -9.999$, $b_3 = 387.700$	297	21 022	0.093	0.097	0.047	0.399
18	$a_1 = 0.083$, $a_2 = -10.347$, $a_3 = 465.369$ $b_1 = 0.000$, $b_2 = 0.004$, $b_3 = 0.364$	544	58 786	0.155	0.135	0.064	0.668
$t_d \geq 60$ min							
12	$a = 0.417$, $b = 0.278$, $n = -0.712$	7977	10810	0.244	0.343	0.554	0.646
13	$a = -0.894$, $b = 6.365$, $n = -0.406$	11571	13711	0.229	0.282	0.668	0.727
14	$a = 2.023$, $b = -0.112$, $n = -0.819$, $m = 0.425$	4304	5032	0.136	0.131	0.407	0.441
15	$a = 0.200$, $b = 6.027$, $n = -1.392$, $m = 0.797$	2077	4039	0.124	0.118	0.283	0.395
16	$a = 1.725$, $b = -0.085$, $n = -0.855$, $m = 0.377$	4164	4798	0.122	0.118	0.401	0.430
17a-c	$a_1 = 0.001$, $a_2 = -0.196$, $a_3 = 27.987$ $b_1 = 0.004$, $b_2 = -1.017$, $b_3 = 147.275$	4772	6147	0.191	0.239	0.429	0.487
18	$a_1 = 0.007$, $a_2 = -0.399$, $a_3 = 124.513$ $b_1 = 0.000$, $b_2 = 0.003$, $b_3 = 0.257$	1743	2923	0.108	0.108	0.259	0.335

*)referred to F (Eq. 19), **)referred to F (Eq. 20).

tion should be performed. Because of the relatively high number of parameters and thus very low efficiency of systematic search method, and due to the polymodal nature of the objective function (recognized in preliminary tests) and thus low efficiency of local search methods, more efficient optimization technique is desirable. Therefore, the CRS2 method for global optimization was applied.

6.4 CRS2 method application

The CRS2 method was applied to find the parameters of a global IDF-relationship described by formulas:

for $t_d \leq 60$ min: $$q = -a \ln p + b \ , \tag{24}$$

for $t_d \geq 60$ min: $$q = a\, p^{-b} \ , \tag{25}$$

where $$a = a_1 t_d^2 + a_2 t_d + a_3 \ , \tag{26}$$

$$b = b_1 t_d^2 + b_2 t_d + b_3 \ . \tag{27}$$

The parameters a_i and b_i ($i = 1, 2, 3$) may take different values for Eqs. 24 and 25, and thus the optimal values of twelve parameters were searched. The limitations for the values of parameters defining the search domain were assumed on the basis of preliminary tests (see Section 6.3). According to Eqs. 26 and 27, the parameters a_i and b_i ($i = 1, 2, 3$) are related to individual components of the quadratic relations, wherein the values of t_d vary in a relatively wide range (from 10 to 240 min). Moreover, a and b describe ether logarithmic or exponential relation between q and p. These factors cause the parameters a_i and b_i of significantly different magnitudes and variations in different ranges. Based on the preliminary stage of analyses, the limitations of the possible parameter values, defining the search domain, were finally determined as:

for $t_d < 60$ min:

$a_1 \in <-10, 10>$, $a_2 \in <-20, 20>$, $a_3 \in <-2000, 2000>$,
$b_1 \in <-10, 10>$, $b_2 \in <-10, 10>$, $b_3 \in <-2000, 2000>$;

for $t_d \geq 60$ min:

$a_1 \in <0, 1>$, $a_2 \in <-20, 0>$, $a_3 \in <0, 2000>$,
$b_1 \in <-0.1, 0.1>$, $b_2 \in <-1, 1>$, $b_3 \in <0, 2>$.

The surrounding of the most probable location of the optimum was deliberately extended into a wider domain, in case the solution found in Section 6.3 (with the use of relatively not effective methods) was not the global optimum.

In order to decrease the number of the parameters to be estimated, the optimization was carried out in two steps: in the first, Eq. 24 for rainfall data $t_d \leq 60$ min, in the second – Eq. 25 for $t_d \geq 60$ min were optimized. For random choice of the trial points, sampling with no replacement according to the uniform probability distribution was applied. The relation $N = 10\,(n + 1)$ was adopted and the points were stored in an array **A** with the use of bubble sort algorithm. As the end-of-optimization criterion the following formula was applied:

$$\left|F_{\text{av}} - F_L\right| < eps \ , \tag{28}$$

where *eps* is the predetermined value, F_L is the value of objective function for the best point in array **A**, and F_{av} denotes the average value of objective function calculated for each of the N points of **A**:

$$F_{\text{av}} = \frac{\sum_{j=1}^{N} F_j}{N} \ . \tag{29}$$

The criterion defines the situation in which the solution is not corrected any more in iterative process, as all next simplexes are constituted of the points of very similar values of objective functions. However, two additional stop criteria were also included, in order to enable to break the calculations if the total number of iterations or/and total number of objective function runs ("cost") was too high. However, the experiments proved the criterion 29 was conclusive in the analyzed problem.

The method was applied with $eps = 0.05$ at first. The CRS2 procedure was applied ten times to recognize the effectiveness of the optimization in terms of finding the global optimum. As the values of optimized parameters for each attempt were very similar, for the final CRS2 run the search domain was narrowed to the closest surrounding of the expected optimum and the value of *eps* was decreased to 0.01. A few representative examples of the results of optimization runs and the final optimal parameter sets for $t_d \leq 60$ min and $t_d \geq 60$ min are presented in Table 7. The values indicate relatively small differences between the parameters (and thus also corresponding error values) obtained in various optimization runs. In fact, the method found global optimum (or at least a point very close to it) in each run. This may confirm the credibility of the results and possibility to decrease the total number of runs leading to satisfactory solution. Taking into account the total number of starting points (here: 70 points in each run), very short duration of calculations (a few minutes for each run) and the obtained result – the method may be considered efficient and the solution reliable.

Table 7

Exemplary results of parameter optimization
with use of CRS2 method for $t_d \leq 60$ min and $t_d \geq 60$ min

No.	Optimal values of parameters [—]						F (Eq. 20) [dm^3/(s·ha)]^2	ε (Eq. 21) [dm^3/(s·ha)]	Cost (number of obj. function runs) [—]
	a_1	a_2	a_3	b_1	b_2	b_3			
$t_d \leq 60$ min									
eps = 0.05									
1	0.011302	–1.5006	69.279	0.0757	–9.0489	374.919	20 651.710	0.395	14 086
2	0.011283	–1.4986	69.245	0.0756	–9.0410	374.783	20 651.704	0.395	14 896
3	0.011273	–1.4985	69.250	0.0756	–9.0386	374.773	20 651.717	0.396	14 187
4	0.011298	–1.5004	69.262	0.0757	–9.0480	374.867	20 651.724	0.396	15 106
eps = 0.01									
5	0.011306	–1.5005	69.263	0.0757	–9.0471	374.845	20 651.668	0.395	6 069
$t_d \geq 60$ min									
eps = 0.05									
1	0.017787	–2.7625	247.618	0.0000	0.0029	0.272	1 585.657	0.247	13 188
2	0.017740	–2.7551	247.367	0.0000	0.0029	0.272	1 585.678	0.247	15 064
3	0.017738	–2.7494	246.919	0.0000	0.0029	0.271	1 585.687	0.247	14 044
4	0.017788	–2.7614	247.600	0.0000	0.0029	0.271	1 585.664	0.247	12 819
eps = 0.01									
5	0.017767	–2.7563	247.284	0.0000	0.0029	0.271	1 585.654	0.247	9 828

For the optimal set of parameters, in the case of $t_d \leq 60$ the value of F (Eq. 19) was 281.8 mm^2 and the values of E_1 and E_2, defined by Eqs. 22 and 23, respectively, were 0.093 and 0.100. The value of ε referred to F (Eq. 19) was 0.046 mm. In the case of $t_d \geq 60$ min, the value of F (Eq. 19) was 927.1 mm^2, the values of E_1 and E_2 were equal to 0.110 and 0.115, respectively, and ε (referred to F; Eq. 19) was equal to 0.189 mm.

Finally, the application of global optimization technique CRS2 enabled determination of the IDF-relationship for GUT station. The obtained formulas are:

for $t_d < 60$ min: $$q = -a \ln p + b \ , \tag{30a}$$

where $$a = 0.011306\, t_d^2 - 1.500482\, t_d + 69.2635 \ , \tag{30b}$$

$$b = 0.075719\, t_d^2 - 9.04707\, t_d + 374.845 \ , \tag{30c}$$

for $t_d \geq 60$ min: $$q = a \cdot p^{-b} \ , \tag{31a}$$

where $$a = 0.017767\, t_d^2 - 2.756281\, t_d + 247.28426\ , \tag{31b}$$

$$b = 0.002902\, t_d + 0.271315\ . \tag{31c}$$

The obtained intensity-frequency curves for different values of aggregation level are shown in Fig. 7. Intensity-duration curves for selected values of p are presented in Fig. 8.

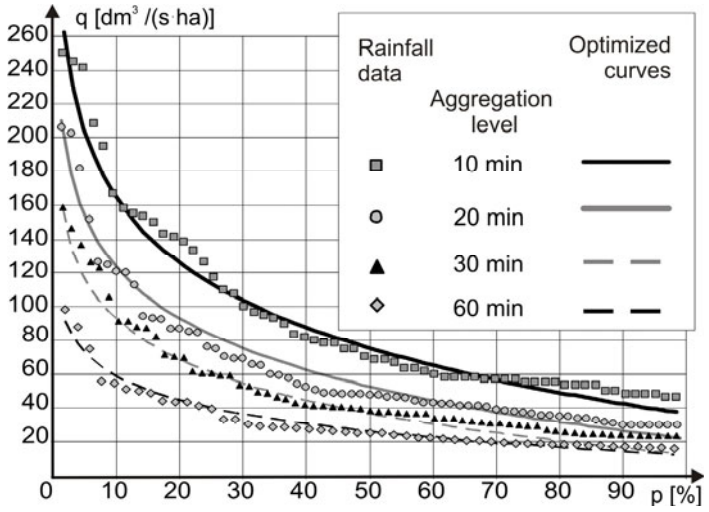


Fig. 7. Intensity-frequency curves for GUT station in Gdańsk (Poland).

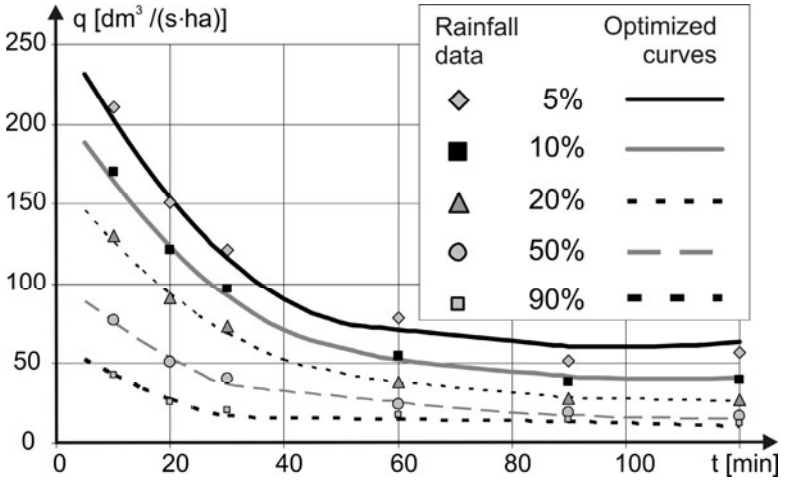


Fig. 8. Intensity-duration curves for GUT station in Gdańsk (Poland).

6.5 Comparison with other optimization methods

In order to verify the legitimacy of the CRS2 method application in the considered case, and for the purpose of its comparison with other algorithms, generally accepted for this type of problem, suitable numerical tests were carried out. In each of the tests the parameters a_1, a_2, a_3, b_1, b_2, and b_3 in Eqs. 26 and 27 were estimated (in the search domain defined as in the case presented in Section 6.4.) with use of different optimization method, but with the same accuracy requirements (defined in stopping criterion). Due to the different nature of the selected methods, the single iteration of the procedure was differently defined in each case. Thus, to enable objective comparison of the efficiency of the methods, the total cost of optimization, the values of the objective function F (Eq. 20), mean error ε (Eq. 21), and average relative errors E_1 (Eq. 22) and E_2 (Eq. 23) were compared for "optimal" sets of the estimated parameters. Each optimization method was applied at least ten times (increasing the number of runs was applied only if the results could affect the conclusions of the analysis). Selected representative examples of the results are presented in Table 8.

In the first stage, the parameters were estimated with the use of "successive search method". Because of a very low effectiveness of the algorithms, strongly depending on the size of the search domain and the length of the search step, the procedure was carried out in several stages, with successively narrowed search domain and reduced exploration step. However, even when the search step for each parameter was established only as 1/10 of the range of values defining boundaries of the search domain, in the case of six parameters to be estimated, the cost of 11^6 at each stage was needed

Table 8

Exemplary results of parameter optimization with use of systematic search (SS), Powell method (PM), modified Powell method (MPM), Newton optimization method (NOM), and evolutionary algorithm (EA) for $t_d \leq 60$ min and $t_d \geq 60$ min

Test no.	Optimal values of parameters [—]						F (Eq. 20) [dm^3/(s·ha)]^2	ε (Eq. 21) [dm^3/(s·ha)]	Cost (number of obj. function runs) [—]
	a_1	a_2	a_3	b_1	b_2	b_3			
$t_d \leq 60$ min									
Successive search method									
SS1	–0.018000	0.6200	37.200	–0.070	1.4600	224.000	101874.377	0.879	>7E+06
SS2	–0.016000	0.9536	22.400	–0.048	1.6960	179.200	138142.264	1.024	>8E+06
SS3	–0.040000	2.5920	11.200	–0.152	8.8000	128.000	172504.429	1.144	>91E+06
SS4	0.011000	–1.4800	69.200	0.075	–9.0000	375.000	20667.316	0.396	1.77E+06

to be continued

Table 8 (continuation)

Powell method									
PM1	0.227188	–24.1799	414.772	0.8546	–91.7490	1627.687	6492794.640	7.010	385
PM2	–0.438900	15.7180	24.590	–1.3180	37.2449	388.400	19640656.910	12.210	379
PM3	–0.118034	12.3048	–141.716	0.5000	–18.9227	259.620	24533293.161	13.645	227
PM4	0.009999	–1.5000	69.382	0.0700	–9.0000	375.000	21606.116	0.405	62
Modified Powell method									
MPM1	–0.147538	8.9843	–43.084	–0.3832	20.497	66.714	1002455.546	2.758	52917
MPM2	0.431536	–29.5952	375.202	1.8811	–135.7820	1826.212	9667618.690	8.566	54439
MPM3	0.502604	–36.2514	582.779	1.3530	–102.5665	1918.379	16688799.303	11.244	56279
MPM4	–0.021248	–6.1265	214.418	–0.4836	1.5628	631.197	9628908.819	8.548	25398
Newton search method									
NOM1	0.014780	–1.7474	72.135	0.0891	–10.0000	385.922	20902.514	0.398	120
NOM2	0.014609	–1.7487	72.613	0.0884	–9.9999	387.699	20946.700	0.399	252
NOM3	–0.058240	3.4380	12.068	–0.1925	10.0000	154.274	138682.640	1.026	120
NOM4	0.011284	–1.4992	69.259	0.0756	–9.0400	374.831	20603.962	0.395	672
Evolutionary algorithm									
EA1	–0.023090	1.0852	35.403	–0.0478	0.2859	251.027	54802.212	0.645	28780*
EA2	–0.014320	0.3238	47.772	–0.0214	–2.137	293.491	36249.521	0.524	28461*
EA3	–0.023595	1.0843	36.569	–0.0497	0.282	256.861	52920.362	0.634	31302*
EA4	–0.022700	0.9357	40.762	–0.0530	0.168	266.765	48456.261	0.606	16070*
$t_d \geq 60$ min									
Successive search method									
SS5	0.000800	–0.2400	172.800	0.000	0.0000	0.534	5412.163	0.456	>7E+06
SS6	0.000000	–0.1280	172.800	0.000	0.0000	0.533	5509.966	0.461	>7E+06
SS7	0.000000	–0.2000	170.000	0.000	0.0000	0.500	5713.677	0.469	>8E+06
SS8	0.017000	–2.7000	248.000	0.000	0.0028	0.280	1596.869	0.248	1.77E+06
Powell method									
PM5	0.017000	–2.7500	251.766	0.0000	0.0029	0.270	1703.623	0.256	1541
PM6	–0.118000	10.500	591.000	0.0000	0.0000	1.000	761966.918	5.422	20013
PM7	0.019619	–1.4638	238.515	0.0007	0.0030	–1.979	91741.570	1.881	139
PM8	0.638100	–2.1253	515.044	0.0818	0.8810	1.831	143716.983	2.355	1329
Modified Powell method									
MPM5	–0.232999	–9.3968	1420.001	0.0025	–0.1682	0.928	102248.066	1.986	105200
MPM6	0.012747	–1.0791	238.299	0.0007	0.0029	–1.964	87257.190	1.835	72739
MPM7	0.148486	–1.8437	667.270	0.0073	–0.0738	0.338	133992.829	2.274	66990
MPM8	0.019999	–3.3971	309.408	0.0000	0.0000	0.682	32287.420	1.116	66095
Newton search method									
NOM5	0.007125	–0.3991	124.513	0.000	0.0028	0.257	2923.069	0.336	492
NOM6	0.999900	–6.5091	1588.983	0.000	0.0977	0.000	133526.940	2.270	48
NOM7	0.000000	–20.0000	0.000	0.000	1.0000	0.000	143717.015	2.355	24
NOM8	0.017756	–2.7516	246.986	0.000	0.0029	0.270	1586.663	0.247	288
Evolutionary algorithm									
EA5	0.059317	–9.9158	556.111	0.000	0.0047	0.211	5280.886	0.452	39705*
EA6	0.038975	–5.7779	357.446	0.000	0.0047	0.174	3007.959	0.341	230259*
EA7	0.036106	–6.1234	400.184	0.000	0.0036	0.254	2833.594	0.330	28936*
EA8	0.029875	–4.8584	344.224	0.000	0.0036	0.227	2083.982	0.284	33705*

*) number of "subproblems" (solver evolutionary algorithm)

to find the "optimum", while the accuracy of such optimization was still very low. The effectiveness of the algorithm depended on the search step, the manner in which the search domain was successively narrowed at each subsequent stage of the calculations and the total number of stages. However, in all the tests conducted the process was very time-consuming and did not guarantee the success in finding the global optimum. In many cases the global optimum was omitted due to relatively large search step (Table 8, tests SS1, SS2, SS3, SS5, SS6, and SS7). The algorithm enabled obtaining the global optimum (with the objective function only slightly higher than in the case of CRS2 method) only in the case of the search domain initially narrowed to the closest neighborhood of the global optimum (Table 8, cases SS4 and SS8). However, even then the total cost of such search was much higher compared to CRS2 and other methods (more than a million objective function evaluations). In a typical situation, such precise localization of the global optimum is not known. Therefore, the only way to obtain a satisfactory result with the use of this method would be one-stage-search with a very small step. However, the cost of such calculation and the time required to obtain the solution would be unacceptable. Therefore, this algorithm may be useful only for rough estimation at the preliminary stage of the calculations and is not recommended in this case.

The second approach to the considered optimization problem was the "Powell method" implementation (Powell 1964, Fletcher 1965, Pierre 1969, Press *et al.* 2007). The typical, widely known concept of the conjugate directions was applied (Powell 1964), according to the procedures presented by Press *et al.* (2007). The initial bracketing of a minimum and the golden section search algorithm was applied as the procedures of one-dimension search for each of the conjugate directions (Press *et al.* 2007). The method required the specification of the initial point from which the searching procedure would start, thus the tests with different starting points were carried out. In each single iteration, the one-dimensional search along each of the conjugate directions was carried out, with the number of objective function evaluations depending on the function properties and required accuracy.

In the case of Powell method application, no satisfactory solution was obtained, as the impact of multimodal character of the objective function was clearly highlighted. In each case of the method run (from different starting points), premature convergence to the closest local optimum was observed. Thus, depending on the place in the search domain from which the optimization procedure was started, different optima were obtained (Table 8, exemplary tests PM1-PM8). The cost of the optimization was relatively low. Despite this, the quality of the solution was unacceptable. The values of the objective function in "optimal" points for most procedure runs were high. For example, for $t_d \leq 60$ min F was even higher than 24 000 000 (Table 8,

test PM3), while the corresponding value for the CRS2 algorithm was 20 651. The relatively good solution was obtained only when the starting point was close to the global optimum (tests PM4 and PM5). In typical situations, such a "lucky" guess is very difficult or impossible to provide. Therefore, the local search procedure (Powell method or any other) is not recommended in the analyzed problem.

In order to improve the quality of optimization using Powell method, the modification of the procedure was introduced. The algorithm of the local search (Powell method) was combined with random procedure of selecting the initial points in the manner similar to initial step of the CRS2 method. In this approach, the initial set of $N = 10\,(n + 1)$ randomly chosen points was established and the Powell method was run from each of the points. Such a test was then repeated several times. Unfortunately, as there was no additional mechanism of random interference in the optimization process (apart from the initial stage), the success of such approach depended on the "chance" of the draw of the initial points (Table 8, exemplary tests MPM1-MPM8) and the algorithm still suffered from all the inconveniences described before. Additionally, the cost of such optimization was much higher compared with CRS2 method. In all tests the solution was of much worse quality than the one obtained with the use of the CRS2 method, which enabled the solution very close to the global optimum in each run.

The final step was to test the effectiveness of the optimization procedures accessible in the embedded solver tools of commonly used spreadsheet packages. The two procedures –Newton search and the evolutionary algorithm – were applied. In the case of Newton method, the procedure was very fast and enabled obtaining "optimal" solution (in the sense of local optimization) in a short time, but the quality (effectiveness) of such optimization (expressed in the value of the objective function calculated for the optimal point) depended on the initial point of the optimization process. Initial points for the tests were selected either deliberately (*e.g.*, the central point of the domain, points close to the global optimum, points previously selected for CRS2 or Powell methods) or randomly. The procedure proved more effective than the Powell method. For some of the starting points, for which the Powell procedure suffered from the premature convergence, the satisfactory solution was obtained. In some cases, the optimum found by the procedure was very close to the one obtained by the CRS2 method (Table 8, tests NOM1-NOM2, NOM5). The further improvement was possible (*e.g.*, Table 8, best results – tests NOM4 and NOM8) when the additional run of the algorithm was carried out started from the previously found "optimum". However, for many other trials, in which the starting points were less favorable, the false "optimum" was found (with the value of objective function close to 100 000 or more, *e.g.*, tests NOM3, NOM6, NOM7). The tests have

shown that this method of optimization may be effective, but the success in finding the global optimum depends on the starting point from which the procedure is initiated.

