

Volume 65 • Number 1 • March 2017

Acta Geophysica

PAN
POLISH ACADEMY OF SCIENCES



Institute of Geophysics
Polish Academy of Sciences



Springer

Centimeter-level precise orbit determination for the HY-2A satellite using DORIS and SLR tracking data

Qiaoli Kong^{1,2,3} · Jinyun Guo^{1,2} · Yu Sun⁴ · Chunmei Zhao⁵ · Chuanfa Chen^{1,2}

Received: 27 December 2016 / Accepted: 29 December 2016 / Published online: 19 January 2017
© Institute of Geophysics, Polish Academy of Sciences & Polish Academy of Sciences 2017

Abstract The HY-2A satellite is the first ocean dynamic environment monitoring satellite of China. Centimeter-level radial accuracy is a fundamental requirement for its scientific research and applications. To achieve this goal, we designed the strategies of precise orbit determination (POD) in detail. To achieve the relative optimal orbit for HY-2A, we carried out POD using DORIS-only, SLR-only, and DORIS + SLR tracking data, respectively. POD tests demonstrated that the consistency level of DORIS-only and SLR-only orbits with respect to the CNES orbits were about 1.81 cm and 3.34 cm in radial direction in the dynamic sense, respectively. We designed 6 cases of different weight combinations for DORIS and SLR data, and found that the optimal relative weight group was 0.2 mm/s for DORIS and 15.0 cm for SLR, and RMS of orbit differences with respect to the CNES orbits in radial direction and three-dimensional (3D) were 1.37 cm and 5.87 cm,

respectively. These tests indicated that the relative radial and 3D accuracies computed using DORIS + SLR data with the optimal relative weight set were obviously higher than those computed using DORIS-only and SLR-only data, and satisfied the requirement of designed precision. The POD for HY-2A will provide the invaluable experience for the following HY-2B, HY-2C, and HY-2D satellites.

Keywords HY-2A satellite · Precise orbit determination · DORIS · SLR · Weight

Introduction

HY-2A was launched on August 16, 2011, and as the first mission of the earth observation satellite HY-2 series (including HY-2A, HY-2B, HY-2C, and HY-2D) focuses on monitoring the marine dynamic environment with microwave sensors to detect sea surface wind field, height, and temperature. It can provide important measures to address climate change. HY-2B has just been approved and will be launched at the end of 2016 and HY-2C will be in 2019 on schedule. Therefore, the precise orbit determination (POD) for HY-2A will provide invaluable experiences for the subsequent satellites. To achieve the orbit radial accuracy up to the centimeter level, the HY-2A satellite loads a Doppler Orbit Determination and Radio positioning Integrated by Satellite (DORIS) receiver, a dual-frequency Global Positioning System (GPS) receiver, and a Satellite Laser Ranging (SLR) retro-reflector array (Zhao et al. 2013; Guo et al. 2013a; Kong et al. 2014).

DORIS is an excellent satellite tracking system supporting precise orbit determination of satellites (Tapley et al. 1994; Tavernier et al. 2006; Zelensky et al. 2010).

✉ Qiaoli Kong
kqlabc3334@hotmail.com

✉ Jinyun Guo
jinyunguo1@126.com

- ¹ College of Geomatics, Shandong University of Science and Technology, Qingdao, China
- ² State Key Laboratory of Mining Disaster Prevention and Control Co-founded by Shandong Province and the Ministry of Science and Technology, Shandong University of Science and Technology, Qingdao, China
- ³ Engineering Laboratory of Spatial Information Technology of Highway Geological Disaster Early Warning in Hunan Province, Changsha University of Science & Technology, Changsha, China
- ⁴ Department of Geoscience and Remote Sensing, Delft University of Technology, Delft 2628CN, The Netherlands
- ⁵ Chinese Academy of Surveying and Mapping, Beijing, China

The DORIS system is jointly developed and deployed by the Centre National d'Études Spatiales (CNES), Institut national de l'information géographique et forestière (IGN), and Groupe de Recherche de Géodésie Spatiale (GRGS), France. It is based on the one-way measurement of Doppler frequency shift. The selection of a one-way system (uplink) allows fully automated operation of beacons and easy communication within the whole system. The signal is transmitted by ground beacons and received by the onboard DORIS package when the satellite transits over the sky above the ground beacons (Seeber 1993). The data are processed on the ground to support the centimeter-level orbit determination for the satellite. DORIS data are also processed on board to provide real-time satellite positioning with an accuracy of some tens of centimeters. DORIS was originally designed to perform the precise orbit determination especially for TOPEX/Poseidon, and the high accuracy orbit derived from this technique played a critical role in ocean altimeter experiments (Choi 2003; Jiang et al. 2014) with an accuracy of 10 cm (Willis 1995). The TOPEX/Poseidon satellite was equipped with the first generation DORIS receiver, and achieved 13 cm radial orbit accuracy for POD (Tapley et al. 1994). With the first flight of the DGXX generation receiver on-board Jason-2, DORIS entered a new era of POD capability (Auriol and Tourain 2010; Zelensky et al. 2010). Zelensky et al. (2010) achieved an average fit value of 1.02 cm in radial direction for Jason-2 using DORIS-only data. The current DGXX generation DORIS receiver onboard HY-2A can simultaneously track 7 DORIS beacons with 2×7 channels (Mercier et al. 2010). This capability can increase the observation quantity, improve the coverage of observations, enhance the geometric strength, and hence improve the accuracy of orbit determination. All of these advantages will help the POD community to achieve the centimeter-level radial accuracy goal.

As a high precise tracking and positioning space geodetic technique, the SLR system is based on a two-way range principle and can provide the absolute range from tracking station to satellite. The international laser ranging service (ILRS) is responsible for the simultaneous observation and scientific product delivery of all satellites equipped with the laser retro-reflector array (LRA) (Pearlman et al. 2002; Gurtner et al. 2005). For nearly four decades, SLR has provided the primary tracking data for numerous geodetic missions such as LAGEOS, TOPEX/Poseidon, Jason, Envisat, HY-2A, and so on. On the panel of HY-2A, a dedicated LRA is equipped facing the Earth. LRA onboard HY-2A is similar to those on Envisat, ERS-2, and Jason-1/2, and plays an important role in proving a calibration tool and improving the POD solution (Luthcke et al. 2002, 2003; Urschl et al. 2005; Zhao et al. 2013; Kong et al. 2014).

If SLR data are combined with the other independent space geodetic data, the geometric structures will be strengthened, the solution will be more stable than those using the single geodetic technique, and the systematic error will be compensated to some extent. Thus, the accuracy of POD will be significantly improved (Zhang et al. 2000). Recent researches focus on using multi-geodetic techniques to estimate satellite orbit (e.g., Choi 2003; Lemoine et al. 2010; Zelensky et al. 2010; Cerri et al. 2010; Melachroinos et al. 2011; Peng et al. 2012; Thaller et al. 2013; Sheng et al. 2014; Hackel et al. 2015; Sosnica et al. 2015). Recent orbit accuracy based on DORIS + SLR data can reach about 2 cm in the radial direction thanks to the development of tracking instruments and the gradual perfection of the physics-based force models. Some of researchers have achieved 1 cm using SLR/DORIS in the radial direction (e.g., Choi 2003; Lemoine et al. 2010; Zelensky et al. 2010; Cerri et al. 2010; Couhert et al. 2015). Radial orbit accuracies for ERS-1 and ERS-2 have reached 2–3 cm (Rudenko et al. 2012). 6 cm of 3D RMS (root mean square error) was achieved for Jason-2 using DORIS + SLR data (Peng et al. 2012). Accuracy better than 2 cm in the radial direction was obtained for Jason-2 using combined GPS, SLR, and DORIS data (Sheng et al. 2014). One decimeter orbit accuracy was achieved for the first two Galileo In-Orbit Validation (IOV) satellites using combined GNSS + SLR observations (Hackel et al. 2015).

To achieve the precise orbit determination for HY-2A, a lot of work has been carried out. Zhao et al. (2013) and Wang et al. (2014) showed that the radial accuracy can be better than 3 cm using SLR-only data. Zhu et al. (2013) showed the radial accuracy of about 2 cm using DORIS-only data. Lin et al. (2014) and Guo et al. (2013b) demonstrated that the radial accuracy better than 2 cm can be achieved using GPS-only data. Jiang et al. (2014) have achieved 4.40 cm radial accuracy using combined DORIS and SLR data.