The comparison of the two approaches – CRS2 method (used once) and Newton search (applied many times) – shows several significant differences. In the latter approach, each single run is much quicker, however does not guarantee the success (in terms of the "guarantee" that non-deterministic method CRS2 can provide; see Section 6.4). Moreover, the restart of Newton search – in the approach applied in the study – suffers from "subjectivity" and some onerousness, because each starting point must be selected by the user, which may be done according to different criterions. What is more, in a typical situation (when the global optimum is not known), it is difficult to judge when to stop restarting the procedure of local optimization and how "good" the so-far "optimal" solution really is. The CRS2 method is more "objective", as it proposes random selection of a relatively high number of initial points (recognized by Price (1983) as sufficient). It also leads to satisfactory solution in each run. In order to compare the results of the methods in the same initial conditions, $10\,(n+1)$ runs of Newton search should be performed (for each run of CRS2 method). That would be much more time-consuming and burdensome, and thus the automation would be necessary. The "cost" of such optimization would be comparable but the time of calculations – longer.

To the authors mind, the CRS2 method is more reliable. The Newton search approach is worth considering, especially in relatively simple applications; however, it should be applied carefully, as there is always the risk of premature convergence.

The last of the analyzed approaches was the application of the evolutionary algorithm available in Excel 2010 package. Unfortunately, this method occurred to be ineffective in the case considered. The evolutionary procedure conducted on the population of $N = 10\,(n+1) = 70$ individuals, did not lead to global optimum in any of the tests (more than 10 trials for each t_d range), although in some of them it was relatively close, compared to successive search and Powell method (Table 8, exemplary results EA1-EA8). As the common spreadsheet packages enable only limited influence on the parameters of the procedures, the improvement of the effectiveness in this case is very difficult. On the other hand, the considered problem (the estimation of the IDF-formula parameters) is not complex enough to apply very sophisticated methods of optimization, especially when other algorithms (*e.g.*, the CRS2 procedure) occurred to be effective. The tests showed that the best approach to the case considered was the method which combined both random and deterministic search in order to overcome all the inconveniences encountered in the calculations.

7. SUMMARY AND CONCLUSIONS

IDF-relationships, in spite of their "simple" nature, play an important role in many engineering applications. The choice of IDF-formulas for design rainfall determination instead of more sophisticated analysis, is often a compromise between practical tendency to use simplified approaches (leading to solution in short, relatively easy and economically beneficial way), and the problem complexity and computational requirements. The presented study has developed the modified empirical approach to IDF-relationship development, based on the global optimization technique of controlled random search applied to the estimation of the parameters of the previously determined functions. The approach is relatively simple and efficient and it enables obtaining satisfactory solution in a relatively short time. In the problem of IDF-formulas determination – seen globally, as an engineering task – the most time-consuming is the first step – rainfall data collection and analysis. Currently available methods of parameter search may enable quick access to efficient solutions, even (or especially) in those cases where the global form of IDF-formula is more complex and typical theoretical distributions do not lead to satisfactory results, as the measured data do not "match" typical hydrological distributions.

The proposed procedure of IDF-relationship determination has proven effective in the case when commonly used simple formulas do not allow an acceptably good fit between simulated rainfall intensities and observations in the whole range of considered values of t_d and p. In this particular case, the "combined" relationship $q(t_d, p)$, in the form more complex than commonly used IDF-formulas, had to be taken into account and the higher number of parameters had to be identified. Therefore, global optimization method was applied. The proposed procedure CRS2 is obviously not the only one possible to use in this case. Numerous experiences in different fields of engineering enabled the development of many optimization procedures, from relatively simple to very sophisticated, dedicated to different kinds of problems. Therefore, the choice of the method should result from the complexity of the task (mainly the features of the objective function), as the chosen algorithm should enable obtaining satisfactory solution with reasonable labor input, corresponding to the problem complexity. The CRS2 algorithm applied in the study was effective in finding global minimum of the assumed function, although it was still simple, compared to further versions of the procedure (CRS3 and higher) and other global optimization techniques. Commonly used approaches to IDF-relationship determination (based on formulas 1-3) produced worse results in the considered case. Although the solution was achieved relatively easily and quickly, the goodness-of-fit of the obtained relationships to observations was much poorer, which was ana-

lyzed in the preliminary stage of IDF-formulas determination. Local optimization procedures (*e.g.*, the widely recommended Powell method) applied to parameter identification in case of more complex IDF-function, led to different, not always optimal solutions, depending on the assumed starting point, due to premature convergence to local optima. Systematic search procedure finally led to the correct result but the process was much more time-consuming. Evolutionary algorithm applied to the analyzed case was also ineffective. Such algorithms are also much more time-consuming and – because of their random nature – it is very difficult to judge whether the solution is really "optimal" in global meaning. The effectiveness of evolutionary algorithms may be improved by detailed recognition of the most suitable (to the analyzed case) procedures of selection, mutation and crossover. However, for the purpose of calculations considered in this paper (IDF-formulas determination), such effort seems redundant and application of such methods – unfounded. The CRS methods, which had been analyzed in detail and compared with selected global search procedures (Ali *et al.* 1997a), proved to be effective in many complex applications. Therefore, the choice of CRS2 in the study was a compromise between labor input and the expected quality of the results.

Acknowledgements. The author would like to thank Associate Professor Dr. Elżbieta Wołoszyn from the Faculty of Civil and Environmental Engineering of the Gdańsk University of Technology (now retired) for sharing her results of the chronological analysis of precipitation data recorded in GUT gauge station during the period 1991-2008. The author would also like to acknowledge the efforts of Katarzyna Rzeszutek (Msc. Eng.), whose work on a master's degree ("Analysis of temporal diversity of precipitation on the basis of 20-year observations on GUT climatologic station", Gdańsk 2011; in Polish) has been helpful at the stage of preparation and analysis of data in this study.

References

Akan, A.O., and R.J. Houghtalen (2003), *Urban hydrology, Hydraulics, and Stormwater Quality: Engineering Applications and Computer Modeling*, John Wiley & Sons, Inc., Hoboken.

Ali, M.M. (1994), Some modified stochastic global optimization algorithms with applications, Ph.D. Thesis, Loughborough University of Technology, Loughborough, England, https://dspace.lboro.ac.uk/2134/13429.

Ali, M.M., and C. Storey (1994), Topographical multilevel single linkage, *J. Global Optim*. **5**, 4, 349-458, DOI: 10.1007/BF01096684.

Ali, M.M., and C. Storey (1997), Aspiration based simulated annealing algorithm, *J. Global Optim*. **11**, 2, 181-191, DOI: 10.1023/A:1008202703889.

Ali, M.M., C. Storey, and A. Törn (1997a), Application of stochastic global optimization algorithms to practical problems, *J. Optimiz. Theory Appl*. **95**, 3, 545-563, DOI: 10.1023/A:1022617804737.

Ali, M.M., A. Törn, and S. Viitanen (1997b), A numerical comparison of some modified controlled random search algorithms, *J. Global Optim*. **11**, 4, 377-385, DOI: 10.1023/A:1008236920512.

Ariff, N.M., A.A. Jemain, K. Ibrahim, and W.Z. Wan Zin (2012), IDF relationships using bivariate copula for storm events in Peninsular Malaysia, *J. Hydrol*. **470-471**, 158-171, DOI: 10.1016/j.jhydrol.2012.08.045.

Arnell, V. (1982), Rain fall data for the design of sewer pipe systems, Report Series A:8, Chalmers University of Technology, Department of Hydraulics, Göteborg, Sweden.

Ben-Zvi, A. (2009), Rainfall intensity-duration-frequency relationship derived from large partial duration series, *J. Hydrol*. **367**, 1-2, 104-114, DOI: 10.1016/j.jhydrol.2009.01.007.

Bonnans, J.F., J.C. Gilbert, C. Lemaréchal, and C.A. Sagastizábal (2006), *Numerical Optimization: Theoretical and Practical Aspects*, Springer-Verlag, Berlin Heidelberg.

Chow, V.T. (1964), *Handbook of Applied Hydrology: A Compendium of Water-resources Technology*, McGraw–Hill Book Co., New York.

Coles, S. (2001), *An Introduction to Statistical Modeling of Extreme Values*, Springer Series in Statistics, Springer, London, 208 pp.

Cunnane, C. (1973), A particular comparison of annual maxima and partial duration series methods of flood frequency prediction, *J. Hydrol*. **18**, 3-4, 257-271, DOI: 10.1016/0022-1694(73)90051-6.

De Michele, C., and G. Salvadori (2005), Some hydrological applications of small sample estimators of Generalized Pareto and Extreme Value distributions, *J. Hydrol*. **301**, 1-4, 37-53, DOI: 10.1016/j.jhydrol.2004.06.015.

Dekkers, A., and E. Aarts (1991), Global optimization and simulated annealing, *Math. Program*. **50**, 1-3, 367-393, DOI: 10.1007/BF01594945.

Delleur, J.W. (2003), The evolution of urban hydrology: Past, present, and future, *J. Hydraul. Eng. ASCE* **129**, 8, 563-573, DOI: 10.1061/(ASCE)0733-9429(2003)129:8(563).

Dysarz, T., and J.J. Napiórkowski (2002), Flood control in Nysa Reservoir System by means of sequential optimisation and CRS method. **In:** *Proc. Workshop on Evolutionary Algorithms and Global Optimization, 23-25 September 2002, Cracow, Poland*, Warsaw Technical University Publ., Warszawa, 27-33.

Dysarz, T., and J.J. Napiórkowski (2003), Optimal flood control of Nysa Kłodzka Reservoir System, *Publs. Inst. Geophys. Pol. Acad. Sc*. **E-3**, 365, **(**J.J. Napiórkowski (ed.), *Modelling and Control of Floods*), 83-96.

El-Sayed, E.A.H. (2011), Generation of rainfall intensity duration frequency curves for ungauged sites, *Nile Basin Water Sci. Eng. J.* **4**, 1, 112-124.

Elsebaie, I.H. (2012), Developing rainfall intensity-duration-frequency relationship for two regions of Saudi Arabia, *J. King Saud Univ. – Eng. Sci.* **24**, 2, 131-140, DOI: 10.1016/j.jksues.2011.06.001.

Endreny, T.A., and N. Imbeah (2009), Generating robust rainfall intensity-duration-frequency estimates with short-record satellite data, *J. Hydrol*. **371**, 1-4, 182-191, DOI: 10.1016/j.jhydrol.2009.03.027.

Fletcher, R. (1965), Function minimization without evaluating derivatives – a review, *Comput. J*. **8**, 1, 33-41, DOI: 10.1093/comjnl/8.1.33.

Fylstra, D., L. Lasdon, J. Watson, and A. Waren (1998), Design and use of the Microsoft Excel Solver, *Interfaces* **28**, 5, 29-55, DOI: 10.1287/inte.28.5.29.

Goldberg, D.E. (1989), *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison–Wesley Publ. Co., Reading, 412 pp.

Grace, R.A., and P.S. Eagleson (1967), A model for generating synthetic sequences of short-time-interval rainfall depths. **In**: *Proc. Int. Hydrology Symposium, 6-8 September 1967, Fort Collins, Colorado, USA*, Vol. 1, Colorado State University, Colorado, 268-276.

Grimaldi, S., and F. Serinaldi (2006), Design hyetograph analysis with 3-copula function, *Hydrol. Sci. J.* **51**, 2, 223-238, DOI: 10.1623/hysj.51.2.223.

Hailegeorgis, T.T., S.T. Thorolfsson, and K. Alfredsen (2013), Regional frequency analysis of extreme precipitation with consideration of uncertainties to update IDF curves for the city of Trondheim, *J. Hydrol*. **498**, 305-318, DOI: 10.1016/j.jhydrol.2013.06.019.

Johansen, L. (1979), Design rainfalls for sewer systems, Rep. 79-2, Dept. of Sanitary Eng., Technical University of Denmark, Copenhagen, Denmark.

Katz, R.W., M.P. Parlange, and P. Naveau (2002), Statistics of extremes in hydrology, *Adv. Water Resour*. **25**, 8-12, 1287-1304, DOI: 10.1016/S0309-1708(02)00056-8.

Kisiel, C.C., L. Duckstein, and M.M. Fogel (1971), Analysis of ephemeral flow in aridlands, *J. Hydraul. Div. ASCE* **97**, 10, 1699-1717.

Koutsoyiannis, D. (2004a), Statistics of extremes and estimation of extreme rainfall: I. Theoretical investigation, *Hydrol. Sci. J.* **49**, 4, 575-590, DOI: 10.1623/hysj.49.4.575.54430.

Koutsoyiannis, D. (2004b), Statistics of extremes and estimation of extreme rainfall: II. Empirical investigation of long rainfall records, *Hydrol. Sci. J*. **49**, 4, 591-610, DOI: 10.1623/hysj.49.4.591.54424.

Koutsoyiannis, D., D. Kozonis, and A. Manetas (1998), A mathematical framework for studying rainfall intensity-duration-frequency relationships, *J. Hydrol.* **206**, 1-2, 118-135, DOI: 10.1016/S0022-1694(98)00097-3.

Langousis, A., and D. Veneziano (2007), Intensity-duration-frequency curves from scaling representations of rainfall, *Water Resour. Res*. **43**, 2, W02422, DOI: 10.1029/2006WR005245.

Licznar, P., and J. Łomotowski (2005), Analysis of instantaneous intensities of design rainfalls in Wroclaw, *Ochr. Sr*. **27**, 2, 25-28 (in Polish).

Licznar, P., and J. Łomotowski (2007), *Rainfall Kinetic Energy Measurements with Impactometer Implementation*, Works and Studies, Vol. 69, Institute of Environmental Engineering, Polish Academy of Sciences, Zabrze, 70 pp.

Licznar, P., T.G. Schmitt, and D.E. Rupp (2011a), Distributions of microanonical cascade weights of rainfall at small timescales, *Acta Geophys*. **59**, 5, 1013-1043, DOI: 10.2478/s11600-011-0014-4.

Licznar, P., J. Łomotowski, and D.E. Rupp (2011b), Random cascade driver rainfall disaggregation for urban hydrology: An evaluation of six models and a new generator, *Atmos. Res*. **99**, 3-4, 563-578, DOI: 10.1016/j.atmosres.2010.12.014.

Llasat, M.-C. (2001), An objective classification of rainfall events on the basis of their convective features: application to rainfall intensity in the northeast of Spain, *Int. J. Climatol*. **21**, 11, 1385-1400, DOI: 10.1002/joc.692.

Lutgens, F.K., and E.J. Tarbuck (2004), *The Atmosphere. An Introduction to Meteorology*, 9th ed., Pearson Education Inc., New Jersey.

Madsen, H., P.F. Rasmussen, and D. Rosbjerg (1997a), Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events: 1. At-site modeling, *Water Resour. Res*. **33**, 4, 747-757, DOI: 10.1029/96WR03848.

Madsen, H., C.P. Pearson, and D. Rosbjerg (1997b), Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events: 2. Regional modeling, *Water Resour. Res*. **33**, 4, 759-769, DOI: 10.1029/96WR03849.

Manzanares-Filho, N., and R.B.F. Albuquerque (2008), Accelerating controlled random search algorithms using a distribution strategy. **In**: *Proc. Int. Conf. Engineering Optimization (EngOpt 2008), 1-5 June 2008, Rio de Janeiro, Brazil*.

Manzanares-Filho, N., C.A.A. Moino, and A.B. Jorge (2005), An improved controlled random search algorithm for inverse airfoil cascade design. **In:** *Proc. 6th World Congress of Structural and Multidisciplinary Optimization (WCSMO-6), 30 May – 3 June 2005, Rio de Janeiro, Brazil*, paper No. 4451.

Maraun, D., T.J. Osborn, and N.P. Gillett (2008), United Kingdom daily precipitation intensity: improved early data, error estimates and an update from 2000 to 2006, *Int. J. Climatol*. **28**, 6, 833-842, DOI: 10.1002/joc.1672.

Marsalek, J. (1978), Research on the design storm concept, ASCE Urban Water Resources Research Program, Technical Memorandum No. 33, New York.

McCuen, R.H. (2005), *Hydrologic Analysis and Design*, 3rd ed., Pearson Prentice Hall, Englewood Cliffs, 859 pp.

Molnar, P., and P. Burlando (2005), Preservation of rainfall properties in stochastic disaggregation by a simple random cascade model, *Atmos. Res.* **77**, 1-4, 137-151, DOI: 10.1016/j.atmosres.2004.10.024.

Mutzner, H. (1991), The significance of areal rainfall distribution for flows from a very small urban drainage catchment, *Atmos. Res*. **27**, 1-3, 99-107, DOI: 10.1016/0169-8095(91)90011-K.

Nagy, J.T. (1991), Evaluation of rainfall characteristics in Bratislava, *Atmos. Res*. **27**, 1-3, 209-217, DOI: 10.1016/0169-8095(91)90020-W.

Pierre, D.A. (1969), *Optimization Theory with Applications*, John Wiley &Sons Inc., New York.

Powell, M.J.D. (1964), An efficient method for finding the minimum of a function of several variables without calculating derivatives, *Comput. J*. **7**, 2, 155-162, DOI: 10.1093/comjnl/7.2.155.

Press, W.H., S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery (2007), *Numerical Recipes: the Art of Scientific Computing*, 3rd ed., Cambridge University Press, New York.

Price, W.L. (1977), A controlled random search procedure for global optimisation, *Comput. J*. **20**, 4, 367-370, DOI: 10.1093/comjnl/20.4.367.

Price, W.L. (1978), Controlled random search procedure for global optimization. **In:** L.C.W. Dixon, and G.P. Szegö (eds.), *Toward Global Optimization 2*, North-Holland Publ. Co., Amsterdam, 71-84.

Price, W.L. (1983), Global optimization by controlled random search, *J. Optimiz. Theor. Appl*. **40**, 3, 333-348, DOI: 10.1007/BF00933504.

Pui, A., A. Sharma, R. Mehrotra, B. Sivakumar, and E. Jeremiah (2012), A comparison of alternatives for daily and sub-daily rainfall disaggregation, *J. Hydrol*. **470-471**, 138-157, DOI: 10.1016/j.jhydrol.2012.08.041.

Reiss, R.-D., and M. Thomas (2007), *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*, 3rd ed., Birkhäuser Verlag, Basel.

Rinaudo, S., F. Moschella, O. Muscato, and M.A. Ahile (1998), Controlled random search parallel algorithm for global optimization with distributed processes on multivendor CPUs. **In:** L. Arkeryd, J. Bergh, Ph. Brenner, and R. Petterson (eds.), *Proc. 10th Conf. European Consortium for Mathematics in Industry "Progress in Industrial Mathematics at ECMI 98", 23-28 June 1998, Gothenburg, Sweden*, 271-278.

Rinnooy Kan, A.H.G., and G.T. Timmer (1987), Stochastic global optimization methods. Part II: Multi level methods, *Math. Program*. **39**, 1, 57-78, DOI: 10.1007/BF02592071.

Romanowicz, R.J., M. Osuch, and M. Grabowiecka (2013), On the choice of calibration periods and objective functions: A practical guide to model parameter identification, *Acta Geophys*. **61**, 6, 1477-1503, DOI: 10.2478/s11600-013-0157-6.

Schilling, W. (1991), Rainfall data for urban hydrology: what do we need?, *Atmos. Res*. **27**, 1-3, 5-21, DOI: 10.1016/0169-8095(91)90003-F.

Sevruk, B., and H. Geiger (1987), Frequency distributions preferred by hydrologists. **In**: B.C. Yen (ed.), *Proc. 4th Int. Conf. in Urban Storm Drainage "Topics in Urban Drainage Hydraulics and Hydrology", Session D, 31 August – 4 September 1987, Lausanne, Switzerland*, 51-52.

Suligowski, R. (2004), Temporal and spatial structure of the rainfall in Poland. Attempt at the regionalization, Publ. of Jan Kochanowski University, Institute of Geography, Kielce (in Polish).

Svensson, C., R.T. Clarke, and D.A. Jones (2007), An experimental comparison of methods for estimating rainfall intensity-duration-frequency relations from fragmentary records, *J. Hydrol*. **341**, 1-2, 79-89, DOI: 10.1016/j.jhydrol. 2007.05.002.

Tanaka, S., and K. Takara (2002), A study of threshold selection in POT analysis of extreme floods. **In:** A. Snorasson, H.P. Finnsdóttir, and M. Moss (eds.), *Proc. Symp. "The Extremes of the Extremes: Extraordinary Floods", Reykjavik, Iceland*, IAHS Publications, No. 271, IAHS Press, Wallingford, 299-306.

Tavares, L.V., and J.E. Da Silva (1983), Partial duration series method revisited, *J. Hydrol*. **64**, 1-4, 1-14, DOI: 10.1016/0022-1694(83)90056-2.

Todorovic, P. (1978), Stochastic models of floods, *Water Resour. Res*. **14**, 2, 345-356, DOI: 10.1029/WR014i002p00345.

Trefry, C.M., D.W. Watkins, and D. Johnson (2005), Regional rainfall frequency analysis for the State of Michigan, *J. Hydrol. Eng. ASCE* **10**, 6, 437-449, DOI: 10.1061/(ASCE)1084-0699(2005)10:6(437).

Tsoulos, I.G., and I.E. Lagaris (2006), Genetically controlled random search: a global optimization method for continuous multidimensional functions, *Comp. Phys. Comm*. **174**, 2, 152-159, DOI: 10.1016/j.cpc.2005.09.007.

Vaes, G., P. Willems, and J. Berlamont (2001), Rainfall input requirements for hydrological calculations, *Urban Water* **3**, 1-2, 107-112, DOI: 10.1016/ S1462-0758(01)00020-6.

Venkata Ramana, R., B. Chakravorty, N.R. Samal, N.G. Pandey, and P. Mani (2008), Development of intensity duration frequency curves using L-moment and GIS technique, *J. Appl. Hydrol*. **21**, 1-2, 88-100.

Vukmirović, V., and J. Petrović (1991), Statistical analysis of storms – a basis for urban runoff modeling. **In:** Č. Maksimović (ed.), *Proc. of Int. Conf. "New Technologies in Urban Drainage UDT'91", 17-21 June 1991, Dubrovnik, Yugoslavia*, Elsevier, London, 13-19.

Walesh, S.G. (1989), *Urban Surface Water Management*, John Wiley & Sons, Inc., New York.

Wenzel, H.G. Jr., and M.L. Voorhees (1978), Evaluation of design storm concept (presented at the AGU Fall Meeting, December 1978, San Francisco, California), *EOS Trans. AGU* **59**, 12, 1070.

Willems, P. (2000), Compound intensity/duration/frequency-relationships of extreme precipitation for two seasons and two storm types, *J. Hydrol*. **233**, 1-4, 189-205, DOI: 10.1016/S0022-1694(00)00233-X.

Wołoszyn, E. (1991), Polish rainfall-runoff investigations and modification of the rational method, *Atmos. Res*. **27**, 1-3, 219-229, DOI: 10.1016/0169-8095 (91)90021-N.

Wołoszyn, E. (2003), The catastrophic flood in Gdańsk on July 2001. **In:** R. Arsov, J. Marsalek, E. Watt, and E. Zeman (eds.), *Urban Water Management: Science Technology and Service Delivery*, NATO Science Series, Vol. 25, Springer, Dordrecht, 115-124, DOI: 10.1007/978-94-010-0057-4_12.

Wołoszyn, E. (2009), Analysis of rainfall data of Gdansk Meteorological Station. **In:** C. Popowska and M. Jovanovski (eds.), *Proc. 11th Int. Symp. on Water Management and Hydraulic Engineering, 1-5 September 2009, Univ. Ss. Cyril and Methodius, Faculty of Civil Engineering, Orhid, Macedonia,* 675-684.

Yevjevich, V. (1972), *Probability and Statistics in Hydrology*, Water Resources Publ., Fort Collins, USA.