The precise orbit of HY-2A can provide a reference frame for the oceanic science research of this mission. To achieve the centimeter-level radial accuracy, independent DORIS and SLR data are selected to determine the HY-2A satellite orbit with the dynamic method. Section 2 of this paper addresses the factors related to HY-2A; Sect. 3 analyzes the performance of DORIS and SLR; Sect. 4 describes the POD strategies; Sect. 5 focuses on analyzing the RMS of orbit differences between the achieved orbits using DORIS-only, SLR-only, and DORIS + SLR data with respect to the precise CNES orbits, and compares the orbit difference RMS between DORIS-only, SLR-only, and DORIS + SLR data. Finally, the optimal relative weights for DORIS and SLR data are recommended for the optimal solution.

The HY-A satellite description

The HY-2A mission was approved jointly by the Ministry of Finance and State Administration of Science, Technology and Industry for National Defense of PRC in January, 2007. This satellite is the first ocean dynamic environment monitoring satellite sponsored by the State Oceanic Administration of China (SOA), and is the third Chinese ocean satellite after HY-1A and HY-1B. The HY-2A mission is one of the key civil space projects in the 11th Five-Year Plan of China. It runs in a near sun-synchronous orbit with a lifetime of more than 3 years. It has an orbital inclination of 99.3° , altitude of 971 km, and an equator crossing time on descending node at 6:00 a.m. In the first stage, its repeat cycle is 14 days with an intersection period of 104.46 min and it runs $13 + 11/14$ revolutions every day. In the second stage, it will cover $13 + 131/168$ revolutions every day with the repeat cycle of 168 days. The main remote sensor payloads on it include one radar altimeter, one microwave scatterometer, one microwave scanning radiometer, one microwave calibrating radiometer, etc.

The attitude error of a satellite can degrade the accuracy of the computed orbit to a great extent (Tseng et al. 2012) because attitude models can not exactly describe the observed attitude in real time and cannot reflect the time lag (Sheng et al. 2014). Normally, attitudes should be taken into consideration in POD strategy. Usually, adjustment of attitude angles of roll, pitch, and yaw is required to compensate the impact of attitude maneuver. For HY-2A satellite, the attitude controller is stable inertial wheels, which can make three-axis pose stability lower than 0.003 %, pointing precision lower than 0.1° , and measurement precision lower than 0.05° in pitch, roll, and yaw in local orbital frame (Kong et al. 2014). The satellite is three-axis stabilized, and is non-spinning relative to the satellite reference frame.

During the operational stage, the satellite mass and the mass center change during the HY-2A maneuver. The detailed information about mass and mass center is available from IDS.¹

HY-2A is equipped with DORIS and LRA as shown in Fig. 1. The onboard DORIS receiver has two frequencies (2 GHz and 400 MHz), and both phase centers of DORIS are listed in Table 1. Accuracies of positions of both phase centers are better than 2 mm. The DORIS antenna axis nominally points to the geocenter. The HY-2A satellite is the second Chinese satellite equipped with LRA following the Compass M1 (Wu et al. 2011). LRA on HY-2A is composed of nine corner cubes made of fused silica, and cubes are symmetrically mounted on a hemispherical surface with one nadir-looking corner cube in the center. The radius and height of each cube are 16.5 and 26.2 mm,

respectively (Zhao et al. 2013). The angle between the normal of the center reflector and the side ones is 48° . The size of LRA is 250 mm \times 88.5 mm and its weight is about 1.41 kg. The phase center and spherical center are listed in Table 1. The range correction of LRA from the spherical center is 0.074 m (Guo et al. 2013b).

In the process of POD, most of satellite parameters come from the file written by CNES. The non-conservative forces like the solar and Earth radiation pressure and atmospheric drag must be taken into account, and these force models have a close relation with the complex geometry and material properties of the spacecraft. The spacecraft is usually treated as a combination of flat plates arranged in the shape of a box, and the disturbances of solar arrays and atmospheric drag can be conveniently calculated. In this paper, HY-2A flat is designed as 13 flat plates. Besides the specular reflectivity and diffuse reflectivity, the properties of emissivity and cold equilibrium are all taken into consideration. We used a more refined macro-model than that on line.² All specular reflectivity coefficients of 13 flat plates are zero, and the optical and infrared properties of these panels are listed in Table 2.