Received 10 July 2013
Received in revised form 23 May 2014
Accepted 9 June 2014

DE GRUYTER OPEN

Acta Geophysica
vol. 63, no. 1, Feb. 2015, pp. 275-300
DOI: 10.2478/s11600-014-0261-2

Exceptionally Hot and Cold Summers in Europe (1951-2010)

Robert TWARDOSZ[1] and Urszula KOSSOWSKA-CEZAK[2]

[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, Warsaw, Poland

A b s t r a c t

The paper explores exceptional thermal conditions, an area of research that has increased in significance in the context of the changes that are being observed in climate. Specifically, the study addressed the frequency, long-term change and spatial coverage of exceptionally hot summers, and exceptionally cool summers (EHS and ECS) in Europe. The statistical criterion of plus/minus two standard deviations from the long-term average was used to identify EHS and ECS at 60 weather stations over the period 1951-2010. The study has demonstrated that EHS are characterised by greater temperature anomalies than ECS and are approximately twice as frequent. They occurred virtually everywhere within the study area, whilst ECS are concentrated in its northern part. Five EHS (1972, 2002, 2003, 2007, and 2010) and three ECS (1956, 1962, and 1976) were large enough to be recorded by at least 10% of the stations.

Key words: cool summers, hot summers, temperature anomaly, Europe.

1. INTRODUCTION

According to the latest IPCC Report (IPCC 2013) the global mean surface air temperature (SAT) increased by 0.85 K between 1880 and 2012. This in-

crease was caused by, among other things, an increased frequency, intensity and duration of heat waves (Perkins *et al.* 2012). The occurrence of heat waves causes non steady-state statistical air temperature distribution. Hansen *et al.* (2012) showed that the distribution of seasonal mean temperature anomalies has shifted toward higher temperatures and the range of anomalies has increased. The number of local record-breaking monthly temperature extremes is now on average five times larger than expected in a climate with no long-term warming (Coumou *et al.* 2013). Kamae *et al.* (2014) clarified why hot summers were becoming more frequent despite the slowdown of global warming. They claimed that the explanation of this effect could be found in decadal variabilities in the Pacific and Atlantic Oceans and in the direct influence of anthropogenic forcing. In interesting recent studies, Sillmann *et al.* (2013, 2014) showed that spatial trend patterns differed between the warm and cold extremes, whereby the warm extremes showed continuous positive trends across the globe and the cold extremes exhibited a coherent cooling pattern across the Northern Hemisphere mid-latitudes. In Europe at the turn of the 21st century, exceptionally hot months and even entire summers also occurred. Kossowska-Cezak and Twardosz (2013) report that during this period exceptionally cool summers were less frequent and did not occur at all between 2001 and 2010 in the central and eastern part of the continent. They also claim that the current climatic warming is accompanied by an increase in temperature variance, which means that the warming is not produced through a complete disappearance of exceptionally cool summers, but primarily through an increased frequency of exceptionally hot summers. This effect is particularly strong in Central Europe where the warming effect is stronger than elsewhere on the continent (Moberg *et al*. 2006).

The occurrence of anomalous thermal conditions is a feature of the temperate zone in Europe, the climate of which is determined by a highly variable atmospheric circulation. Hot summer weather is directly caused by stationary anticyclones, which are normally associated with the advection of tropical air. The combination of reduced cloudiness typical of high-pressure weather and the long summer days produces long sunshine duration. The root causes of hot summers are typically linked with the thermal conditions of the ocean (Black and Sutton 2007), especially its surface layer (Della-Marta *et al.* 2007, Feudale and Shukla 2011, Dole *et al.* 2011, Zveryaev *et al.* 2012), and with the increased concentration of greenhouse gases in the atmosphere (Schär *et al.* 2004, Stott *et al.* 2004, Jones *et al.* 2008, Gruza and Ran'kova 2011).

Cool summer spells, on the other hand, are mostly a result of any of the following: (i) advection of fresh maritime air from the Atlantic Ocean, often accompanied by intensive rainfall; (ii) northern advection (Kuziemska 1975,

Kossowska-Cezak 1997, Jaagus 2006, Sidorenkov and Orlov 2008, Anisimov *et al.* 2011, Twardosz *et al.* 2011); or (iii) the cool advection involved in so-called "diving" Scandinavian cyclones (Elizbarashvili *et al.* 2007, Isayev and Sherstyukov 2008). The literature also mentions an effect of volcanic eruptions on cool summers, such as an anomalously low temperature in Europe in the summer of 1816 after the eruption of Tambora in 1815 (Piervitali *et al.* 1997, Trigo *et al.* 2009). Climatologists have given particular attention to long periods with very high temperatures, such as heat waves (Stott *et al.* 2004, Chase *et al.* 2006, Hutter *et al.* 2007, Kyselý and Huth 2008, Révész 2008, Founda and Giannakopoulos 2009, Rebetez *et al.* 2009, Kossowska-Cezak 2010, Kyselý 2010, Barriopedro *et al.* 2011, Dole *et al.* 2011, Krzyżewska and Wereski 2011, Twardosz 2009, Friedrich and Bissolli 2012, Twardosz and Batko 2012, Unal *et al.* 2013) or anomalously hot months and summers (Gerstengarbe and Werner 1992, Piervitali *et al.* 1997, Filipiuk and Kaszewski 2000, Luterbacher *et al.* 2004, Elizbarashvili *et al.* 2007, Kossowska-Cezak and Twardosz 2012a, b, Twardosz and Kossowska-Cezak 2013a, b, c, Wehry 2010). In view of the unfavourable biometeorological conditions, as well as droughts and wildfires, during such periods this field of research has great merit. Indeed, sustained heat stress has an exceptionally adverse influence on human health and well-being and has been demonstrated to lead to higher mortality rates (Błażejczyk and Mc Gregor 2007, Hutter *et al.* 2007, D'Ippoliti *et al.* 2010, Twardosz 2009, Muthers *et al.* 2010, Barriopedro *et al.* 2011, Zvyagintsev *et al.* 2011, Revich and Shaposhnikov 2012). Droughts tend to cause losses in agriculture and contribute to forest fires, sometimes over large areas, such as in the summers of 2003 and 2010 (Hodzic *et al.* 2007, Zvyagintsev *et al.* 2011).

The results of modelling research indicate that the trend to an increase in the frequency of these weather conditions may not only continue in future, but also that heat waves could become even more severe (Meehl and Tebaldi 2004, Jones *et al.* 2008, Révész 2008, Kürbis *et al.* 2009, Fischer and Schär 2010, Bardin 2011, Kamae *et al.* 2014, Sillmann *et al.* 2014). Perhaps we have already experienced a confirmation of these predictions during the events of 2003 and 2010, which have been classified as "mega-heatwaves" (Barriopedro *et al.* 2011).

Much less attention has been given to exceptionally cool summer months and seasons. Their effects on the economy primarily include smaller and delayed harvests of poorer quality produce, as reported in the cold and wet summer of 1816 (Trigo *et al.* 2009). As far as tourism is concerned, such summers affect not only the comfort of holidaymakers, but also their safety in the case of floods (de Freitas 2003, Scott and Lemieux 2010). Cool summers and summer months were particularly frequent in the 19th and early 20th century (Gerstengarbe and Werner 1992, Filipiuk and Kaszewski 2000,

Elizbarashvili *et al.* 2007) and the two coolest summers in recent history were 1816 (dubbed "the year without a summer") (*e.g.*, Piervitali *et al.* 1997) and 1913 (Filipiuk and Kaszewski 2000). Cool summer seasons did also occur recently, such as in 2009 in the USA (Hansen *et al.* 2012).

This study was designed to learn about the frequency, duration (frequency change) and spatial extent of exceptionally hot and exceptionally cool summers (EHS and ECS). The working hypothesis was that the current climatic warming has been manifested in a decreasing frequency of exceptionally cool summers and an increasing frequency of exceptionally hot summers. The study builds on the authors' earlier research (Kossowska-Cezak and Twardosz 2012a, b, Twardosz and Kossowska-Cezak 2013a, b, c) that focused on the central-eastern part of the continent and expands it to the entire continent of Europe. Such a comprehensive approach to thermal anomalies using Europe-wide observation data is a new contribution to a body of research that otherwise includes multiple studies on specific areas and different periods.

2. DATA AND METHODOLOGY

The study is based on average seasonal and monthly temperatures in summer (June to August) and on average maximum and minimum temperatures recorded at 60 stations over a 60-year period between 1951 and 2010 (Table 1). The data were obtained from the on-line European Climate Assessment & Dataset (ECA&D) (www.eca.knmi.nl). This database is one of four publicly available data sources (Moberg *et al.* 2006). It also stands out from the other such data bases, *e.g.*, CRUTEM4 (Jones *et al.* 2012) and HadEX2 (Donat *et al.* 2013), with its superior spatial resolution and it has been used to build other data bases, *e.g.*, HadEX2 (Donat *et al.* 2013). The ECA&D data come directly from national weather services (Klein Tank *et al.* 2002). They have been verified for uniformity using four statistical tests (Wijngaard *et al.* 2003) the outcomes of which are available directly from the ECA&D website. Some data gaps found at a small number of stations have been filled in using records from neighbouring stations processed with a generally established climatological method (quotient method).

The stations selected for the study cover the entire territory of the continent, from the Mediterranean to the northern end of the Scandinavian Peninsula and from the Atlantic Ocean to the Ural Mountains, and the British Isles. Mountain weather stations were excluded. Among the 60 stations, 35 were located at altitudes up to 100 m above sea level (including Astrakhan in a depression at –23 m), 17 stations at between 101 and 200 m, and 5 stations at between 201 and 300 m, which gives a total of 57 stations in the lowlands (Table 1, Fig. 1). Only 3 stations were located in uplands: Madrid (667 m),

Table 1

Long-term average summer temperatures and the numbers of EHS and ECS in Europe (1951-2010)

Station		WMO no.	t_{av} [°C]	ECS	EHS	Station		WMO no.	t_{av} [°C]	ECS	EHS
No.	Name					No.	Name				
1	Lisbon	8535	21.9	2	1	31	De Bilt	6260	16.5	1	1
2	Almeria	8487	24.8	1	1	32	Berlin	10381	17.7	–	2
3	Crotone		24.8	–	2	33	Warsaw	12375	17.8	–	3
4	Athens	16716	27.3	2	1	34	Minsk	26850	17.3	–	3
5	La Coruña	8001	18.2	1	2	35	Kiev	33345	19.3	–	1
6	Madrid	8223	23.4	1	1	36	Kursk	34009	18.4	–	2
7	Bordeaux	7510	19.5	–	1	37	Saratov	31172	21.0	–	2
8	Barcelona	8181	22.4	–	2	38	Orenburg	35121	20.9	1	2
9	Marseilles	7650	23.0	–	1	39	Edinburgh	3160	14.1	1	–
10	Rome	16235	23.4	–	2	40	Oslo	1492	15.8	–	4
11	Split	14445	24.8	1	1	41	Copenhagen	6180	17.0	2	1
12	Belgrade	13274	21.7	–	3	42	Stockholm	2485	16.5	1	1
13	Sofia	15614	19.6	1	2	43	Liepaja	26406	16.1	2	2
14	Konstanca	15480	21.8	–	2	44	St. Petersburg	26063	16.8	2	2
15	Istanbul	17062	23.1	–	2	45	Moscow	27612	17.5	1	2
16	Simferopol	33990	21.0	–	2	46	Vologda	27037	15.7	–	2
17	Sochi	37099	22.2	–	1	47	Kazan	27595	18.3	–	2
18	Makhachkala	37385	23.6	2	1	48	Yekaterinburg	28440	16.8	–	2
19	Brest	7110	15.6	3	2	49	Bergen	1317	14.2	–	3
20	Paris	7156	19.1	1	2	50	Trondheim	1271	13.5	–	2
21	Zurich	6660	17.1	–	1	51	Vaasa	2911	14.7	1	2
22	Würzburg	10655	17.9	1	1	52	Kajaani	2869	14.2	1	1
23	Vienna	11035	19.8	–	2	53	Arkhangelsk	2250	14.1	2	2
24	Debrecen	12882	20.3	1	2	54	Syktyvkar	23804	15.2	2	1
25	Chernivtsi	33658	18.8	1	2	55	Ivdel	23921	15.3	1	1
26	Zaporozhe	33805	21.4	1	2	56	Bodo	1152	12.0	1	1
27	Rostov on the Don	34731	22.3	–	3	57	Sodankyla	2836	12.8	1	–
28	Astrakhan	34880	24.1	1	1	58	Naryan-Mar	2320	10.7	1	1
29	Valentia	3953	14.6	1	2	59	Pechora	23418	13.3	–	–
30	London	3776	16.4	–	2	60	Vardö	1098	8.4	–	2

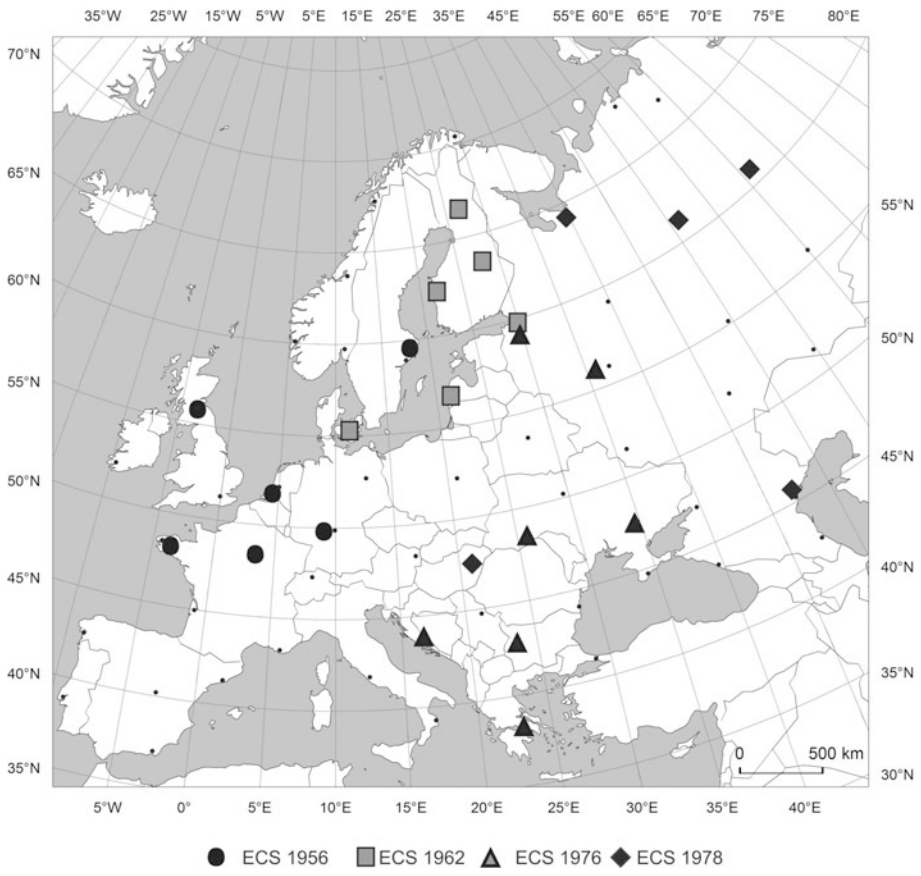


Fig. 1. Stations with ECS: 1956, 1962, 1976, and 1978.

Sofia (595 m), and Zurich (569 m). The stations were coded from 1 to 60 going from west to east in belts spanning 5 degrees of latitude, *i.e.*, up to 40°N – 4 stations (nos. 1 to 4), 40-45°N – 14 stations (5-18), 45-50°N – 10 stations (19-28), 50-55°N – 10 stations (29-38), 55-60°N – 10 stations (39-48), 60-65°N – 7 stations (49-55), 65-70°N – 4 stations (56-59), and above 70°N – 1 station (no. 60) (Table 1, Fig. 1).

Summers were classified as exceptionally cool or hot when the average seasonal temperature at a given station differed from its respective 60-year counterpart (1951-2010) by two or more standard deviations (*i.e.*, EHS: $t \geq t_{av} + 2\sigma$; ECS: $t \leq t_{av} - 2\sigma$). The same approach was followed when identifying exceptional months. This definition is analogous to the one used when identifying exceptionally hot months and seasons in Central and Eastern Europe (Kossowska-Cezak and Twardosz 2012a, b, 2013, Twardosz and

Kossowska-Cezak 2013a, b, c). By referencing average thermal conditions in defining thermally anomalous periods, the study managed to take into account both the concept of a thermal anomaly in the form established in climatology and the actual thermal conditions to which the populations of the various European climatic regions were acclimatised. The method also offered the benefit of obtaining information on very important characteristics of the climate, such as the scale of the anomaly and its frequency of occurrence. For this latter reason the authors decided, right from the start, not to employ the increasingly popular (and strongly popularised) percentile method which, by definition, did not offer these characteristics.

The numbers of days with the maximum temperature above 30°C (hot days) and with the minimum temperature above 20°C (tropical nights) were used to characterise the EHS and ECS identified.

The method yielded a number of exceptionally hot and cool summers, as well as calendars of their occurrence and the areas where they occurred.

3. FREQUENCY OF EHS AND ECS

Using the criterion of average temperature plus/minus two standard deviations, 42 ECS were identified in 13 years of the study period. ECS only occurred at 32 of the 60 stations (mostly in northern Europe), mostly only once (23 stations) with just Brest, on the western tip of the continent, recording three occurrences (Table 2).

Table 2

The number of EHS and ECS and the number of their occurrences

10-years	ECS		EHS	
	Number of years	Number of occurrences	Number of years	Number of occurrences
1951-1960	2	7	2	4
1961-1970	2	9	–	–
1971-1980	5	20	3	12
1981-1990	2	3	3	7
1991-2000	2	3	7	14
2001-2010	–	–	9	63
1951-2010	**13**	**42**	**24**	**100**

The largest number of five ECS was recorded in the decade 1971-1980 (Table 2), and during the period 1969-1978 there were six examples (Table 3). During the decade 1971-1980, 20 stations recorded an ECS (Table 2) and in the decade 1969-1978 there were 23 stations with an ECS, which ac-

Table 3

Exceptionally cool summers (ECS) (1951-2010)

Year	Number of stations	Stations (no. according to Table 1)
1954	1	19
1956	6	19, 20, 22, 31, 39, (42)
1962	6	(41), (43), (44), (51), (52), (57)
1969	3	53, 54, 58
1972	3	1, 19, 29
1975	1	56
1976	7	4, 11, 13, 25, 26, (44), (45)
1977	4	1, 2, (5), 6
1978	5	(24), 28, (53), (54), (55)
1984	1	4
1987	2	41, (43)
1993	1	18
1994	2	18, 38

Note: station codes in parentheses denote stations without an ECM during an ECS.

counts for more than half of all station/ECS. This means that ECS not only occurred the most frequently during that period, but also that they tended to cover wider areas. There were no ECS in the last decade in the study, but the decline in their occurrence had begun much earlier: after 1978 (4 stations reporting an ECS) exceptionally cool summers would only occur at 1-2 stations and not a single occurrence has been recorded since 1995. The greatest area with an ECS was recorded in 1976 when it covered 7 stations in South-Eastern Europe and in the central part of European Russia, and that was followed by 2 years (1956 and 1962) with 6 stations each.

EHS were far more frequent and were recorded at 100 cases in 24 years of the study period (Table 2). Only stations at Edinburgh, Sodankyla, and Pechora did not record any EHS, while most stations recorded two (29 stations) or three EHS (21 stations), with Oslo standing out with four occurrences. The greatest frequency of EHS was noted during the last decade (2001-2010) with nine such seasons (including every year from 2002) recorded at 63 stations (more than 63% of all stations) (Table 2). This means that the EHS of that period were not just very frequent, but also covered extended areas. The preceding decade (1991-2000) had seven EHS, but with just 14 stations involved the seasons clearly had a smaller spatial extent. Exceptionally hot summers were not recorded at all during the decade 1961-

1970 and in fact even during 1960-1971 (Table 4). While the period of most frequent occurrence of ECS (1971-1980) did not overlap with the period of lowest frequency (in fact lack of) EHS (1961-1970), the reverse correlation, *i.e.*, lack of ECS and the most frequent EHS, overlapped fully towards the latter part of the study period, *i.e.*, beginning virtually in 1995. The largest EHS included 20 stations in Eastern Europe (without its northern tip) in

Table 4

Exceptionally hot summers (EHS) (1951-2010)

Year	Number of stations	Stations (no. according to Table 1)
1955	3	29, 30, 40
1959	1	30
1972	9	27, 36, 37, 44, 45, 46, 52, 53, 60
1974	1	(53)
1976	2	19, 20
1981	3	(47), 48, 55
1988	2	48, 54
1989	2	1, 2
1992	4	23, 24, 32, 33
1993	1	58
1995	1	**29**
1997	4	40, 49, 50, 51
1998	2	**3**, 38
1999	1	34
2000	1	12
2002	8	34, 40, 42, 43, 49, 50, 51, 56
2003	17	3, 4, 6, 7, **8**, **9**, **10**, **11**, 12, 19, 20, **21**, **22**, **23**, 31, 32, 49
2004	1	60
2005	1	5
2006	5	5, 8, 33, 40, 41
2007	9	12, 13, 14, 15, 16, 24, (25), 26, (27)
2008	1	13
2009	1	10
2010	20	14, 15, 16, **17**, 18, 25, 26, 27, **28**, (33), 34, **35**, **36**, 37, **38**, 43, 44, **45**, 46, **47**

Notes: station codes in parentheses denote stations without an EHM during an EHS, station numbers printed in bold mean that the average temperature met the criterion $t \geq t_{av} + 3\sigma$.

2010; 16 stations in Western and Central Europe in 2003; 9 stations in 1972 (Northern and Eastern Europe), 2002 (Scandinavian Peninsula and Central Europe), and 2007 (South-Western Europe).

There were 3 years during the study period when ECS and EHS coincided. In 1972, there was an ECS in the far west of the continent (3 stations) and an EHS in Northern and Eastern Europe (9 stations). In 1976, the largest ECS of the period (7 stations) occurred in Central Europe, while an EHS occurred in Western Europe (2 stations). In 1993, an ECS over the Caspian Sea (1 station) coincided with an EHS in the far north-east (1 station).

In summary, the greater frequency of exceptionally hot than exceptionally cool summers is an obvious consequence of the properties of the summertime statistical distribution of temperature. Indeed, many authors (*e.g.*, Piervitali *et al.* 1997, Luterbacher *et al.* 2004, Elizbarashvili *et al.* 2007, Kossowska-Cezak and Twardosz 2013) agree that this distribution is characterised by a right-skewed asymmetry, *i.e.*, sporadic occurrence of temperatures considerably greater than the long-term average compensated by a greater frequency of lower discrepancies below the average.

These properties of the long-term average summer temperatures are linked with factors influencing temperature in this geographic zone. At a time of persistent high pressure systems that tend to prevent increased cloudiness, thus ensuring high insolation on long summer days, advection of hot tropical or polar continental air may lead to very high and persistent increases of air temperature. On the other hand, an influx of much cooler arctic or polar maritime air, even with increased cloudiness, does not produce as much of a negative thermal anomaly because it is offset by the solar conditions characteristic of summer. Even in subpolar areas the air remains warm due to the long day and to the influence of the Arctic Ocean.