Doris and SLR tracking data performance

In this paper, we applied DORIS range rate data and SLR normal point data from September 8, 2012 to December 21, 2012 to compute the precise orbits of HY-2A. During these 105 days, HY-2A was tracked by 50 DORIS beacon stations and 25 SLR stations. 50 DORIS beacon stations are evenly distributed over the globe, whereas most SLR tracking sites are concentrated in Europe and North America, with a few stations in Asia and the southern hemisphere. The DORIS system has no weather restriction and beacons are designed to operate for a long time with little human intervention. Therefore, sufficient observations are available from the HY-2A satellite. However, only 3 DORIS beacon stations have tracked HY-2A on October 26, 2012. We have not computed the orbits for HY-2A on this day using DORIS-only and DORIS + SLR data.

For the low Earth orbit satellite, the SLR tracking normal point data are rare. Gaps of several hours occur very frequently because of bad weather conditions near SLR stations. During these 105 days, 25 SLR stations have tracked this mission. There are 11 days in which the number of SLR passes per day is less than 6, and we have not solved the orbits in these days for HY-2A using SLR-only data. The SLR data used in the process of POD are

¹ <ftp://ftp.ids-doris.org/pub/ids/satellites/h2amass.txt>.

² ftp://cddis.gsfc.nasa.gov/doris/cb_mirror/satellites/DORISatelliteModels.pdf.

Fig. 1 HY-2A scheme for DORIS receiver and LRA

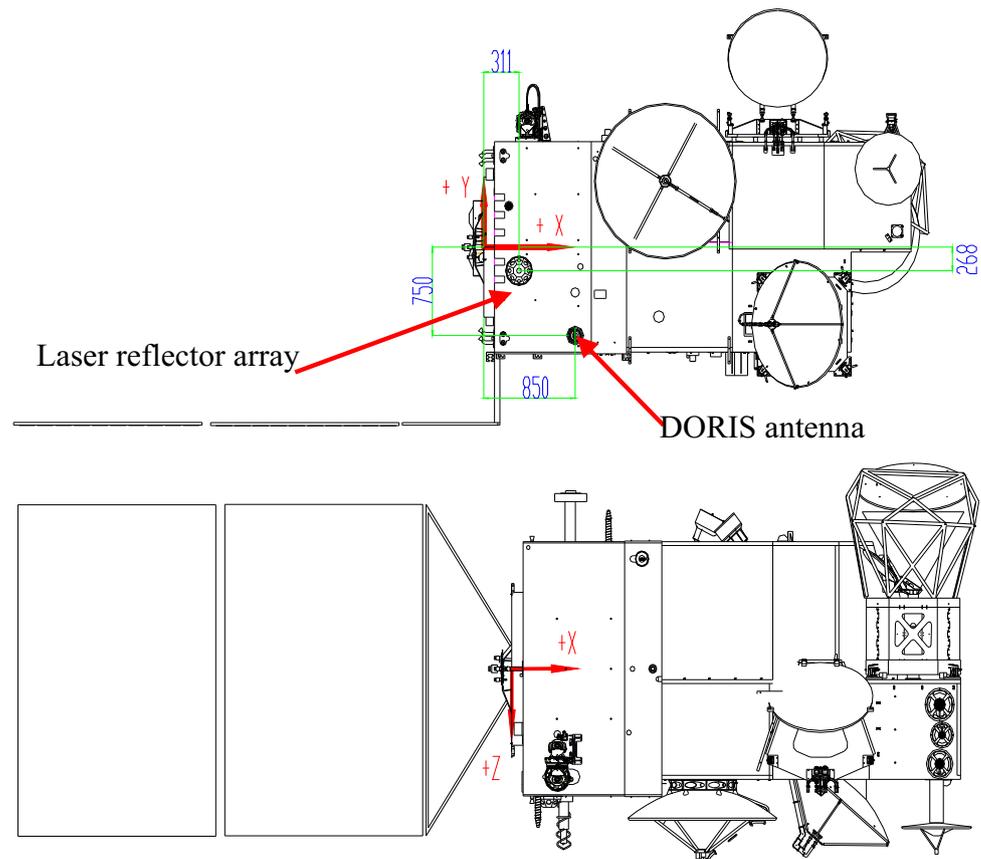


Table 1 Positions of LRA optical center, DORIS and LRA phase centers in the satellite reference frame (m)

| Payload | X | Y | Z |
|-----------------------------|-------|--------|-------|
| DORIS receiver phase center | | | |
| 2 GHz | 0.850 | -0.750 | 1.306 |
| 400 MHz | 0.850 | -0.750 | 1.144 |
| LRA phase center | 0.311 | -0.268 | 1.068 |
| LRA spherical center | 0.311 | -0.215 | 0.984 |

normal point data compressed from original tracking data for HY-2A.