4. THERMAL CHARACTERISTICS OF EHS AND ECS

The full calendar of years with EHS and ECS, including the number and code of the stations, is provided in Tables 3 and 4. As has already been noted, ECS occurred in 13 years of the study period, including 4 years at 1 station and 2 years at 2 stations (Table 3). These 6 years in which ECS were recorded at less than 3 stations (*i.e.*, less than 5% of stations) were omitted in the data on thermal characteristics given in this section. This leaves 7 years when ECS covered areas with between 3 and 7 stations.

EHS were much more frequent than ECS occurring in 24 years during the study period. In a similar manner to ECS, exceptionally hot summers were also frequently recorded at just one (10 years) or two (4 years), typically neighbouring, stations in a given year (Table 4). These occurrences were also omitted in this section. The EHS in the remaining 10 years covered areas with between 3 and 20 stations.

The characteristics of individual EHS and ECS include not just their location and approximate extent indicated by the recording stations, but also the scale of the temperature anomaly (Δt) measured as the difference between the average temperature of the summer in question and the long-term average. Additionally, the authors noted the number of exceptionally hot or cool months, if any, during the EHS or ECS ($t \le t_{av} - 2\sigma$ or $t \ge t_{av} + 2\sigma$).

4.1 Exceptionally cool summers (ECS)

The **ECS of 1956** covered part of north-western Europe (5 stations) plus Stockholm (Fig. 1). August was an exceptionally cool month. The temperature anomaly (Δt) ranged from –1.3 to 1.7°C in the west and –2.6°C in the east of the area (Table 5).

The **ECS of 1962** encompassed eastern Scandinavia, the eastern Baltic coast (5 stations altogether) and Copenhagen (Fig. 1). It included no exceptionally cool months. The temperature anomaly (Δt) ranged from –2.0°C in Copenhagen to –2.8°C in Sondankyla (Table 5).

The **ECS of 1969** was noted at 3 stations in north-eastern Europe (Arkhangelsk, Syktyvkar, and Naryan-Mar). June was an exceptionally cool month. The temperature anomaly ranged from –2.6°C in the west to –3.7°C in the north-east.

The **ECS of 1972** covered an area with 3 stations on the Atlantic coast: Lisbon ($\Delta t = -1.8$°C), Brest ($\Delta t = -1.7$°C), and Valentia ($\Delta t = -1.3$°C). The anomaly at Valentia may be smallest in absolute terms, but in terms of standard deviation at it stood out at $\Delta t = t \le t_{av} - 3\sigma$. June was exceptionally cool. This ECS coincided with an EHS in Central and Eastern Europe.

Table 5

Thermal characteristic of the exceptionally cool summers (ECS)

Year	Station		Temperature					Number of days with temperature	
			T_{av}	Δt		T_{max}	T_{min}		
	No.	Name	°C	°C	σ	°C	°C	$T_{max} > 30$°C	$T_{min} > 20$°C
1956	20	Paris	16.5	–2.6	2.2	20.8	12.2	3	–
	22	Würzburg	15.7	–2.2	2.0	20.4	11.5	–	–
1962	41	Copenhagen	15.0	–2.0	2.0	18.6	11.2	–	–
	44	St. Petersburg	14.2	–2.6	2.2	18.1	10.7	–	–
1976	4	Athens	25.4	–1.9	2.3	30.9	20.8	53	59
	44	St. Petersburg	14.3	–2.5	2.1	17.9	11.1	–	1
	45	Moscow	14.8	–2.7	2.1	19.4	10.6	–	–
1978	28	Astrakhan	21.9	–2.2	2.2	28.1	16.6	29	11
	53	Arkhangelsk	11.6	–2.5	2.0	16.8	7.4	1	–

The **ECS of 1976** covered the largest area in the study period and included 7 stations from Athens in the south across the Balkan Peninsula to southern Ukraine, Moscow, and St. Petersburg (Fig. 1). August was exceptionally cool over the whole area, while Sofia also had an exceptionally cool June and Zaporozhe also had an exceptionally cool July. The temperature anomaly ranged from –2.0 to –3.0°C (Table 5). In the same year, an EHS was recorded in Brest and Paris.

The **ECS of 1977** was recorded by stations on the Iberian Peninsula (4 stations). All stations recorded at least one exceptionally cool month, but in Almeria it was all three (June-August), and in Madrid, June, and August. The temperature anomaly reached between –1.6 and –1.9°C, but in Madrid up to –3.5°C.

A total of 5 stations in three different areas recorded **the ECS of 1978**: in the north-east (Arkhangelsk, Syktyvkar, Ivdel; there were no exceptionally cool months, Δt ranged from –2.5 to –3.0°C); in Astrakhan (exceptionally cool June, $\Delta t = -2.5$°C); and at Debrecen (no exceptionally cool months, $\Delta t = -2.1$°C) (Fig. 1).

A review of ECS covering at least 5% of the weather stations showed that while some ECS occurred in a single area, others affected several discrete ones. Perhaps the reason for this uneven spatial distribution of the thermal anomalies is linked to the atmospheric circulation. Understanding of this influence would require access to regional catalogues of circulation types, which are not always available. The scale of thermal anomalies identified was not great and mostly did not exceed –3.0°C with a maximum of –3.5°C (Madrid in 1977).

4.2 Exceptionally hot summers (EHS)

The **EHS of 1955** occurred at 3 stations in Western Europe (Table 4), including Valentia, London, and Oslo. August was exceptionally hot at all 3 stations, as also was July in Oslo. The temperature anomaly at Valentia stood at $\Delta t = 1.4$°C and elsewhere at 2.5-2.3°C.

The **EHS of 1972** covered large areas of Eastern Europe (9 stations) from Rostov to Kursk and Saratov to Moscow and St. Petersburg all the way to the northern tip of Scandinavia (Fig. 2). Exceptionally hot months included August (at most stations), June (Rostov), and July (Kajaani and St. Petersburg – together with August). The temperature anomaly ranged from 2.4-2.7°C in the north to 3.3-3.7°C in the centre and peaked at Saratov (Table 6).

The **EHS of 1981** covered 3 stations in Eastern Europe: Kazan, Yekaterinburg, and Ivdel. The latter two recorded an exceptionally hot August. The temperature anomaly Δt ranged from 2.7 to 3.2°C.

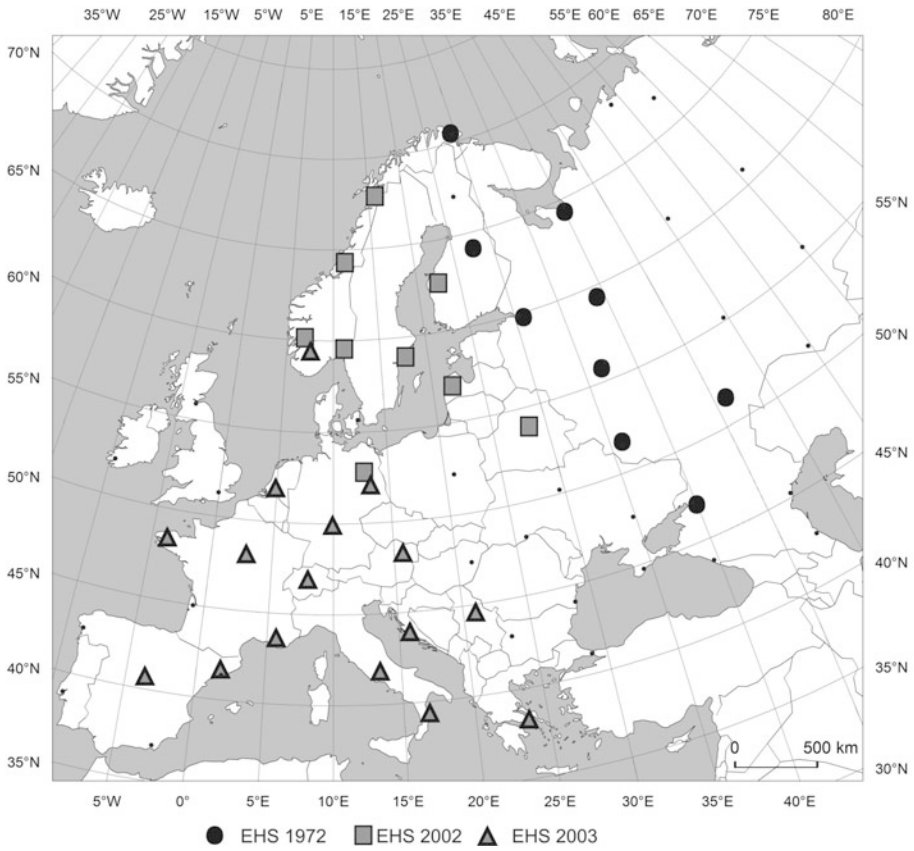


Fig. 2. Stations with EHS: 1972, 2002, and 2003.

The **EHS of 1992** included 4 stations in Central Europe: Berlin, Warsaw, Vienna, and Debrecen. August was exceptionally hot (including $t \geq t_{av} + 3\sigma$ at Vienna and Debrecen). The temperature anomaly Δt ranged from 2.2 to 2.4°C. While high, it did not constitute the highest summer temperature anomaly of the 20th century in Central Europe. Indeed, Kossowska-Cezak (1993) and Filipiuk and Kaszewski (2000) found that the greatest anomalies were recorded in the summer of 1939. Exceptionally high temperatures in the north of Central Europe were caused by the southern advection of tropical air, which was augmented by a foehn effect across the mountain barrier of the Sudeten and Carpathian Mts. (Kuziemska 1987, Kossowska-Cezak 1993). This type of weather is linked to a wedge of the Azores' high (Fink *et al.* 2004).

The **EHS of 1997** also involved 4 stations, but this time in central Scandinavia, including: Oslo, Bergen (exceptionally hot August), Trondheim, and

Vaasa (exceptionally hot July and August). The temperature anomaly ranged from 2.5 to 2.8°C, but peaked in Oslo at 3.9°C.

Table 6

Thermal characteristic of the exceptionally hot summers (EHM) temperature

Year	Station		Air temperature					Number of days with temperature	
			T_{av}	Δt		T_{max}	T_{min}		
	No.	Name	°C	°C	σ	°C	°C	$T_{max} > 30$°C	$T_{min} > 20$°C
1972	27	Rostov on the Don	25.2	+2.9	2.2	31.8	19.3	67	37
	36	Kursk	21.6	+3.2	2.0	27.2	16.2	21	6
	44	St. Petersburg	20.1	+3.3	2.8	25.0	15.8	16	14
	45	Moscow	20.7	+3.2	2.4	26.4	14.7	24	6
	53	Arkhangelsk	16.8	+2.7	2.1	22.8	11.1	11	6
2002	42	Stockholm	19.3	+2.8	2.2	24.2	15.4	2	1
	49	Bergen	16.6	+2.4	2.4	20.6	13.6	–	1
	56	Bodo	14.7	+2.7	2.4	18.1	11.6	–	–
2003	4	Athens	29.0	+1.7	2.0	33.1	22.9	86	83
	8	Barcelona	**26.6**	+4.2	3.5	31.2	22.1	71	75
	10	Rome	**26.9**	+3.5	3.7	33.0	20.9	76	62
	20	Paris	22.6	+3.5	3.0	27.6	17.6	20	16
	21	Zurich	**21.7**	+4.5	4.2	27.8	16.3	25	2
	32	Berlin	20.0	+2.3	2.3	25.9	14.1	18	–
	49	Bergen	16.2	+2.0	2.0	20.4	13.0	2	–
2006	8	Barcelona	24.8	+2.4	2.0	28.5	21.0	32	65
	41	Copenhagen	18.9	+1.9	2.0	23.4	14.8	2	2
2007	15	Istanbul	25.8	+2.7	2.8	30.7	21.6	53	66
	27	Rostov on the Don	25.4	+3.1	2.4	32.0	19.0	63	35
2010	15	Istanbul	25.4	+2.3	2.3	29.7	21.5	51	62
	18	Makhachkala	26.0	+2.5	2.8	30.2	21.7	49	67
	27	Rostov on the Don	26.1	+3.8	2.9	32.5	20	64	46
	28	Astrakhan	**27.5**	+3.4	3.4	34.9	20.4	84	55
	36	Kursk	**24.2**	+5.8	3.9	30.3	18.6	52	37
	44	St. Petersburg	19.8	+3.0	2.5	24.1	15.8	20	21
	45	Moscow	**22.2**	+4.7	3.6	27.8	16.4	44	25

Note: a value in bold means that the temperature meets the criterion $t \geq t_{av} + 3\sigma$.

The **EHS of 2002** covered north-central Europe from Minsk to the Baltic coast and a majority of the Scandinavian Peninsula to the town of Bodo (Fig. 2). August was exceptionally hot everywhere (including $t > t_{av} + 3\sigma$ at Bergen), except in Bodo when the hot month was June. The scale of the anomaly was distributed in a rather irregular way and ranged from 2.1 to 3.3°C.

The **EHS of 2003** was the second largest in terms of territory; it included 16 stations in Western and Southern Europe (Fig. 2). Its intensity was exceptional too, as the anomaly exceeded three standard deviations ($t > t_{av} + 3\sigma$) in Barcelona, Marseilles, Rome, Split, Würzburg, and Vienna, and four standard deviations in Zurich (Table 6). At most of the stations June and/or August were exceptionally hot and at many others the scale of the anomaly exceeded 3σ. In August, the greatest temperature anomaly was recorded in the south-western part of the Massif Central, France, at Gourdon (Friedrich and Bissolli 2012). In Barcelona and Rome all three summer months were exceptionally hot. The scale of the anomaly ranged from approximately 2°C on the Atlantic coast, to 2-3°C in Western Europe, to more than 3°C in all other locations and it peaked in Zurich at $\Delta t = 4.6$°C. In that part of Europe, this was the hottest summer of the study period and at many of the stations also the hottest in the history of measurement (Luterbacher *et al.* 2004). During that EHS, Western Europe experienced exceptionally hot weather, which caused an unprecedented increase in mortality rates and rapid thawing of Alpine glaciers. Indeed, the scale and intensity of this heat wave was researched in multiple studies (*e.g.*, BUWAL 2004, Fink *et al.* 2004, Schär *et al.* 2004, Stott *et al.* 2004, Chase *et al.* 2006, D'Ippoliti *et al.* 2010, Twardosz 2009, Twardosz and Batko 2012). These weather conditions were caused by a highly stable blocking situation over Europe (Fink *et al.* 2004, Twardosz 2009). A strong high-pressure system centred over Scandinavia with a wedge penetrating into Central Europe dominated the weather, forcing a southern circulation with hot tropical air to stream into Western Europe.

The **EHS of 2006** covered 5 stations split between two different areas: northern Spain (La Coruña and Barcelona) and a more central area with Oslo, Copenhagen and Warsaw (Fig. 3). July was exceptionally hot in both areas and their temperature anomalies were similar $\Delta t = 1.8$-2.4°C (Table 6).

The **EHS of 2007** covered an area that extended from Debrecen and Belgrade in the west to Rostov on the Don in the east and Istanbul in the south (9 stations) (Fig. 3). The most intensive heat of that summer was recorded in the Balkan Peninsula (Founda and Giannakopoulos 2009). June was exceptionally hot at almost all of the stations, but so was either July or August at some of them. The temperature anomaly ranged from 2.2 to 3.2°C (Table 6)

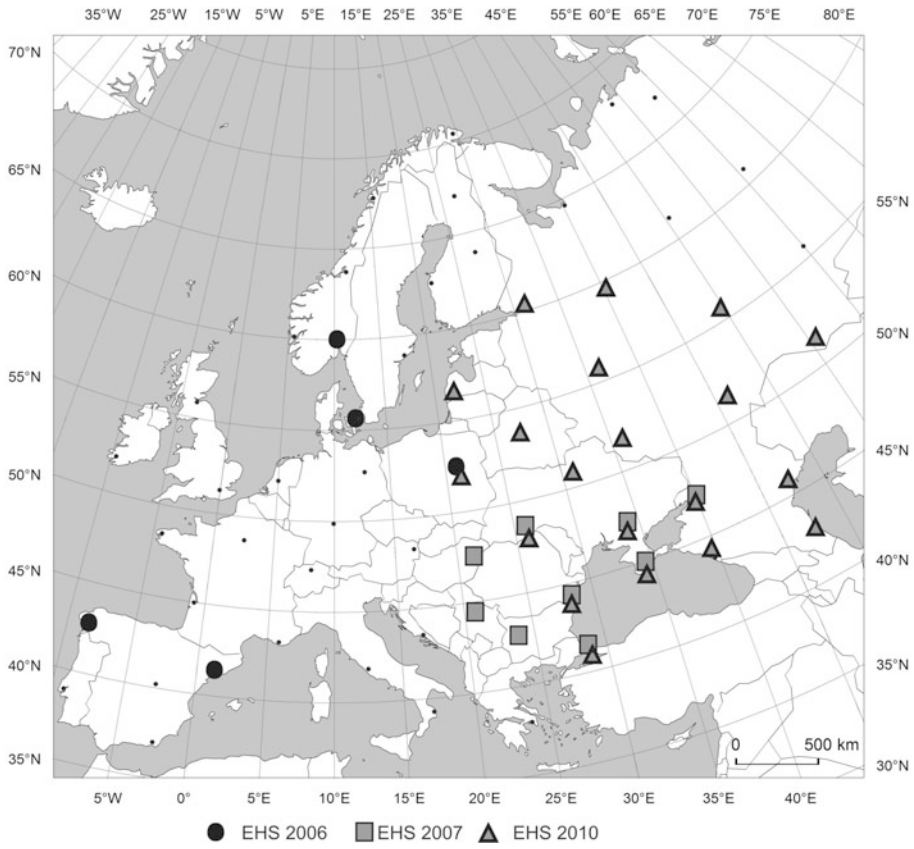


Fig. 3. Stations with EHS: 2006, 2007, and 2010.

without any apparent regularity in their spatial distribution. The highest anomaly was recorded in Simferopol at $\Delta t = 3.2$°C.

The **EHS of 2010** was the largest in terms of territory, as it involved 20 stations located east of Liepaja and Warsaw and extending all the way to Kazan and Orenburg, and from Istanbul and Makhachkala in the south to St. Petersburg and Vologda in the north (Fig. 3), in other words: Central Europe and European Russia bar its northernmost part. At seven of the 20 stations (in Russia and the Ukraine) the average temperature of the summer exceeded the long-term averages by three or more standard deviations ($t \geq t_{av} + 3\sigma$). In various parts of this extensive area, different months were exceptionally hot and at most of the stations there were two such months (including August almost everywhere), but at Sochi, Astrakhan, Kiev, Kursk, Saratov, Orenburg, Moscow, and Kazan all three months had exceptional temperatures. The temperature anomaly increased from 2.0°C at the western limit of the

area towards the centre of European Russia to reach 5.0°C at Kazan, 5.3°C at Saratov, and 5.8°C at Kursk (Table 6). The EHS of 2010 was therefore not just the largest, but also the most intense. The persistence of a high degree of heat stress caused adverse conditions for human health and negative economic effects, including widespread fires, which, in turn, caused a 2-3 times increase in air pollution over Moscow (Zvyagintsev *et al.* 2011). Just as in other similar cases in Europe, this exceptionally hot summer had been caused by long-lasting anticyclones (Bardin 2007), the advection of tropical air and long sunshine hours (WMO 2010). A powerful high-pressure bank over Russia created a stable blocking situation that forced southern circulation (Masters 2010, Gruza and Ran'kova 2011).

5. EXCEPTIONALLY HOT/COOL MONTHS (EHM AND ECM)

For an exceptionally hot or cool summer to occur there must be a temperature anomaly that is sufficient in scale and duration. For this reason, EHS and (especially) ECS are neither very frequent, nor do they involve large areas in Europe. This is also the reason why there are far more anomalous months (ECM and EHM) than entire summers.

The study has found that there were 119 occurrences of an ECM being recorded during 49 months of the study period. Only 32 of these occurrences (27%) took place during an ECS, while the remaining 87 were isolated months. June was exceptionally cool 18 times and at 56 stations, in July on 19 times and in 29 stations and August on 12 times and 34 stations. This means that while June and July were equally likely to be exceptionally cool, the June ECM typically covered a far larger area than in July or August.

ECM were most frequently recorded at either one or 2 stations (34 of 49 months, *i.e.*, nearly 70%) and the two largest ECM covered areas represented by 8 stations (August 1956 coinciding with an ECS at 5 stations in Central Europe) and 7 stations (ECM of June in 1982 in north-western Russia and northern Scandinavia).

ECM were the most frequent in the decade 1971-1980, which included 15 of the 49 ECM and 39 out of 119 occurrences. On both counts, ECM were more numerous than EHM during that decade, even though EHM were much more frequent overall. The lowest number of ECM (7) was recorded in the last of the decades studied, *i.e.*, 2001-2010.

Indeed, EHM were nearly three times as numerous as ECM at 368 occurrences in 83 months of the study period. 144 of these (39%) were recorded during exceptionally hot summers, which is a much greater percentage than in exceptionally cool summers.

This means that thermally anomalous summers without thermally anomalous months were much more frequent on the cool side than on the hot side

of the year. Indeed, it was exceedingly rare for two ECM to occur in an ECS (5 occurrences) and there was only one ECS (Almeria in 1977) with all three months exceptionally cool. The picture is the reverse with exceptionally hot summers, which mostly included at least one EHM, quite often two and in the 21st century even all three. Two such seasons with all EHM included the Western-European EHS of 2003 at 2 stations (Barcelona and Rome) and the Eastern-European EHS of 2010 at 5 stations (from Kiev to Astrakhan and Orenburg).

June and July were EHM 26 times each and August 31 times, while the station count increased from June to August with 103, 121, and 144, respectively. The fact that June had 47% of the ECM occurrences and August had 39% of the EHM suggests that the gradual warming of the air and ground by solar radiation has an influence on thermal anomalies, even if their fundamental causes are linked with specific types of atmospheric circulation (Kossowska-Cezak 1997).

Just like ECM, EHM also tended to cover areas with one or 2 stations, but significantly less frequently. Only 36 of the 83 (43%) EHM fell into that category, while the remaining EHS covered much greater territories than ECM. The three largest EHM encompassed 21 stations (August 2003), 20 stations (June 2010), and 17 stations (July 2006 and August 2010). As a comparison, only two ECM covered more than 6 stations (4%), while 18 EHM (22%) fitted in that category.

The frequency of EHM over the study period (1951-2010) rose from 38 EHM and 116 occurrences in 1951-1990, to 19 and 69 occurrences in 1991-2000, and to 26 and 183 occurrences in 2001-2010. This means that the last decade of the study accounts for one half of all EHM occurrences and the last two decades account for 69% of such occurrences. It must be added that during the last 20 years, and more precisely from 1987, an EHM was recorded somewhere in Europe every year. The rapid increase in EHM occurrence in the 21st century is linked with the fact that they are not only more frequent, but also cover ever greater territories. Another phenomenon of the 21st century worthy of note is the unprecedented scale of thermal anomalies during EHM and EHS, which exceed three or even four standard deviations.

There is also an interesting category of summers with a “both-way anomaly”, when an EHM in one part of Europe coincides with an ECM in another. There were 16 such months when some of the 49 ECM and 83 EHM overlapped in a single month. Examples include June and August 1972: a June ECM was recorded at 6 stations in the far west of Europe, while an EHM occurred at the 2 discrete stations of Bodo and Rostov; in August there was an ECM in Barcelona, while 11 Russian stations recorded an EHM. In 1991, June was an ECM at 5 stations on the North Sea coast and in southern Scandinavia, but it was also an EHM at 4 stations in the far east

of Europe. 2003 was a reverse of that situation, as Western Europe experienced its hottest and largest June EHM (11 stations), while at the same time 4 Russian stations recorded an ECM.

6. CONCLUSIONS

The study used a relatively precise statistical criterion ($t \leq t_{av} - 2\sigma$ or $t \geq t_{av} + 2\sigma$) to identify exceptionally hot and cool summers (EHS and ECS) during the 60-year study period from 1951 to 2010 at 60 European weather stations. Alongside the entire summer seasons the study also looked at exceptional months.

It was found that EHS were approximately twice as frequent as ECS and occurred virtually in the whole study area, while the ECS were concentrated in its northern part. ECS were observed in 13 years, while EHS in 24 years, but a half of both types of exceptional summer were only recorded at 1 or 2 stations (6 and 14, respectively). Such thermally anomalous summers rarely covered large areas, as only three ECS (1956, 1962, and 1976) and five EHS (1972, 2002, 2003, 2007, and 2010) were recorded by at least 10% of the stations.

EHS were not just more frequent than ECS and likely to occur anywhere in Europe, but also their extreme occurrences covered much larger territories with up to one quarter or even one third of all stations included in the study. What is more, the extreme occurrences of EHS featured higher temperature anomalies than the respective ECS (in excess of 5.0°C compared to less than 4.0°C in the case of the most extreme ECS).

The greater frequency, greater territorial coverage and the extreme scale of the thermal anomalies of both EHM and EHS, when compared to ECM and ECS, suggests a potential for prolonged periods of radically increased air temperatures in Europe during the summer season, *i.e.*, at a time of maximum sunshine duration. This increase is also linked with the advection of hot tropical or continental air. On the other hand, an influx of much cooler arctic or polar maritime air is sufficiently offset by the sunshine potential on long summer days to prevent an equivalent drop of air temperature.

The frequencies of EHS and ECS varied over the study period. The most distinctive feature of this variability was the gradual disappearance of ECS and the simultaneous increase in the frequency of EHS beginning in the mid-1990s. The latter not only grew in frequency, but also in size and severity, as the most extensive and the hottest EHS occurred in the 21st century, namely in 2003 in Western Europe and in 2010 in Eastern Europe.

The results of this study are of a preliminary nature and the authors intend to embark on a comprehensive study involving all four seasons with

anomalous temperatures in Europe using data from a much greater pool of stations.

Acknowledgements. We thank Mr. Paweł Pilch and Dr. Martin Cahn for reviewing the English.

References

Anisimov, O.A., I.I. Borzenkova, E.L. Zhil'tsova, O.K. Zakharova, V.A. Kokorev, S.A. Reneva, and Yu.G. Strel'chenko (2011), Hydrometeorological conditions of the Volga region and current climate changes, *Rus. Meteorol. Hydrol.* **36**, 5, 307-314, DOI: 10.3103/S1068373911050049.

BUWAL (2004), Auswirkungen des Hitzesommers 2003 auf die Gewässer. Dokumentation, Schriftenreiche Umwelt, Nr. 369, Gewässerschutz, Bundesamt für Umwelt, Wald und Landschaft, Bern (in German).

Bardin, M.Yu. (2007), Anticyclonic quasi-stationary circulation and its effect on air temperature anomalies and extremes over western Russia, *Rus. Meteorol. Hydrol.* **32**, 2, 75-84, DOI: 10.3103/S106837390702001X.

Bardin, M.Yu. (2011), Scenary forecasts of air temperature variations for the regions of the Russian Federation up to 2030 using the empirical stochastic climate models, *Rus. Meteorol. Hydrol.* **36**, 4, 217-228, DOI: 10.3103/S1068373911040017.

Barriopedro, D., E.M. Fischer, J. Luterbacher, R.M. Trigo, R. García-Herrera (2011), The hot summer of 2010: Redrawing the temperature record map of Europe, *Science* **332**, 6026, 220-224, DOI: 10.1126/science.1201224.

Black, E., and R. Sutton (2007), The influence of oceanic conditions on the hot European summer of 2003, *Clim. Dyn.* **28**, 1, 53-66, DOI: 10.1007/s00382-006-0179-8.

Błażejczyk, K., and G. McGregor (2007), Bio-thermal conditions and mortality in selected European agglomerations, *Prz. Geogr*. **79**, 3-4, 627-649 (in Polish).

Chase, T.N., K. Wolter, R.A. Pielke Sr., and I. Rasool (2006), Was the 2003 European summer heat wave unusual in a global context?, *Geophys. Res. Lett*. **33**, 23, L23709, DOI: 10.1029/2006GL027470.

Coumou, D., A. Robinson, and S. Rahmstorf (2013), Global increase in record-breaking monthly-mean temperatures, *Climatic Change* **118**, 3-4, 771-782, DOI: 10.1007/s10584-012-0668-1.

de Freitas, C.R. (2003), Tourism climatology: evaluating environmental information for decision making and business planning in the recreation and tourism sector, *Int. J. Biometeorol.* **48**, 1, 45-54, DOI: 10.1007/s00484-003-0177-z.

Della-Marta, P.M., J. Luterbacher, H. von Weissenfluh, E. Xoplaki, M. Brunet, and H. Wanner (2007), Summer heat waves over western Europe 1880-2003,

their relationship to large-scale forcings and predictability, *Clim. Dyn.* **29**, 2-3, 251-275, DOI: 10.1007/s00382-007-0233-1.

D'Ippoliti, D., P. Michelozzi, C. Marino, F. de'Donato, B. Menne, K. Katsouyanni, U. Kirchmayer, A. Analitis, M. Medina-Ramón, A. Paldy, R. Atkinson, S. Kovats, L. Bisanti, A. Schneider, A. Lefranc, C. Iñiguez, and C.A. Perucci (2010), The impact of heat waves on mortality in 9 European cities: results from the EuroHEAT project, *Environ. Health* **9**, 37, 1-9, DOI: 10.1186/1476-069X-9-37.

Dole, R., M. Hoerling, J. Perlwitz, J. Eischeid, P. Pegion, T. Zhang, X.W. Quan, T. Xu, and D. Murray (2011), Was there a basis for anticipating the 2010 Russian heat wave?, *Geophys. Res. Lett.* **38**, 6, L06702, DOI: 10.1029/2010GL046582.

Donat, M.G., L.V. Alexander, H. Yang, I. Durre, R. Vose, R.J.H. Dunn, K.M. Willett, E. Aguilar, M. Brunet, J. Caesar, B. Hewitson, C. Jack, A.M.G. Klein Tank, A.C. Kruger, J. Marengo, T.C. Peterson, M. Renom, C. Oria Rojas, M. Rusticucci, J. Salinger, A.S. Elrayah, S.S. Sekele, A.K. Srivastava, B. Trewin, C. Villarroel, L.A. Vincent, P. Zhai, X. Zhang, and S. Kitching (2013), Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset, *J. Geophys. Res.* **118**, 5, 2098-2118, DOI: 10.1002/jgrd.50150.

Elizbarashvili, E.Sh., R.Sh. Meskhiya, and M.E. Elizbarashvili (2007), Dynamics of occurrence frequency of extreme anomalies of monthly mean air temperature in Georgia in the 20th century and its effect on precipitation and on the river water discharge, *Rus. Meteorol. Hydrol.* **32**, 1, 71-74, DOI: 10.3103/S1068373907010116.

Feudale, L., and J. Shukla (2011), Influence of sea surface temperature on the European heat wave of 2003 summer. Part I: an observational study, *Clim. Dyn.* **36**, 9-10, 1691-1703, DOI: 10.1007/s00382-010-0788-0.

Filipiuk, E., and B.M. Kaszewski (2000), Hot and cold summers in Central Europe (1871-1990), *Prace Geogr.* **108**, 149-154.

Fink, A.H., T. Brücher, A. Krüger, G.C. Leckebusch, J.G. Pinto, and U. Ulbrich (2004), The 2003 European summer heatwaves and drought – synoptic diagnosis and impacts, *Weather* **59**, 8, 209-216, DOI: 10.1256/wea.73.04.

Fischer, E.M., and C. Schär (2010), Consistent geographical patterns of changes in high-impact European heatwaves, *Nat. Geosci.* **3**, 398-403, DOI: 10.1038/ngeo866.

Founda, D., and C. Giannakopoulos (2009), The exceptionally hot summer of 2007 in Athens, Greece – A typical summer in the future climate?, *Global Planet. Change* **67**, 3-4, 227-236, DOI: 10.1016/j.gloplacha.2009.03.013.

Friedrich, K., and P. Bissolli (2012), Analysis of temperatures and precipitation recorded at stations in Eastern Europe during the heat wave in summer 2010, Deutscher Wetterdienst, Business Area Climate and Environment, http://www.dwd.de/bvbw/generator/DWDWWW/Content/Oeffentlichkeit/

KU/KU2/KU23/rcc-cm/products/BesondereWetterereignisse/European/20110124__Hitzewelle__Russland__de,templateId=raw,property=publicationFile.pdf/20110124_Hitzewelle_Russland_de.pdf.

Gerstengarbe, F.W., and P.C. Werner (1992), The time structure of extreme summers in Central Europe between 1901 and 1980, *Meteorol. Z.* **1**, 6, 285-289.

Gruza, G.V., and E.Ya. Ran'kova (2011), Estimation of probable contribution of global warming to the genesis of abnormally hot summers in the European part of Russia, *Izv. Atmos. Ocean. Phys.* **47**, 6, 661-664, DOI: 10.1134/S0001433811060065.

Hansen, J., M. Sato, and R. Ruedy (2012), Perception of climate change, *Proc. Natl. Acad. Sci. U.S.A*. **109**, 37, E2415-E2423, DOI: 10.1073/pnas.1205276109.

Hodzic, A., S. Madronich, B. Bohn, S. Massie, L. Menut, and C. Wiedinmyer (2007), Wildfire particulate matter in Europe during summer 2003: meso-scale modeling of smoke emissions, transport and radiative effects, *Atmos. Chem. Phys.* **7**, 15, 4043-4064.

Hutter, H.P., H. Moshammer, P. Wallner, B. Leitner, and M. Kundi (2007), Heatwaves in Vienna: effects on mortality, *Wien. Klin. Wochenschr.* **119**, 7-8, 223-227, DOI: 10.1007/s00508-006-0742-7.

IPCC (2013), Summary for policymakers. **In:** 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.), *Climate Change 2013: The Physical Science Basis. Working Group I. Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, 3-32.

Isayev, A.A., and B.G. Sherstyukov (2008), Mean and extreme characteristics of Moscow climate at the end of the 20th century, *Rus. 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.

Jones, G.S., P.A. Stott, and N. Christidis (2008), Human contribution to rapidly increasing frequency of very warm Northern Hemisphere summers, *J. Geophys. Res*. **113**, D2, D02109, DOI: 10.1029/2007JD008914.

Jones, P.D., D.H. Lister, T.J. Osborn, C. Harpham, M. Salmon, and C.P. Morice (2012), Hemispheric and large-scale land-surface air temperature variations: An extensive revision and an update to 2010, *J. Geophys. Res*. **117**, D5, D05127, DOI: 10.1029/2011JD017139.

Kamae, Y., H. Shiogama, M. Watanabe, and M. Kimoto (2014), Attributing the increase in Northern Hemisphere hot summers since the late 20th century, *Geophys. Res. Lett.* **41**, 14, 5192-5199, DOI: 10.1002/2014GL061062.

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 surface air temperature and precipitation series for the European Climate Assessment, *Int. J. Climatol*. **22**, 12, 1441-1453, DOI: 10.1002/joc.773.

Kossowska-Cezak, U. (1993), Summer 1992 in Poland against summer seasons of 120-years period, *Prz. Geofiz.* **38**, 1, 67-74 (in Polish).

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. (2010), Occurrence of hot weather in Warsaw (1951-2009), *Prz. Geofiz.* **2010**, 1-2, 61-75 (in Polish).

Kossowska-Cezak, U., and R. Twardosz (2012a), Exceptionally hot summer months and seasons in Central and Eastern Europe (1951-2010). Part I. Exceptionally hot summer months, *Prz. Geofiz.* **2012**, 3-4, 299-324 (in Polish).

Kossowska-Cezak, U., and R. Twardosz (2012b), Exceptionally hot summer months and seasons in Central and Eastern Europe (1951-2010). Part II. Exceptionally hot summer seasons, *Prz. Geofiz.* **2012**, 3-4, 325-342 (in Polish).

Kossowska-Cezak, U., and R. Twardosz (2013), Exceptionally cool summers in Central and Eastern Europe (1951-2010), *Prz. Geofiz.* **2013**, 1-2, 25-39 (in Polish).

Krzyżewska, A., and S. Wereski (2011), Heat waves and frost waves in selected Polish stations against bioclimatic regions background (2000-2010), *Prz. Geofiz.* **2011**, 1-2, 99-109 (in Polish).

Kürbis, K., M. Mudelsee, G. Tetzlaff, and R. Brázdil (2009), Trends in extremes of temperature, dew point, and precipitation from long instrumental series from central Europe, *Theor Appl. Climatol.* **98**, 1-2, 187-195, DOI: 10.1007/s00704-008-0094-5.

Kuziemska, D. (1975), Atmospheric precipitation in cool, temperate and warm months of the year in Warsaw, Temat 1.05.04, Problem 157, Institute of Meteorology and Water Management (IMGW), Warszawa, Poland, 1-28 (in Polish).

Kuziemska, D. (1987), The air temperature differentiation on Poland area and atmospheric circulation types over central Europe, *Prz. Geofiz.* **3**, 277-287 (in Polish).

Kyselý, J. (2010), Recent severe heat waves in central Europe: how to view them in a long-term prospect?, *Int. J. Climatol.* **30**, 1, 89-109, DOI: 10.1002/joc.1874.

Kyselý, J., and R. Huth (2008), Relationships of surface air temperature anomalies over Europe to persistence of atmospheric circulation patterns conducive to heat waves, *Adv. Geosci*. **14**, 243-249, DOI: 10.5194/adgeo-14-243-2008.

Luterbacher, J., D. Dietrich, E. Xoplaki, M. Grosjean, and H. Wanner (2004), European seasonal and annual temperature variability, trends, and extremes since 1500, *Science* **303**, 5663, 1499-1503, DOI: 10.1126/science.1093877.

Masters, J. (2010), Causes of the Russian heat wave and Pakistani floods, Dr. Jeff Master's WunderBlog, http://www.wunderground.com/blog/JeffMasters/comment.html?entrynum=1576 (retrieved on 9 February 2011).

Meehl, G.A., and C. Tebaldi (2004), More intense, more frequent, and longer lasting heat waves in the 21st century, *Science* **305**, 5686, 994-997, DOI: 10.1126/science.1098704.

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.

Muthers, S., A. Matzarakis, and E. Koch (2010), Climate change and mortality in Vienna – A human biometeorological analysis based on regional climate modeling, *Int. J. Environ. Res. Public Health* **7**, 7, 2965-2977, DOI: 10.3390/ijerph7072965.

Perkins, S.E., L.V. Alexander, and J.R. Nairn (2012), Increasing frequency, intensity and duration of observed global heatwaves and warm spells, *Geophys. Res. Lett*. **39**, 20, L20714, DOI: 10.1029/2012GL053361.

Piervitali, E., M. Conte, and M. Colacino (1997), Summer air temperature anomalies in Europe during the century 1811-1910, *Nuovo Cimento C* **20**, 2, 195-208.

Rebetez, M., O. Dupont, and M. Giroud (2009), An analysis of the July 2006 heatwave extent in Europe compared to the record year of 2003, *Theor. Appl. Climatol.* **95**, 1-2, 1-7, DOI: 10.1007/s00704-007-0370-9.

Révész, A. (2008), Stochastic behaviour of heat waves and temperature in Hungary, *Appl. Ecol. Environ. Res.* **6**, 4, 85-100, DOI: 10.15666/aeer/0604_085100.

Revich, B.A., and D.A. Shaposhnikov (2012), Climate change, heat waves, and cold spells as risk factors for increased mortality in some regions of Russia, *Stud. Rus. Econom. Develop.* **23**, 2, 195-207, DOI: 10.1134/S1075700712020116.

Schär, C., P.L. Vidale, D. Lüthi, C. Frei, C. Häberli, M.A. Liniger, and C. Appenzeller (2004), The role of increasing temperature variability in European summer heatwaves, *Nature* **427**, 6972, 332-336, DOI: 10.1038/nature02300.

Scott, D., and C. Lemieux (2010), Weather and climate information for tourism, *Procedia Environ. Sci.* **1**, 146-183, DOI: 10.1016/j.proenv.2010.09.011.

Sidorenkov, N.S., and I.A. Orlov (2008), Atmospheric circulation epochs and climate changes, *Rus. Meteorol. Hydrol.* **33**, 9, 553-559, DOI: 10.3103/S1068373908090021.

Sillmann, J., V.V. Kharin, X. Zhang, F.W. Zwiers, and D. Bronaugh (2013), Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate, *J. Geophys. Res.* **118**, 4, 1716-1733, DOI: 10.1002/jgrd.50203.

Sillmann, J., M.G. Donat, J.C. Fyfe, and F.W. Zwiers (2014), Observed and simulated temperature extremes during the recent warming hiatus, *Environ. Res. Lett*. **9**, 6, 064023, DOI: 10.1088/1748-9326/9/6/064023.

Stott, P.A., D.A. Stone, and M.R. Allen (2004), Human contribution to the European heatwave of 2003, *Nature* **432**, 7017, 610-614, DOI: 10.1038/nature03089.

Trigo, R.M., J.M. Vaquero, M.J. Alcoforado, M. Barriendos, J. Taborda, R. García-Herrera, and J. Luterbacher (2009), Iberia in 1816, the year without a summer, *Int. J. Climatol.* **29**, 1, 99-115, DOI: 10.1002/joc.1693.

Twardosz, R. (2009), Extraordinary heat waves in 21st century Europe, *Prz. Geofiz.* **2009**, 3-4, 193-204 (in Polish).

Twardosz, R., and A. Batko (2012), Heat waves in Central Europe (1991-2006), *Int. J. Global Warm.* **4**, 3-4, 261-272, DOI: 10.1504/IJGW.2012.049430.

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. Kadioglu (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 (2013c), Exceptional thermal anomalies in the Atlantic-European area of the sub-polar zone, *Probl. Klimatol. Polar.* **23**, 93-105 (in Polish).

Twardosz, R., E. Łupikasza, and T. Niedźwiedź (2011), *Temporal Variability in the Form and Type of Precipitation in Kraków in Relation to Circulation Patterns*, Wydawnictwo UJ, Kraków, 174 pp. (in Polish).

Unal, Y.S., E. Tan, and S.S. Mentes (2013), Summer heat waves over western Turkey between 1965 and 2006, *Theor. Appl. Climatol.* **112**, 1-2, 339-350, DOI: 10.1007/s00704-012-0704-0.

Wehry, W. (2010), Der heisse Sommer 2010 in Russland – Teil 2: Meteorologische Ursachen der Waldbrände, Beitr. Berliner Wetterkarte 52/10 (SO 24/10) (in German).

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**, 6, 679-692, DOI: 10.1002/joc.906.

WMO (2010), Extreme weather events, World Meteorological Organization, http://www.wmo.int/pages/mediacentre/news/extremeweathersequence_en.html (retrieved on 9 February 2011).

Zveryaev, I., Yu. Zyulyaeva, S. Gulev, and P. Koltermann (2012), Intercomparison of the Russian summer heatwaves of 2010 and 1972. **In:** *EGU General Assembly, 22-27 April 2012, Vienna, Austria*, EGU2012-9714.

Zvyagintsev, A.M., O.B. Blum, A.A. Glazkova, S.N. Kotel'nikov, I.N. Kuznetsova, V.A. Lapchenko, E.A. Lezina, E.A. Miller, V.A. Milyaev, A.P. Popikov, E.G. Semutnikova, O.A. Tarasova, and I.Yu. Shalygina (2011), Air pollution over European Russia and Ukraine under the hot summer conditions of 2010, *Izv., Atmos. Ocean. Phys.* **47**, 6, 699-707, DOI: 10.1134/S0001433811060168.

Received 19 August 2014
Received in revised form 5 November 2014
Accepted 18 November 2014

DE GRUYTER OPEN

Acta Geophysica
vol. 63, no. 1, Feb. 2015, pp. 301-318
DOI: 10.2478/s11600-014-0251-4

Mitigation of Oceanic Tidal Aliasing Errors in Space and Time Simultaneously Using Different Repeat Sub-Satellite Tracks from Pendulum-Type Gravimetric Mission Candidate

Basem ELSAKA[1,2], Karl Heinz ILK[3], and Abdulaziz ALOTHMAN[1]

[1]Space and Aviation Research Institute, King Abdulaziz City for Science and Technology (KACST), Riyadh, Saudi Arabia
e-mail: balsaka@kacst.edu.sa (corresponding author)

[2]National Research Institute of Astronomy and Geophysics (NRIAG), Helwan, Cairo, Egypt

[3]Institute of Geodesy and Geoinformation, University of Bonn, Bonn, Germany

A b s t r a c t

This contribution investigates two different ways for mitigating the aliasing errors in ocean tides. This is done, on the one hand, by sampling the satellite observations in another direction using the pendulum satellite mission configuration. On the other hand, a mitigation of the temporal aliasing errors in the ocean tides can be achieved by using a suitable repeat period of the sub-satellite tracks.

The findings show, firstly, that it is very beneficial for minimizing the aliasing errors in ocean tides to use pendulum configuration; secondly, optimizing the orbital parameter to get shorter repeat orbit mode can be effective in minimizing the aliasing errors. This paper recommends the pendulum as a candidate for future gravity mission to be launched in longer repeating orbit mode with shorter "sub-cycle" repeat periods to improve the temporal resolution of the satellite mission.

Key words: gravity field recovery, repeat sub-satellite tracks, ocean tides aliasing.

1. INTRODUCTION

The products that the twin-satellite GRACE (Gravity Recovery and Climate Experiment) mission (Tapley *et al.* 2004) has provided in the form of numerous model series such as the EIGEN and ITG-GRACE series (*e.g.*, Förste *et al.* 2008, Mayer-Gürr *et al.* 2010), respectively have not yet matched pre-mission expectations in terms of error level and error isotropy. The main limitation of the quality of the resolved temporal gravity field estimates is mostly controlled not only by the instrument noise, but also the anisotropy of spatial sampling and temporal aliasing errors (*i.e.*, the errors in the modeling of mass changes due to the high frequency signals such as atmosphere, ocean, and ocean tides). The latter effects are related to the GRACE orbital configuration because of the inhomogeneous sampling in time and space. Therefore, to minimize the errors in the temporal gravity field models, *e.g.*, ocean tides, one has to optimize and/or improve the determination of the gravitational signal spatially and temporally. On the one hand, a spatially homogeneous sampling can be achieved via the selection of an alternative mission type whose satellite observables are sensitive in other directions (*e.g.*, radial and/or cross-track) compared to the GRACE along-track observable. On the other hand, adjusting the orbit and formation parameters can improve the temporal sampling of the mission. These parameters include the orbital altitude, the inter-satellite distance, the inclination, the repeat mode of sub-satellite tracks (as projected satellite orbits on the Earth's surface) and, of course, the choice of the number of satellites and satellite links to create a possible multi-satellite/formation mission. By means of an appropriate choice of these parameters, isotropy can be enlarged and aliasing effects (the most problematic issue of GRACE mission) can be reduced.

It should be mentioned here that there are common approaches for reducing the temporal aliasing effects, such as smoothing techniques with Gaussian and/or de-correlation filters (see *e.g.*, Wahr *et al.* 2004, Swenson and Wahr 2006, Kusche 2007). However, it was found that the impact of such filters is partially undesirable as a part of the desired gravity signal is smoothed besides the errors.

It is important first to mention that a variety of studies was published in the previous years which have investigated the performance of the basic types of satellite formation missions, *e.g.*, pendulum, cartwheel and LISA (see, *e.g.*, Sharifi *et al.* 2007, Sneeuw *et al.* 2008, Wiese *et al.* 2009, Elsaka *et al.* 2012, 2014a). All these studies have found that the latter three missions would provide a lower error spectrum with improved isotropy. In addition, the arrangement of a second, inclined satellite pair in the so-called "Bender design" was studied by Bender *et al.* (2008), Visser *et al.* (2010), Wiese *et al.* (2011a, 2012), and Elsaka *et al.* (2014b). All of the above-mentioned

studies show a common result that a significant increase in accuracy and sensitivity is expected when a future formation will be flown in an alternative configuration, different from the GRACE leader-follower scheme. Furthermore, Visser *et al.* (2010), Wiese *et al.* (2011b), and Elsaka (2014) have studied the feasibility of estimating low resolution gravity fields at short periods via a single and double pairs of satellites similar to GRACE to reduce the effect of temporal aliasing errors from mass variations with large spatial scales.

The aim of this paper is to focus on reducing the aliasing errors in the ocean tide models in space and time simultaneously, *i.e.*, spatially via selecting the "cross-track" pendulum configuration as an alternative mission candidate for future satellite gravimetry (Elsaka *et al.* 2012, 2014a), and temporally via choosing an appropriate repeat mode of the sub-satellite tracks.

Based on the above-mentioned studies, this paper simulates satellite observations of the "cross-track" pendulum mission, in addition to the GRACE configuration as a reference mission (for a comparative reason), since the other radial configurations are more technically challenging. It is well known that when the GRACE data are analyzed, a set of background models for tidal and non-tidal oceanic, atmospheric and hydrologic mass change are applied to mitigate the aliasing effects; however, remaining errors in these models still alias into the monthly GRACE solutions and manifest themselves as artefacts.

Therefore, this paper applies two oceanic tidal models: FES2004 (Lyard *et al*. 2006) and EOT2008a (Savcenko and Bosch 2008), assuming that the differences of two state-of-the-art oceanic tidal models are representative of their errors. The simulation scenario has been performed using the IGG's GROOPS (Gravity Recovery Object Oriented Programming System) software (Mayer-Gürr 2006), and the gravity results are analyzed in the spectral and spatial domain complete up to degree and order (d/o) 60/60.

This paper is organized as follows: in Section 2, a review concerning the computation of the repeat periods of sub-satellite tracks and the applied repeat modes in this contribution is outlined. The pendulum mission configuration is discussed in Section 3. The simulation strategy used for the gravity field analysis is introduced in Section 4. The gravity field solutions in terms of ocean tide aliasing errors are presented in Section 5. Finally, on this basis, a relevant conclusion is outlined in Section 6.

2. COMPUTATION OF THE REPEAT PERIOD OF SUB-SATELLITE TRACKS

Repeat sub-satellite track means simply that the sub-satellite track retraces itself exactly after a certain time. If the satellite orbit should repeat itself

whilst the Earth was not rotating or the satellite orbital plane was fixed in the Earth's fixed frame, the two satellite crossings at the equator would occur at the same site. Since neither Earth rotation nor the precession of the orbital plane can be neglected, the shift between two ascending nodes takes place. Because the precession of the ascending nodes is much slower than the Earth's rotation, a nodal day differs slightly from a solar day, and in case of a sun-synchronous orbit (*e.g.*, $i = 95°$) they are equal (see Bezděk *et al.* 2009). The nodal day corresponds to the time that the Earth takes to complete one revolution with respect to the orbital plane, while the solar day corresponds to the time required for the Earth to complete one revolution with respect to the Sun–Earth line.

Taking the Earth's rotation rate ($\dot{\omega}$) with respect to the satellite's orbital plane, the notion of a nodal period (P_n) is therefore orbit-dependent and is defined as (after Rees 2001)

$$P_n = \frac{2\pi}{\dot{\omega}} = \frac{2\pi}{(\omega - \dot{\Omega})}\,, \tag{1}$$

where ω is the angular velocity of the Earth, and $\dot{\Omega}$ is the precession rate of the satellite's line of node. After an integer number (α) of Earth rotations in the time required for the satellite to make an integral number of orbits (β), the condition in Eq. 1 can be written as

$$P_n(\omega - \dot{\Omega}) = 2\pi\frac{\alpha}{\beta}\,. \tag{2}$$

The nodal period is equal to the Keplerian period if the perturbations were absent. However, in presence of perturbations, the secular change in the satellite's argument of perigee $\dot{\omega}$ and the secular change in the satellite's mean anomaly $\dot{M}$ must be taken into account (*i.e.*, $P_n = 2\pi/(\dot{\omega} + \dot{M})$). Thus, Eq. 2 can be rewritten in terms of the classical orbital elements as

$$P_n = 2\pi\frac{\alpha}{\beta(\omega - \dot{\Omega})} = \frac{2\pi}{(\dot{\omega} + \dot{M})}\,, \tag{3}$$

and hence,

$$\alpha(\dot{\omega} + \dot{M}) \approx \beta(\omega - \dot{\Omega})\,. \tag{4}$$

Equation 4 represents the repeat period condition and establishes the synchronicity between the Earth rotation and the satellite rotation in a way that the satellite completes β nodal revolutions while the Earth performs α rotations. The three secular changes, $\dot{\omega}$, $\dot{M}$, and $\dot{\Omega}$, in Eq. 2 are calculated according to Kaula (1966) as

$$\begin{aligned}
\frac{\mathrm{d}\Omega}{\mathrm{dt}} &= \frac{3nC_{20}R^2}{2a^2(1-e^2)^2}\cos i\ , \\
\frac{\mathrm{d}\omega}{\mathrm{dt}} &= \frac{3nC_{20}R^2}{4a^2(1-e^2)^2}(1-5\cos^2 i)\ , \\
\frac{\mathrm{d}M}{\mathrm{dt}} &= n-\frac{3nC_{20}R^2}{2a^2(1-e^2)^{3/2}}(3\cos^2 i-1)\ ,
\end{aligned} \tag{5}$$

where n is the satellite's mean motion, C_{20} is the second zonal term of the geopotential, R is the Earth's radius, a is the orbital semi-major axis, and e is the orbital eccentricity. When the Earth takes approximately 365.25 days to orbit around the Sun, this means that it performs exactly 1.002737925 rev (*i.e.*, 1 + 1/365.25) per day. In this way, the duration of the sidereal day becomes exactly 23.9345 hours and the Earth's angular velocity reads $\omega = 2\pi/23.6345 \times 60 \times 60 = 7.2921 \times 10^{-5}$ rad/s.

The repeat period condition (Eq. 3) is related to the mean motion of the satellite, after neglecting the term e^2 (due to its small value) in Eq. 5, using the nodal period of the satellite $2\pi/(\dot{\omega}+\dot{M})$, and the nodal day $2\pi/(\omega-\dot{\Omega})$, as

$$n = \frac{\alpha}{\beta}(\omega-\dot{\Omega})-(\dot{\omega}+\dot{M})\ . \tag{6}$$

After one inserts the secular rates calculated from Eq. 5 into Eq. 4, the ratio β/α can be easily computed. Equation 6 reads for the first order in J_2 (where J_2 is the normalized coefficient $C_{20}\sqrt{5}$) according to Bezděk *et al.* (2009) as

$$n = \frac{\beta}{\alpha}\omega\left(1-\frac{3}{2}j_2\left[\frac{R}{a}\right]^2\left[4\cos^2 i-\frac{\beta}{\alpha}\cos i-1\right]\right)\ . \tag{7}$$

The ratio β/α in Eq. 7 depends on the orbital inclination and altitude. In this paper, an eccentricity $e = 0.001$ and an inclination of $i = 89.5°$ for both pendulum and GRACE configurations are selected, so one can focus here on the change of repeat sub-satellite tracks according to different orbital altitudes. Figure 1 shows different satellite orbits revolutions and their projected sub-satellite tracks on an Earth's map with the ratio β/α equal to 31/2, 170/11, 448/29, 247/16, and 108/7, corresponding to the orbital altitudes of 407, 410, 212, 415, and 420 km, respectively.

The selection of those orbital altitudes is based on the fact that the gravitational signal rapidly decays due to the so-called inverse attenuation factor, $[r/R_e]^{2n+1}$, which is a function of the spherical harmonic degree n. The term r

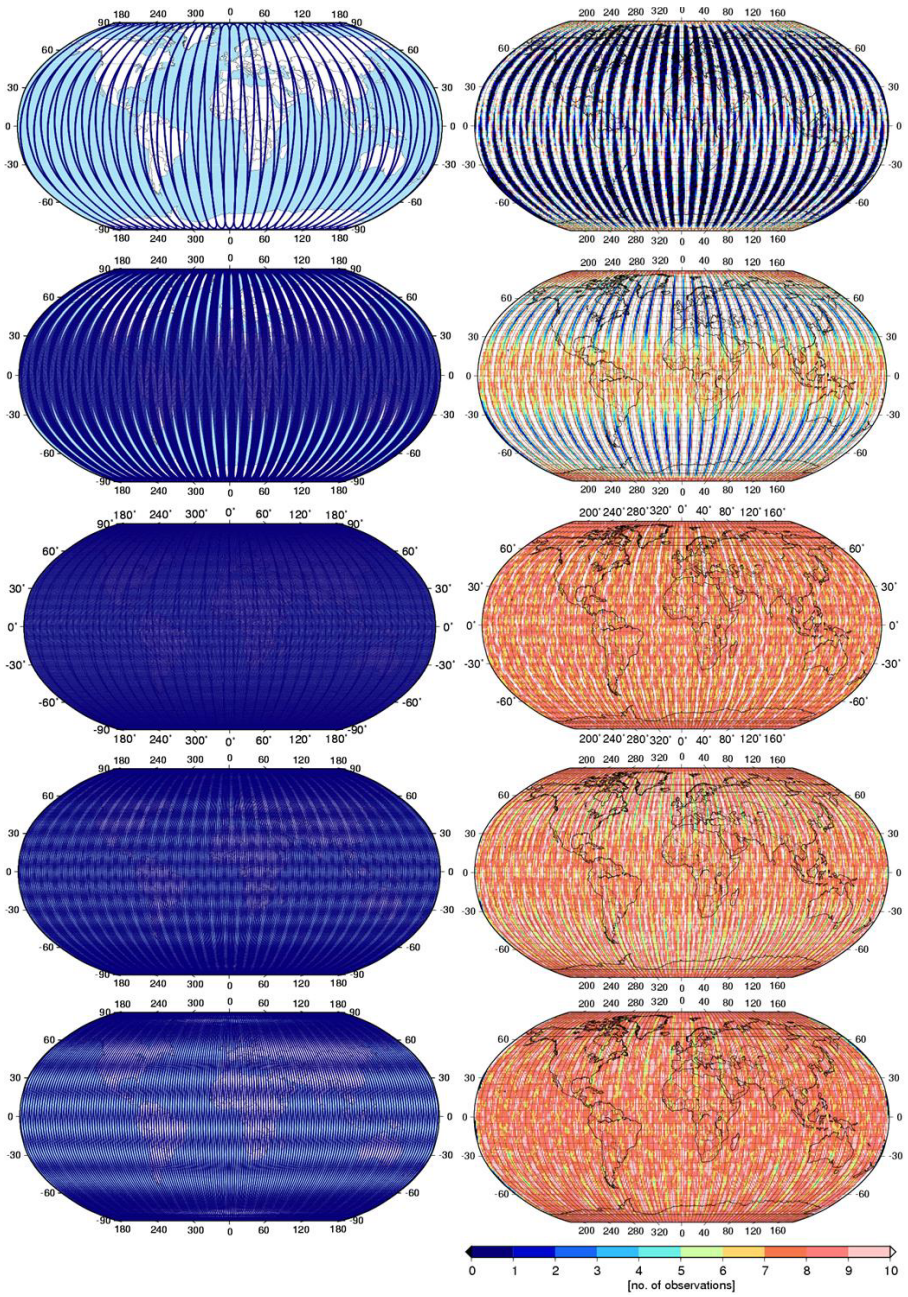


Fig. 1. Different satellite orbit revolutions (left column) and their corresponding projected sub-satellite tracks on the Earth's surface (right column) for different orbital altitudes; from top to bottom: repeat sub-satellite tracks with the ratios β/α of 31/2 (407 km), 170/11 (410 km), 448/29 (412 km), 247/16 (415 km), and 108/7 (420 km).

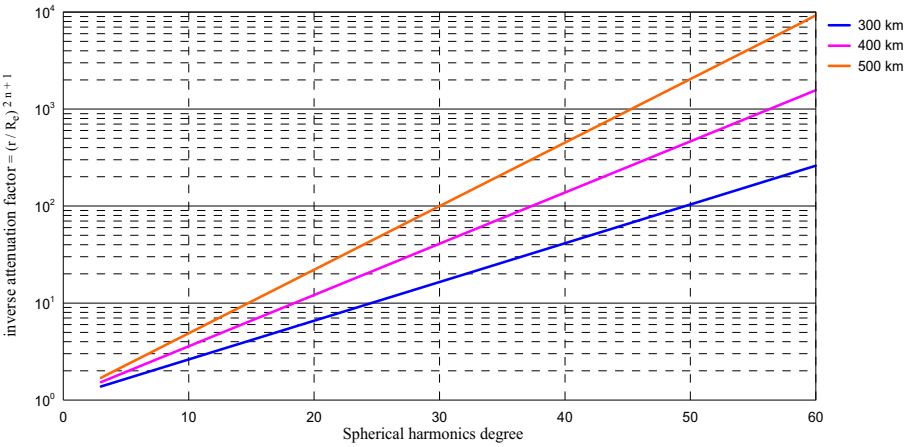


Fig. 2. The inverse attenuation factor showing the effect of orbital altitude on the gravity field recovery.

stands for the orbit altitude and R_e stands for the Earth's radius. The exponent n indicates that the higher the orbital altitude (and consequently the larger the r), the worse the resolution of the gravity field recovery (see, *e.g.*, Fig. 2).

Therefore, nearby orbital altitudes between 407 and 420 km are selected here in order to achieve different repeat modes without affecting the strength of the gravitational signal. The selected orbital altitudes represent different sub-satellite tracks that cover the Earth, insufficiently and sufficiently, as shown in Fig. 1.

3. PENDULUM AND GRACE MISSION CONFIGURATIONS

The GRACE-type configuration (Fig. 3, left panel) is considered as a simple collinear formation flying with two identical satellites separated from each other by approximately 220 km. The GRACE observables are sensitive only in an along-track direction. This is considered as the main drawback of the configuration geometry that no measurement information of the gravitational signal is collected in the cross-track and the radial direction. Therefore, one sees clearly a distortion of the monthly GRACE gravity estimates in the form of a longitudinal striping pattern. A common measure to counteract these effects is filtering the solutions (see, *e.g.*, Swenson and Wahr 2006, Kusche 2007); however, it is known that these procedures remove not only errors but signals as well.

It is generally accepted that follow-on missions should improve in sensitivity and isotropy by involving cross-track or radial information in the satellite observables. Radial information can be gathered via cartwheel and LISA

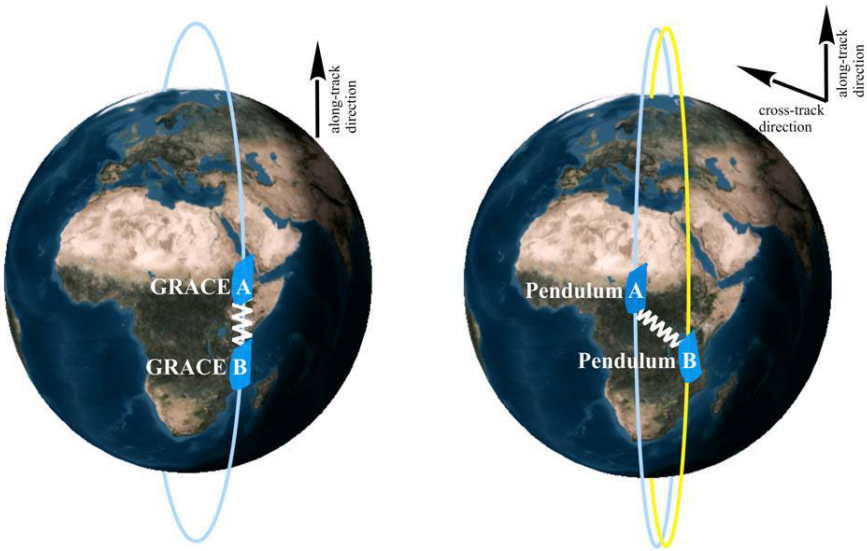


Fig. 3. The investigated satellite mission configurations: GRACE (left panel) and pendulum (right panel).

satellite mission scenarios, while cross-track information can be obtained via the pendulum mission scenario. Elsaka *et al.* (2012) has recommended the pendulum configuration as a future satellite-gravity mission candidate due to the relatively lower relative motion between its both satellites compared to the cartwheel and LISA configurations. Additionally, the pendulum mission was highly recommended by different scientific institutions as a future mission after GOCE (Gravity recovery and steady-state Ocean Circulation Explorer) era as proposed by the E.motion (Earth System Mass Transport Mission) team (see Panet *et al.* 2013) during the ESA (European Space Agency) call for proposals for Earth Explorer Opportunity Mission EE-8. However, they did not investigate the aliasing effects of ocean tides in their study. Therefore, the pendulum configuration at different sub-satellite track periods is considered here to investigate the aliasing errors of ocean tidal signal as a temporal gravity field. The cross-track link has been achieved between the two pendulum satellites by setting different angles of the right ascension of ascending nodes ($\Delta\Omega$) and of the mean anomaly ΔM (to achieve a non-zero cross-track component and to avoid the risk of collision of the two satellites). Another possibility to obtain cross-track formation with non-zero differential inclination is also achievable. However, this option is not guaranteed, since a non-trivial solution of the linear equation system exists (Sneeuw *et al.* 2008). Table 1 indicates the orbital parameters applied for both the GRACE and pendulum satellites. For comparative purposes, the

Table 1

Differential Keplerian orbital parameters
for the pendulum and GRACE missions

Orbital parameters	Formation flights	
	GRACE	Pendulum
Δa	0	0
Δe	0	0
Δi [deg]	0	0
$\Delta\Omega$ [deg]	0	0.45
$\Delta\omega$ [deg]	0	0
ΔM [deg]	1.72	1.72

inter-satellite distance of the pendulum configuration has been set similar to that of GRACE, *i.e.*, approximately 200 km (Fig. 4). At majority of the time, *i.e.*, from pole to equator or *vice versa*, the inter-satellite distance of the pendulum configuration contains a varying cross-track component (see Fig. 3, right panel). Yet, the pendulum configuration investigated here still fulfills requirements suggested by Elsaka *et al.* (2012), where the separation angle between the two satellites is 0.45° and with inter-satellite velocities of approximately ±8 m/s.

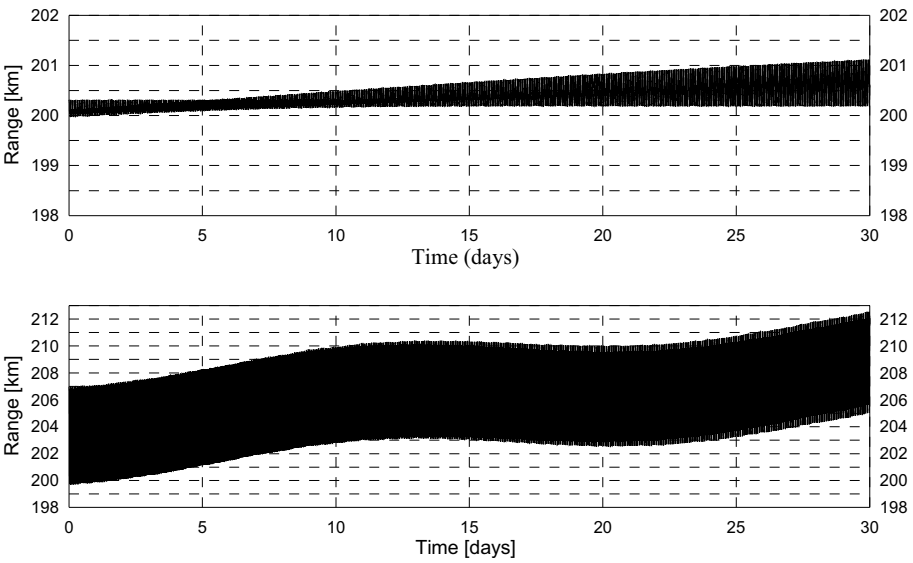


Fig. 4. The inter-satellite range (in km) for the investigated satellite configurations: GRACE (top panel) and pendulum (bottom panel).

4. SIMULATION STRATEGY

In order to compare the performance of both GRACE and pendulum configuration scenarios, numerical simulations using the Gravity Recovery Object Oriented Programming System (GROOPS) software have been performed. GROOPS has been developed in the Astronomical, Physical, and Mathematical Geodesy group at Bonn University to estimate gravity field parameters from satellite measurements, and it is routinely used to compute, *e.g.*, the ITG-GRACE solutions (see, *e.g.*, Mayer-Gürr *et al.* 2010). The mathematical description of GROOPS is given in details in Mayer-Gürr (2006).

For this study, all satellite orbits were integrated at 10-second steps using ITG-GRACE03s (Mayer-Gürr *et al.* 2010) as background mean gravity field model. In addition, background time-variable gravity field models have been used for ocean tidal forces as well as "non-tidal" atmospheric, "non-tidal" oceanic and hydrologic mass variations, as shown in Table 2. In the gravity analysis step, the same set of force models has been used except for the ocean tidal models to simulate the aliasing errors in the ocean tidal masses, as described in Section 1. Subsequently, each satellite orbit was corrupted with a Gaussian standard deviation of 1 cm to mimic errors of POD (precise orbit determination). The range measurements were nominally contaminated by white noise with standard deviation 50×10^{-9} m assuming the next generation of gravity missions will be equipped with a laser interferometric ranging system (Bender *et al.* 2003) as it is currently being developed for the GRACE follow-on mission. And therefore, this paper has not used accelerometer data for the pendulum scenario, as it has been set-up in this paper to be a drag-free mission. Moreover, applying higher noise levels (*e.g.*, micro levels in SST K-Band) leads to larger gravity field recovery errors than

Table 2

Background models applied to mean and time-variable simulation scenarios

Simulation scenario / Force function model	Measurement noise case		Oceanic tidal aliasing case	
	Orbit integration step	Gravity analysis step	Orbit integration step	Gravity analysis step
Mean field	ITG-GRACE03s	ITG-GRACE03s	ITG-GRACE03s	ITG-GRACE03s
Atmosphere	ECMWF	ECMWF	ECMWF	ECMWF
Ocean	OMCT	OMCT	OMCT	OMCT
Hydrology	WGHM	WGHM	WGHM	WGHM
Ocean tide	**FES2004**	**FES2004**	**FES2004**	**EOT08a**

the simulated tide model differences (*e.g.*, ocean tides) (see Visser *et al.* 2010, Elsaka 2010, p. 108).

Since this paper is focused mainly on the mitigation of aliasing errors induced in the oceanic tidal signal, we have decided to separate the impact of the error sources from the other time-varying non-tidal signals. This would help us to interpret the results more accurately regarding the effect of the shorter orbital repeating modes for minimizing the errors. Therefore, all models applied in the orbit integration step were reduced in the gravity analysis step. However, the investigated scenarios contain in reality some remaining aliasing errors that are resulting from the non-tidal signals, since we apply 6-hourly atmosphere and ocean models and daily hydrological model during the orbit integration step.

5. RESULTS

The results are provided in the "long-to-medium" spectral domain in terms of spherical harmonic coefficients up to degree/order (d/o) 60/60. The gravity field estimates are further visualized in terms of error degree-variances of the geoid heights, as shown in Figs. 5 and 6 for the measurement noise and oceanic tidal aliasing scenarios, respectively. In the spatial domain, geoid error maps are constructed in Fig. 7 for the oceanic tidal aliasing case. Moreover, Fig. 8 shows the gravity field solutions of the aliasing case considering only the individual semi-diurnal M2 tidal constituent determined by the pendulum configuration. The corresponding statistics in terms of global root mean square (RMS) values are given separately in Table 3. All error curves shown in Figs. 5 and 6 are obtained from the difference between outputs (estimates) and input (true model, *i.e.*, ITG-GRACE03s). This means that the mean field based on ITG-GRACE03s had to be removed from the monthly recovered solution in order to obtain the residual monthly gravity signal, due to the simulated error of the ocean tides signal.

As expected, all estimated solutions for the pendulum mission scenario perform approximately a half to one full order of magnitude (Table 3) better than the GRACE reference solution, in particular, at the medium wavelength range, as seen in Fig. 5 for the measurement noise case. However, Table 3 shows that the GRACE solutions at orbital heights of 410, 415, and 420 surpass those determined by the pendulum scenario. The reason is clearly identified in Fig. 5 that the pendulum configuration recover the harmonic coefficients at the long wavelength range (up to d/o 8) worse than the GRACE configuration. Yet the pendulum surpasses GRACE within the remaining long as well as medium range (*i.e.*, from d/o 9 up to d/o 60). According to the oceanic tidal aliasing case (Fig. 6), the pendulum solutions outperform the GRACE ones. This can be clearly seen in Table 3 and spatially on the

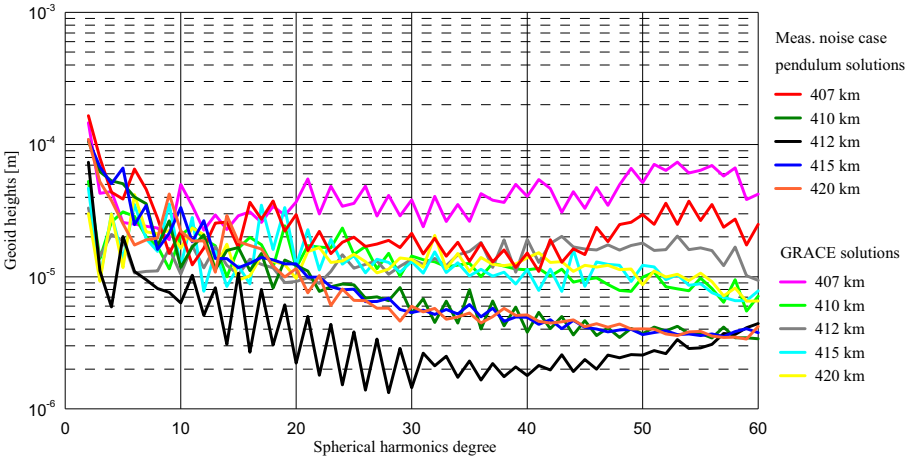


Fig. 5. Gravity solutions from the GRACE and pendulum mission scenarios in terms of error degree-variances of geoid heights for the measurement noise case.

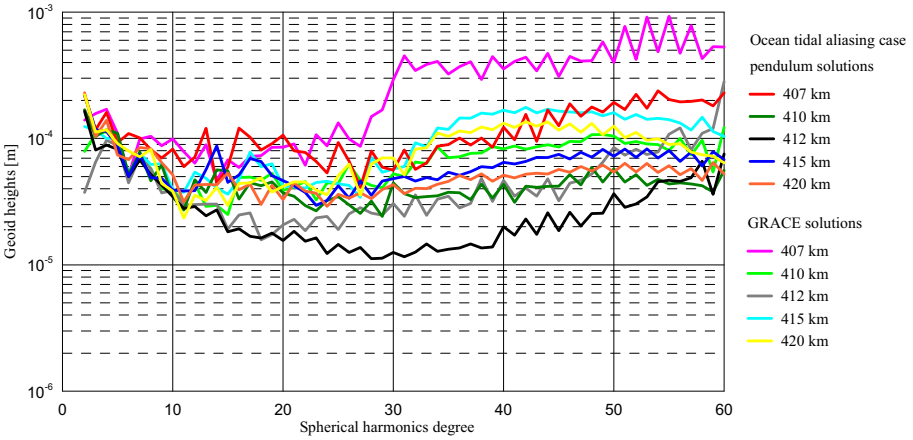


Fig. 6. Gravity solutions from the GRACE and pendulum mission scenarios in terms of error degree-variances of geoid heights for the ocean tidal aliasing case.

Earth's map in Fig. 7, which shows the global geoid height errors. The reason for this performance is that the pendulum configuration adds measurement information in the cross-track directions, and hence it helps in improving the retrieval of the mean and temporal gravity signal.

The north-south striping pattern has been obviously reduced via the pendulum gravity estimates in comparison to the GRACE reference solutions, which display a stronger striping pattern, as expected (see Fig. 7). This means that the pendulum configuration would not only be able to reduce aliasing errors but also to resolve the temporal signal better.

Fig. 7. Geoid height differences (in m) between the simulated static gravity field ITG-GRACE03s and the recovered solutions of the oceanic tidal aliasing case for the GRACE (left column) and pendulum (right column) mission scenarios according to the orbital heights from top to bottom: 407, 410, 412, 415, and 420 km.

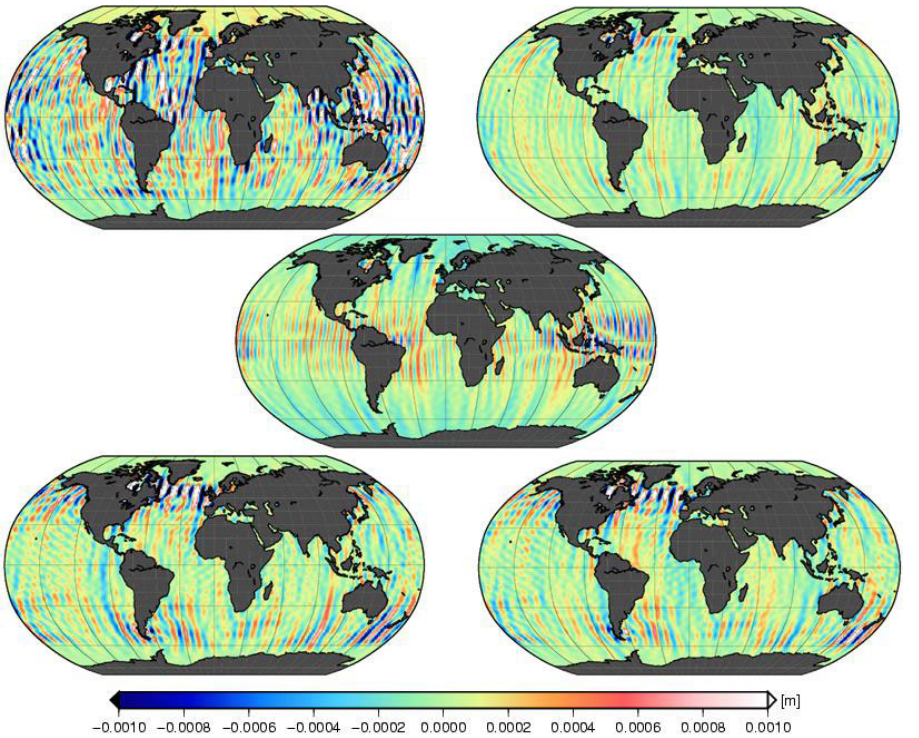


Fig. 8. Recovered solutions concerning the aliasing errors in M2 tidal constituent in terms of geoid height differences (in mm) for the pendulum mission scenarios according to the orbital heights of 407 km (top-left), 410 km (top-right), 412 km (middle), 415 km (bottom-left), and 420 km (bottom-right).

Table 3

Statistical values (RMS) in terms of geoid heights [mm] for both cases given in Table 2 determined by GRACE and pendulum configuration types at d/o 60 in addition to the M2 aliasing case

Orbital altitude [km]	Repeat period [days]	Measurement noise case (RMS)		Oceanic tidal aliasing case (RMS)		M2 aliasing case
		GRACE	Pendulum	GRACE	Pendulum	Pendulum
407	2	0.37	0.27	3.05	1.04	0.43
410	11	0.12	0.16	0.59	0.39	**0.16**
412	29	**0.11**	**0.07**	**0.51**	**0.32**	0.19
415	16	0.13	0.18	0.91	0.55	0.31
420	7	0.12	0.16	0.74	0.50	0.26

Note: bolded are the smallest RMS values.

According to the performance of the gravity field retrievals concerning the different repeating modes of sub-satellite tracks, one finds that the 29-day repeating mode provides the least oceanic tidal aliasing errors for both GRACE and pendulum configurations. This is due to the sufficient satellite observables and the adequate Earth coverage. Strong improvements have been found also for the 11-day repeating for mitigating the aliasing errors as seen in Table 3, which behaves similar to the 29-day repeating.

Since the pendulum configuration provides the least oceanic tidal aliasing errors, it has been additionally examined here how different repeating modes could mitigate the aliasing errors of the individual tides. For this case, the semi-diurnal M2 tidal constituent was selected. One can find here that both the 29-day and 11-day repeating still result in similar improvements; however, the 11-day repeating outperforms the 29-day repeating. This can be seen in Fig. 8 showing that the M2 aliasing errors have been minimized by the 11-day repeating mode better than by the other repeating modes.

One can infer that although the 29-day solution may be the best choice for minimizing the temporal aliasing errors, its temporal resolution is still not enough to resolve errors in the individual temporal signals. Therefore, it is recommended that the pendulum, as a candidate for a future gravity mission for detecting the temporal variations of the Earth's gravity field, is flown in a 29-day repeating mode allowing sufficient satellite observations with, *e.g.*, 11-day sub-cycle repeating at the same time to allowing a better understanding of the temporal variations of the individual time-varying signals.

6. CONCLUSION

In the course of this paper, the effect of different repeating modes concerning the sub-satellite tracks of the pendulum configuration has been investigated for the mitigation of the oceanic tidal aliasing errors.

The first conclusion is that the aliasing errors in the ocean tides reduce significantly if moderate cross-track components are added to the SST observable. This is in principle expected to be beneficial when the satellite pair flies in alternative configuration, such as the pendulum mission. One can also confirm the earlier findings that the GRACE formation is sub-optimal in terms of the gravity field retrievals. Second, it has been found that the selection of proper "shorter" repeating sub-satellite tracks provides a better understanding of the aliasing errors than the longer repeating mode. The 29-day repeating mode ensures a sufficient global Earth coverage; however, it has been found that it cannot resolve the temporal aliasing effects of the individual ocean tides constituents (*e.g.*, M2 semi-diurnal signal). On the other hand, the 11-day repeating mode could resolve the ocean tidal aliasing errors of individual constituents better.

Finally, it is recommended that a future gravity mission is launched in an orbital altitude of 29-day repeating cycle implementing at the same time 11-day repeating sub-cycle, which would support the detection of mass variations at higher temporal frequencies.

Acknowledgements. The authors would like firstly to thank the Managing Editor, Dr. Z. Wisniewski, for his efforts during the publication process of this manuscript and would like also to thank the reviewers for their valuable comment. The financial support of King Abdulaziz City for Science and Technology (KACST) is gratefully acknowledged.

References

Bender, P.L., J.L. Hall, J. Ye, and W.M. Klipstein (2003), Satellite-satellite laser links for future gravity missions, *Space Sci. Rev.* **108**, 1-2, 377-384, DOI: 10.1023/A:1026195913558.

Bender, P.L., D.N. Wiese, and R.S. Nerem (2008), A possible dual-GRACE mission with 90 degree and 63 degree inclination orbits. **In:** *Proc. Third Int. Symp. on Formation Flying, Missions and Technologies, 23-25 April 2008, Noordwijk, Netherlands*, 23-25.

Bezděk, A., J. Klokočnik, J. Kostelecky, R. Floberghagen, and C. Gruber (2009), Simulation of free fall and resonances in the GOCE mission, *J. Geodyn.* **48**, 1, 47-53, DOI: 10.1016/j.jog.2009.01.007.

Elsaka, B. (2010), Simulated satellite formation flights for detecting the temporal variations of the Earth's gravity field, Ph.D. Thesis, University of Bonn, Bonn, Germany.

Elsaka, B. (2014), Sub-monthly gravity field recovery from simulated multi-GRACE mission type, *Acta Geophys*. **62**, 1, 241-258, DOI: 10.2478/s11600-013-0170-9.

Elsaka, B., J. Kusche, and K.-H. Ilk (2012), Recovery of the Earth's gravity field from formation-flying satellites: Temporal aliasing issues, *Adv. Space Res.* **50**, 11, 1534-1552, DOI: 10.1016/j.asr.2012.07.016.

Elsaka, B., J.-C. Raimondo, P. Brieden, T. Reubelt, J. Kusche, F. Flechtner, S. Iran Pour, N. Sneeuw, and J. Müller (2014a), Comparing seven candidate mission configurations for temporal gravity field retrieval through full-scale numerical simulation, *J. Geod.* **88**, 1, 31-43, DOI: 10.1007/s00190-013-0665-9.

Elsaka, B., E. Forootan, and A. Alothman (2014b), Improving the recovery of monthly regional water storage using one year simulated observations of two pairs of GRACE-type satellite gravimetry constellation, *J. Appl. Geophys.* **109**, 195-209, DOI: 10.1016/j.jappgeo.2014.07.026.

Förste, C., R. Schmidt, R. Stubenvoll, F. Flechtner, U. Meyer, R. König, H. Neumayer, R. Biancale, J.-M. Lemoine, S. Bruinsma, S. Loyer, F. Barthelmes, and S. Esselborn (2008), The GeoForschungsZentrum Potsdam/Groupe de Recherche de Gèodésie Spatiale satellite-only and combined gravity field models: EIGEN-GL04S1 and EIGEN-GL04C, *J. Geod.* **82**, 6, 331-346, DOI: 10.1007/s00190-007-0183-8.

Kaula, W.M. (1966), *Theory of Satellite Geodesy*. *Applications of Satellites to Geodesy*, Blaisdell Publ. Co., Waltham.

Kusche, J. (2007), Approximate decorrelation and non-isotropic smoothing of time-variable GRACE-type gravity field models, *J. Geod.* **81**, 11, 733-749, DOI: 10.1007/s00190-007-0143-3.

Lyard, F., F. Lefevre, T. Letellier, and O. Francis (2006), Modelling the global ocean tides: modern insights from FES2004, *Ocean Dynam.* **56**, 5-6, 394-415, DOI: 10.1007/s10236-006-0086-x.

Mayer-Gürr, T. (2006), Gravitationsfeldbestimmung aus der Analyse kurzer Bahnbögen am Beispiel der Satellitenmissionen CHAMP und GRACE, Ph.D. Thesis, University of Bonn, Bonn, Germany.

Mayer-Gürr, T., A. Eicker, E. Kurtenbach, and K.-H. Ilk (2010), ITG-GRACE: Global static and temporal gravity field models from GRACE data. **In:** F.M. Flechtner, T. Gruber, A. Güntner, M. Mandea, M. Rothacher, T. Schöne, and J. Wickert (eds.), *System Earth via Geodetic-Geophysical Space Techniques*, Springer, Berlin Heidelberg, 159-168, DOI: 10.1007/978-3-642-10228-8_13.

Panet, I., J. Flury, R. Biancale, T. Gruber, J. Johannessen, M.R. van den Broeke, T. van Dam, P. Gegout, C.-W. Hughes, G. Ramillien, I. Sasgen, L. Seoane, and M. Thomas (2013), Earth system mass transport mission (e.motion): A concept for future earth gravity field measurements from space, *Surv. Geophys*. **34**, 2, 141-163, DOI: 10.1007/s10712-012-9209-8.

Rees, W.G. (2001), *Physical Principles of Remote Sensing*, 2nd ed., Cambridge University Press, Cambridge, 343 pp.

Savcenko, R., and W. Bosch (2008), EOT08a – empirical ocean tide model from multi-mission satellite altimetry, Rep. No. 81, Deutsches Geodätisches Forschungsinstitut (DGFI), München, Germany.

Sharifi, M., N. Sneeuw, and W. Keller (2007), Gravity recovery capability of four generic satellite formations. **In:** A. Kiliçoglu, and R. Forsberg (eds.), *Proc. Symp. "Gravity Field of the Earth", General Command of Mapping*, *June 2007, Ankara, Turkey*, Spec. Issue 18, 211-216.

Sneeuw, N., M.A. Sharifi, and W. Keller (2008), Gravity recovery from formation flight missions. **In:** P. Xu, J. Liu, and A. Dermanis (eds.), *VI Hotine-Marussi Symposium on Theoretical and Computational Geodesy*, International Association of Geodesy Symposia, Vol. 132, Springer, Berlin Heidelberg, 29-34, DOI: 10.1007/978-3-540-74584-6_5.

Swenson, S., and J. Wahr (2006), Post-processing removal of correlated errors in GRACE data, *Geophys. Res. Lett.* **33**, 8, L08, 402, DOI: 10.1029/2005GL025285.

Tapley, B.D., S. Bettadpur, M. Watkins, and C. Reigber (2004), The gravity recovery and climate experiment: Mission overview and early results, *Geophys. Res. Lett.* **31**, 9, DOI: 10.1029/2004GL019920.

Visser, P.N.A.M., N. Sneeuw, T. Reubelt, M. Losch, and T. van Dam (2010), Space-borne gravimetric satellite constellations and ocean tides: aliasing effects, *Geophys. J. Int.* **181**, 2, 789-805, DOI: 10.1111/j.1365-246X.2010.04557.x.

Wahr, J., S. Swenson, V. Zlotnicki, and I. Velicogna (2004), Time-variable gravity from GRACE: First results, *Geophys. Res. Lett.* **31**, 11, L11501, DOI: 10.1029/2004GL019779.

Wiese, D.N., W.M. Folkner, and R.S. Nerem (2009), Alternative mission architectures for a gravity recovery satellite mission, *J. Geod.* **83**, 6, 569-581, DOI: 10.1007/s00190-008-0274-1.

Wiese, D.N., R.S. Nerem, and S.-C. Han (2011a), Expected improvements in determining continental hydrology, ice mass variations, ocean bottom pressure signals, and earthquakes using two pairs of dedicated satellites for temporal gravity recovery, *J. Geophys. Res.* **116**, B11, B11405, DOI: 10.1029/2011JB008375.

Wiese, D.N., P. Visser, and R.S. Nerem (2011b), Estimating low resolution gravity fields at short time intervals to reduce temporal aliasing errors, *Adv. Space Res*. **48**, 6, 1094-1107, DOI: 10.1016/j.asr.2011.05.027.

Wiese, D.N., R.S. Nerem, and F.G. Lemoine (2012), Design considerations for a dedicated gravity recovery satellite mission consisting of two pairs of satellites, *J. Geod.* **86**, 2, 81-98, DOI: 10.1007/s00190-011-0493-8.

Received 30 October 2013
Received in revised form 19 August 2014
Accepted 3 September 2014

DE GRUYTER OPEN

Acta Geophysica
vol. 63, no. 1, Feb. 2015, pp. 319-328
DOI: 10.2478/s11600-014-0246-1

Ionospheric Response to the Acoustic Gravity Wave Singularity

Olga N. SAVINA[1] and Peter A. BESPALOV[2]

[1]National Research University, Higher School of Economics, Nizhny Novgorod, Russia; e-mail: ONSavina@mail.ru (corresponding author)

[2]Institute of Applied Physics, Russian Academy of Sciences, Nizhny Novgorod, Russia; e-mail: PBespalov@mail.ru

A b s t r a c t

An original model of atmospheric wave propagation from ground sources to the ionosphere in the atmosphere with a realistic high-altitude temperature profile is analyzed. Shaping of a narrow domain with elevated pressure in the resonance region where the horizontal phase wave velocity is equal to the sound velocity is examined theoretically within the framework of linearized *Eq.s*. Numerical simulations for the model profiles of atmospheric temperature and viscosity confirm analytical result for the special feature of wave fields. The formation of the narrow domain with plasma irregularities in the *D* and low *E* ionospheric layers caused by the acoustic gravity wave singularity is discussed.

Key words: atmosphere, realistic temperature profile, ionosphere, acoustic gravity waves, *D* and *E* layers.

1. INTRODUCTION

An actual problem for many applications is explaining the role of acoustic gravity waves (AGW) in the transfer of oscillating processes from the Earth's surface to the upper atmosphere. The cause for the surface sources of these waves can be earthquakes, explosions, sea waves, and other artificial and natural processes (Blanc 1985). Both infrasound and internal waves play

an important role in different atmospheric phenomena (Francis 1975). Many important tasks for acoustics gravity waves in the Earth's atmosphere have been studied theoretically. A general form of wave *Eq.s* is too complex for analytical solution, even in the linear approximation. This is caused by the real inhomogeneity of the atmosphere. Altitude profiles of the atmospheric parameters are taking into account different analytical models (see, *e.g.*, Savina 1996). At present, the numerical methods of the solution of this problem are developed successfully and give a possibility of describing a propagation AGW taking into account different important factors, including the nonlinearity (see, *e.g.*, Lund and Fritts 2012).

In this paper, one of the aspects of the problem in question is examined. We carry out an analysis of AGW behavior near the resonance level, at which the condition of equality of horizontal phase wave velocity to the local value of the sound velocity is satisfied. We have shown that in the Earth's atmosphere the temperature profile is such that there is a range of wave phase velocities (or a frequency range with fixed horizontal dimensions of the source) in which the wave does not pass through a resonance domain. We have found that local disturbances of the wave pressure component are formed at the narrow resonance level.

The resonance mentioned above is the reason for ionospheric irregularities with relatively small vertical scales in the *D* and *E* layers. As will be shown in what follows, such irregularities, flattened vertically, contribute to the generation of sporadic *E* layers of the ionosphere with a large range of translucency.

Sections 2, 3, and 4 of this paper are complementary. Some analytical properties of the waves in a non-isothermal atmosphere are discussed in Section 2. Analytical conclusions on specific features of the fields of acoustic gravity waves at the resonance level are confirmed by numerical calculations, whose results are given in Section 3. Consideration of the ionospheric effects in the *D* and *E* ionospheric layers in Section 4 is based on the approximation of passive impurity and the results obtained in the foregoing sections.

2. WAVE FIELD NEAR THE RESONANCE LEVEL

The linearized system of equations of gas dynamics for the pressure perturbations $p_\sim$, the horizontal velocity $v_\sim$, and the vertical velocity $w_\sim$ is well known. We select axis z in the vertical direction and axis x in the horizontal direction. To simplify equations, it is convenient to introduce new field variables, namely: $V = (\rho/\rho_E)^{1/2} v_\sim$, $W = (\rho/\rho_E)^{1/2} w_\sim$, and $P = (\rho/\rho_E)^{1/2} p_\sim$, where ρ and ρ_E are the basic state densities in the current layer and at the ground level, respectively. Field variables are proportional to $\exp(i\omega t + ik_\perp x)$ for

a monochromatic signal with frequency ω in plane atmospheric layers. Then the linearized system of equations for the wave perturbation can be reduced to the following form (Gossard and Hooke 1975):

$$\begin{aligned}\left[-\omega^2+\omega_g^2(z)\right]W - i\frac{\omega}{\rho_E}\left[\frac{\partial}{\partial z}+\Gamma(z)\right]P = 0\ ,\\ \left[-\omega^2+c_s^2(z)k_\perp^2\right]P - i\omega\rho_E c_s^2(z)\left[\frac{\partial}{\partial z}-\Gamma(z)\right]W = 0\ .\end{aligned} \tag{1}$$

The perturbation of horizontal velocity can be determined as $V = (k_\perp/\omega)P$. In Eq. 1, i is the imaginary unit, $k_\perp$ is the horizontal wave number, $c_s = \sqrt{\gamma p/\rho}$ is the sound velocity, where

$$\begin{aligned}\omega_g^2 &= (\gamma-1)g^2/c_s^2 + (g/T)(\partial T/\partial z)\ ,\\ \Gamma &= (2-\gamma)g/2c_s^2 - (1/2T)(\partial T/\partial z)\ ,\end{aligned}$$

γ is the ratio of specific heats at constant pressure and volume, respectively, g is the acceleration due to gravity, and $T(z)$ is the basic state temperature of the atmosphere. In adiabatic approximation, we have

$$c_s(z) = 20.05\cdot T(z)^{1/2}\ [\mathrm{m/s}]\ ,$$

where T is Kelvin temperature. It is important for the following analysis that in the steady atmosphere (Lighthill 1978) the Ekkard parameter $\Gamma > 0$ and $\omega_g^2 > 0$ where ω_g is the Brunt Väisälä frequency. According to Eq. 1, the averaged vertical energy flux $S = (1/2)(PW^*+P^*W) = \mathrm{const}$, *i.e.*, S does not depend on altitude (Rapoport *et al.* 2004) a superscript asterisk is the sign of complex conjugation. Note that the basic-state temperature of the atmosphere $T(z)$ is altitude dependent. The dependence $\rho(z)$ is governed by the hydrostatic election $dp/dz = -\rho g$, where p is the basic state pressure. Taking into account the ideal gas law, we have the basic state density dependence in the form

$$\rho(z) = \rho_E T_E T^{-1}(z)\exp\left(-\int_0^z H^{-1}(z')dz'\right)\ ,$$

where $H(z)$ is the pressure scale height.

Some altitude domain near $z = z_*$, where

$$0 < \left[\omega - c_s(z)k_\perp\right]^2/\omega^2 < \varepsilon$$

and ε is small in the mathematical meaning with reasonably high accuracy, $W \approx 0$ the sound velocity and the Ekkard parameter are almost constants, and system 1 is reduced to the following form:

$$\left[\frac{\partial}{\partial z} + \Gamma(z)\right] P = 0\,, \quad z \neq z_*\,, \tag{2a}$$

$$\left[-\omega^2 + c_s^2(z)k_\perp^2\right] P = 0\,, \quad z = z_*\,. \tag{2b}$$

Consider in greater detail the processes near the resonance level $z = z_*$, which are described by the equation system 2. Pressure disturbances both above and below the level $z = z_*$ are satisfied according to Eq. 2a, but not according to Eq. 2b, because $\omega \neq c_s k_\perp$, and W is infinitely small (see equation system 1), but its derivative can be finite. Exactly at the level, $z = z_*$ the conditions $W = 0$ and $dW/dz = 0$ are fulfilled. Therefore, as follows from Eq. 2b, P can have an arbitrary value and Eq. 2a is not defined at this point. The absence of disturbances of both the vertical velocity and its derivative, leads to a conclusion that above level $z = z_*$ the solutions both for W and for P are identically equal to zero. Consequently, in order to balance the pressure jump at the level in question, the finite mass should be concentrated at level $z = z_*$, which is taken into account in the solution using a delta function.

A formal solution of equation system 2 near the level $z = z_*$ can be written as follows (Savina and Bespalov 2014):

$$\begin{aligned} &P = P_* E\left(z_* - z\right)\exp\left(\Gamma_*(z_* - z)\right) + \frac{P_* c_{s*}^2}{g}\delta\left(z - z_*\right)\,, \\ &W\left(z = z_*\right) = 0\,, \end{aligned} \tag{3}$$

where $E(z - z_*)$ is the Heaviside step function and $\delta(z - z_*)$ is the Dirac delta function; by a subscript asterisk we mark the values of the variables at the layer $z = z_*$. Solution 3 depends on a single constant P_*, which is determined by the boundary condition on the Earth's surface. Thus, local disturbance (Eq. 3) consists of a special feature in the altitude distribution of the pressure disturbance and characteristic structures $P(z)$ and $W(z)$ adjoining it from below. Actually, one or more resonance levels can exist in the real atmosphere for specified ω and $k_\perp$. Perturbations of the vertical velocity and pressure are absent ($W = 0$ and $P = 0$) everywhere above the first resonance level. Hence, if for the wave perturbation in the nonisothermal atmosphere at

some level $z = z_*$ a condition $\omega = c_s(z_*)k_\perp$ is satisfied, then the averaged vertical energy flux is equal to zero. Above the first of such levels, wave perturbation are absent along the vertical propagation path. This effect is responsible for the formation of a waveguide channel between the Earth's surface and the resonance level for the waves whose horizontal phase velocity is equal to the local sound velocity.

Under the actual conditions, the resonance in the form of a delta function in the pressure disturbance, as well as in the horizontal velocity perturbation that is proportional to it ($V = P/c_s(z = z_*)\rho_E$), is smeared due to the molecular viscosity and the nonlinearity. In the numerical calculations given below the molecular viscosity was taken into account.

3. FULL-WAVE CALCULATIONS FOR RESONANCE ATMOSPHERIC PERTURBATION

In this section, we determine perturbations of the pressure and vertical velocity by means of full-wave numerical calculations. We assume that on the ground level there is a monochromatic source of vertical velocity and that at the altitudes higher than 200 km level the atmosphere is isothermal. Numerical calculation of Eq. 1 made it possible to find the high-altitude distribution of the wave perturbations.

The wave fields are conveniently calculated in dimensionless variables, which we selected as follows:

$$\tilde{\omega} = \omega/\omega_{gE}\,, \quad \tilde{k}_\perp = k_\perp c_{sE}/\omega_{gE}\,, \quad \tilde{c}_s = c_s(z)/c_{sE}\,, \quad \tilde{W} = W/W_E\,, \quad \tilde{P} = P/\rho_E c_{sE}^2\,.$$

Here, subscript E indicates the value of the variable on the Earth's surface.

The temperature profile depends on many factors in the real atmosphere. We selected the altitude temperature profile following the MSIS-E-90 model (Hedin 1991). For example, we chose one temperature profile $T(z)$ for 10 August 2012 at 15:00 UT for geographic latitude 65° and longitude 45°. We determine the analytical function for this temperature profile (Fig. 1a) and kinematic viscosity profile (Fig. 1b) (Kikoin 1976) using the model values marked by boxes and stars in Fig. 1, respectively. The temperature and viscosity profiles were approximated by polynomials of the tenth order.

We give in Fig. 2 the results of numerical simulation for the inhomogeneous plane wave with the fixed frequency $\tilde{\omega} = 0.52$ and the horizontal wave number $\tilde{k}_\perp = 0.5$. Therefore, simulations are one-dimensional, and task is two-dimensional. These results correspond to theoretical examination in Section 2. Numerical calculation is conducted in much the same way as in papers Rapoport *et al.* (2004) and Bespalov and Savina (2012).

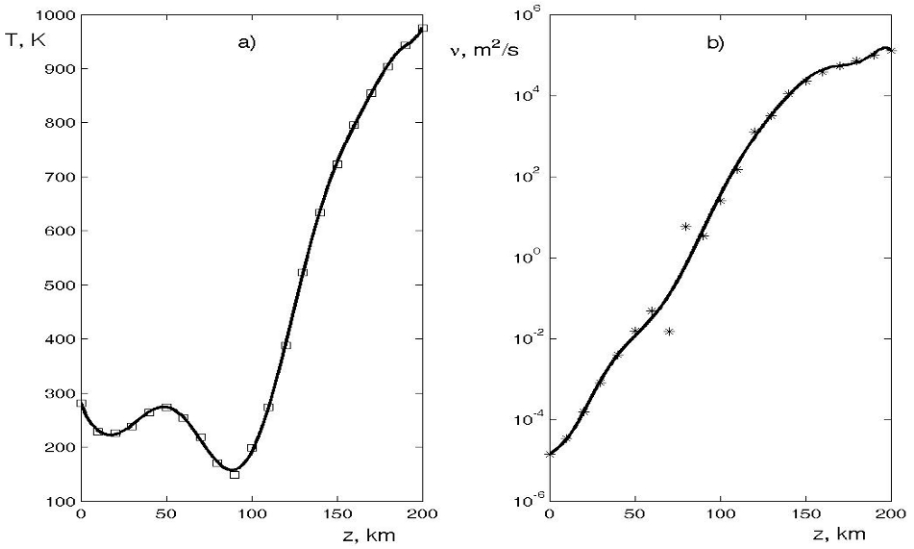


Fig. 1. Typical atmosphere temperature profile $T(z)$ as obtained from the MSIS-E-90 model marked by boxes and spline curve $T(z)$ (a), and typical altitude profile of molecular kinematic viscosity $\nu(z)$ in the Earth's atmosphere (b).

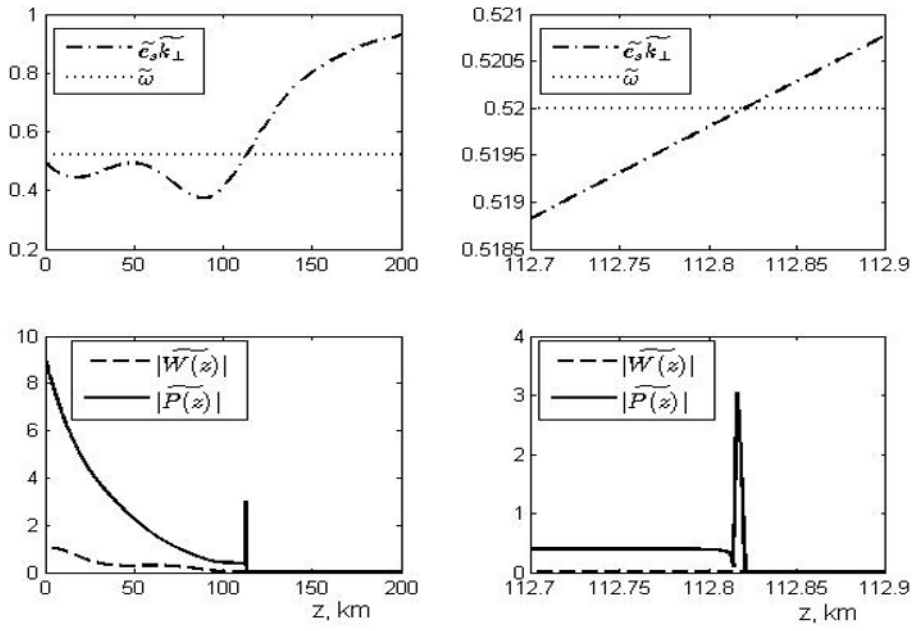


Fig. 2. Results of numerical calculations of AGW in the atmosphere (left column) and near the local disturbance (right column).

Figure 2 shows the local disturbance at an altitude near 113 km. In the upper panels, the dashed dotted curves are the altitude dependence $\tilde{c}_s(z)\tilde{k}_\perp$ and the dotted lines are the frequency $\tilde{\omega}$. In the bottom panels, the solid curves are the altitude dependence $\left|\tilde{P}(z)\right|$ and the dashed curves are the altitude dependence $\left|\tilde{W}(z)\right|$. In Figure 2 on the right bottom panel the dependencies $\left|\tilde{P}(z)\right|$ and $\left|\tilde{W}(z)\right|$ are given in greater detail for altitudes near 113 km, which correspond to the local disturbance. Calculations show that in this case the averaged vertical energy flux $\tilde{S} = 0$.

The local disturbance at the resonance level, which corresponds to solution 3, is limited by the atmospheric viscosity. Numerical calculations were carried out within the framework of the linearized system of *Eq.s*, which is correct for sufficiently small Reynolds numbers $\mathrm{Re} = \upsilon_\sim L/\nu$, where $\upsilon_\sim = \sqrt{\rho_E/\rho}\, P/c_s\rho_E$ is the disturbance of horizontal velocity, and L is a typical vertical scale of local disturbances. For conditions near the local disturbances, $\upsilon_\sim \simeq 10$ m/s, $L \simeq 20$ m, $\nu \simeq 300$ m^2/s, and $\mathrm{Re} < 1$.

4. IONOSPHERIC RESPONSE TO THE ATMOSPHERIC WAVE SINGULARITY

At the altitudes of the ionosphere, acoustic gravity waves are the reason for the occurrence of ionospheric irregularities. It was shown above that at the resonance level a singularity appears in the atmospheric gas pressure disturbance, knowing which one could estimate the horizontal velocity $\upsilon_\sim$. As numerical simulations show, resonance layers can be formed lower than the level given as an example in Fig. 2. This depends on the profile of temperature and values of $k_\perp$ and ω. The processes occurring in the neutral medium influence are passed here to the electrons and ions through collisions. At the altitudes of the *D* and *E* layers, the frequency of the acoustic gravity wave is much lower than the frequencies ν_{in} and ν_{en} of the ion and electron collisions with neutrals ($\nu_{\mathrm{in}} \approx 10^3$ s^{-1} and $\nu_{\mathrm{en}} \approx 10^4$ s^{-1} at an altitude close to 110 km). Then, neglecting the collisions between charged particles, for the velocity of drag of the electrons and ions by neutrals in the absence of external electric fields and sharp density gradients one can write (Gershman 1974)

$$\vec{u}_{\mathrm{e,i}} = \frac{\nu_{\mathrm{en,in}}^2}{\nu_{\mathrm{en,in}}^2 + \omega_{He,\mathrm{i}}^2}\left\{\vec{u}_{\mathrm{n}} \mp \frac{\omega_{He,\mathrm{i}}}{\nu_{\mathrm{en,in}}}\left[\vec{u}_{\mathrm{n}}\vec{h}_0\right] + \frac{\omega_{He,\mathrm{i}}^2}{\nu_{\mathrm{en,in}}^2}\vec{h}_0\left(\vec{h}_0\vec{u}_{\mathrm{n}}\right)\right\}, \qquad (4)$$

where $\vec{u}_{\mathrm{e,i,n}}$ are velocities of electrons, ions and neutrals, $\vec{h}_0$ is a unit vector along the geomagnetic field direction, and $\omega_{He,i}$ are the electron and ion gyrofrequencies values, respectively. The minus sign is chosen for the electron velocity and the plus sign is chosen for the ion velocity. The disturbance of the densities $N \approx N_{\mathrm{i}} \approx N_{\mathrm{e}}$ of the ionospheric plasma, which we assume to be quasi-neutral, can be estimated from the continuity equation:

$$\frac{\partial N}{\partial t} + \mathrm{div}\left(N\vec{u}_{\mathrm{i,e}}\right) = 0 \ . \tag{5}$$

At the altitudes of the *D* layer and in the lower part of the *E* layer, where the conditions $\nu_{\mathrm{en,in}} > \omega_{He,i}$ are fulfilled, it can be assumed that the ionospheric plasma is completely dragged by neutrals and $\vec{u}_{\mathrm{i}} \approx \vec{u}_{\mathrm{e}} \approx \vec{u}_{\mathrm{n}}$. The time τ of onset of forced distributions of electrons and ions can be estimated from the formula

$$\tau = \frac{L^2}{D} \ , \quad \text{where} \quad D = \frac{2\kappa T}{m_{\mathrm{i}}\nu_{\mathrm{in}} + m_{\mathrm{e}}\nu_{\mathrm{en}}}$$

is the diffusion coefficient, κ is Boltzmann's constant, *T* is the absolute temperature, and $m_{\mathrm{e,i}}$ are the electron and ion masses, respectively. A rough estimate shows that for the altitudes of the *D* and lower *E* layers, the time τ is of the order of a few seconds, which is much less than the period of the considered atmospheric gas oscillations. This means that the disturbances of the ionospheric plasma density in our case have temporal and horizontal spatial scales corresponding to the disturbance of neutral gas. Considering that in the linear approximation the disturbances of the electrons and ions densities are small and assuming that velocities of the charged particles are close to $v_{\sim}$, Eq. 5) leads to the following formula

$$N = \frac{k_{\perp}V}{\omega}\sqrt{\frac{\rho_E}{\rho}}N_0 = \frac{P}{c_s^2}\frac{N_0}{\sqrt{\rho\rho_E}} \ , \tag{6}$$

where N_0 is the basic state of the ionospheric plasma density. In the derivation of the Eq. 6 it was taken into account that the vertical velocity disturbances of the atmospheric gas, which are related to the effect considered in the previous sections, is absent in the resonance region, and the equality $\omega = c_s k_{\perp}$ is justified. Since the disturbance of ionospheric plasma density is proportional to the pressure disturbance of the neutral gas, it should be expected that thin (several ten meters) and extended (of the order of the acoustic gravity wave horizontal scale determined by the horizontal scales of the source on the Earth's surface) ionospheric irregularities with a periodically varied plasma density will be generated in the ionospheric *D* layer with

a weak gradient of the background density in the resonance level. The impact of this resonance effect on the plasma at the altitudes of the ionospheric E layer, where $\nu_{in} \sim \omega_{Hi}, \nu_{en} << \omega_{He}$ (electrons are strongly magnetized) and there are conditions for the sporadic E layers generation, is more difficult for analysis. This is due to the fact that the electron density gradient should be taken into account and the Whitehead force, which leads to a still greater reduction of the domain thickness, should play an important role. In this case the formation of finite mass on the resonance level (Fig. 2) must lead to the formation of narrow (in vertical directive) horizontal ionospheric irregularities. Irregularities of such a type can be observed on the ionograms in the form of mildly sloping weakly diffuse sporadic layers with a large range of translucency (Fatkullin *et al.* 1985).

5. SUMMARY

Local acoustic-gravity disturbance in the nonisothermal atmosphere has been studied. According to the analytical results, the pressure wave amplitude has a local wave singularity near the layer at which the horizontal phase velocity is equal to the sound velocity. It is shown for the real altitude temperature profile in the atmosphere that near this layer the wave pressure component has a singularity, and the vertical velocity in the disturbance becomes zero.

In real conditions, many resonance layers (for different ω and $k_\perp$) can exist. The fact is that the atmospheric viscosity and nonlinearity limits the pressure wave singularity. We discussed the case when the viscosity influence on the density spike is more important. For such a situation, the vertical dimensions of singularity domain will be of the order of the mean free path of the molecules at the resonance level.

Estimates show that at altitudes of about one hundred kilometers the viscosity limits the scales to several hundreds of meters. Ionospheric plasma at these altitudes and below behaves as a passive impurity and therefore ionospheric irregularities with the same characteristic features as the neutral gas disturbances should be observed, including the case where the condition of the resonance feature generation is fulfilled. It is shown in paper Erukhimov and Savina (1980) that irregularities such a structure are the main reason for the formation of weakly diffuse sporadic E layers with a large range of translucence. In the framework of the model in question, the time of existence, altitude, and space scales of such layers are fundamentally dependent on the parameters of the ground-based acoustic gravity waves sources and ionospheric conditions.

Acknowledgments. This work has been partly funded by research RFBR grant No. 14-05-00565. P.A. Bespalov thanks Program No. 22 of RAS.

References

Bespalov, P.A., and O.N. Savina (2012), Possibility of magnetospheric VLF response to atmospheric infrasonic waves, *Earth Planets Space*, **64**, 6, 451-458, DOI: 10.5047/eps.2011.05.024.

Blanc, E. (1985), Observations in the upper atmosphere of infrasonic waves from natural or artificial sources: A summary, *Ann. Geophys.* **3**, 6, 673-687.

Erukhimov, L.M., and O.N. Savina (1980), The role of small-scale irregularities in the formation of radio reflections from the mid-latitude sporadic E layer, *Ionosfer. Issled.* **30**, 80-86 (in Russian).

Fatkullin, M.N., K.N. Vasilev, T.I. Zelenova, and O.N. Savina (1985), The E-scattering phenomenon in the midlatitude ionosphere, *Geomagn. Aeron.* **25**, 388-393 (in Russian).

Francis, S.H. (1975), Global propagation of atmospheric gravity waves: A review, *J. Atmos. Terr. Phys.* **37**, 6-7, 1011-1054, DOI: 10.1016/0021-9169(75) 90012-4.

Gershman, B.N. (1974), *Dynamics of the Ionospheric Plasma*, Nauka, Moscow, 256 pp. (in Russian).

Gossard, E.E., and W.H. Hooke (1975), Waves in the Atmosphere. Atmospheric Infrasound and Gravity Waves: Their Generation and Propagation, Elsevier Sci. Publ. Co., Amsterdam, 456 pp.

Hedin, A.E. (1991), Extension of the MSIS thermosphere model into the middle and lower atmosphere, *J. Geophys. Res.* **96**, A2, 1159-1172, DOI: 10.1029/ 90JA02125.

Kikoin, I.K. (ed.) (1976), *Tables of Physical Quantities. Handbook*, Atomizdat, Moscow (in Russian).

Lighthill, J. (1978), *Waves in fluids*, Cambridge University Press, Cambridge.

Lund, T.S., and D.C. Fritts (2012), Numerical simulation of gravity wave breaking in the lower thermosphere, *J. Geophys. Res.* **117**, D21, 105, DOI: 10.1029/ 2012JD017536.

Rapoport, V.O., P.A. Bespalov, N.A. Mityakov, M. Parrot, and N.A. Ryzhov (2004), Feasibility study of ionospheric perturbations triggered by monochromatic infrasonic waves emitted with a ground-based experiment, *J. Atmos. Sol.-Terr. Phys.* **66**, 12, 1011-1017, DOI: 10.1016/j.jastp.2004.03.010.

Savina, O.N. (1996), Acoustic-gravity waves in an atmosphere with a realistic temperature distribution, *Geomagn. Aeron.* **36**, 218-224.

Savina, O.N., and P.A. Bespalov (2014), Filtering features of long acoustic-gravity waves in a windless atmosphere, *Radiophys. Quantum Electr.* **57**, 2, 117-124, DOI: 10.1007/s11141-014-9497-6.

Received 16 December 2013
Received in revised form 11 July, 2014
Accepted 14 July 2014