HY-2A pod strategy using Doris and SLR data

POD was carried out using 10 s integrated Doppler data and the normal point data of SLR from September 8, 2012 to December 21, 2012. The accuracy of orbit determination is affected by several factors, such as tracking station distribution, quantity and quality of precise tracking data, and the accuracy of available forces and measurement models (e.g., gravity field, atmospheric drag, solar radiation pressure, phase center corrections, station coordinates,

etc.) (Tapley et al. 1994). The even beacon distribution and the high precise observations of DORIS provide the possibility of POD. Considering the quantity of observation and the strength of the geometric structure of tracking stations (Švehla and Rothacher 2003; Kong et al. 2014), the dynamic method was selected for POD of HY-2A. Dynamic orbit determination is a very precise method, and the accuracy of the derived orbit for HY-2A strongly depends on the accuracy of force models. The exact modeling of ocean tides of the Earth is a difficult problem because of the complex hydrodynamic response to the tidal forces. Approximately the same as the solid Earth tides, the dynamical effect of ocean tides is most easily described as periodic variations in the normalized geopotential coefficients. Both solar radiation and Earth albedo radiation have significant influence on the low Earth orbit satellite and neither of them can be neglected (Zelensky et al. 2010). For DORIS and SLR data, the tropospheric refraction corrections were computed using the modified Hopfield model (Goad and Goodman 1974) and the Mendes-Pavlis model (Mendes et al. 2002; Mendes and Pavlis 2004), respectively. Remaining errors of the non-gravitational force models can cause secular and long-term changes in the orbital elements, which can be effectively compensated with the RTN empirical force model (Doornbos et al. 2002;

Table 2 The optical and infrared properties of the macro-model and the plate surfaces

| Surface (m ²) | Normal in satellite reference frame | | | Optical properties | | Infrared properties | |
|---------------------------|-------------------------------------|----|----|--------------------|------------|---------------------|------------|
| | X | Y | Z | Diffuse | Emissivity | Diffuse | Emissivity |
| 2.50 | 1 | | | 0.54 | 0.46 | 0.31 | 0.69 |
| 2.92 | −1 | | | 0.54 | 0.46 | 0.31 | 0.69 |
| 5.85 | | 1 | | 0.54 | 0.46 | 0.31 | 0.69 |
| 6.74 | | −1 | | 0.54 | 0.46 | 0.31 | 0.69 |
| 4.93 | | | 1 | 0.54 | 0.46 | 0.31 | 0.69 |
| 4.60 | | | −1 | 0.54 | 0.46 | 0.31 | 0.69 |
| 9.06 | | −1 | | 0.36 | 0.64 | 0.16 | 0.84 |
| 9.06 | | 1 | | 0.06 | 0.94 | 0.06 | 0.94 |
| 0.71 | 1 | | | 0.85 | 0.15 | 0.21 | 0.79 |
| 0.60 | 1 | | | 0.85 | 0.15 | 0.21 | 0.79 |
| 0.89 | | 1 | | 0.73 | 0.27 | 0.13 | 0.87 |
| 1.50 | | | 1 | 0.85 | 0.15 | 0.21 | 0.79 |
| 1.80 | | | 1 | 0.85 | 0.15 | 0.21 | 0.79 |

Kang et al. 2006; Liu 2013; Kong et al. 2014). Cowell II numerical integration (Balmino and Barriot 1989) was used during the HY-2A satellite orbit determination, and the integration step size was 20 s in Geodyn II software. The output interval of orbit ephemeris was 60 s.

In the process of POD, all unknown parameters were estimated simultaneously, including the coordinates, velocities, atmospheric drag coefficients, solar radiation coefficients, and empirical acceleration coefficients. The RTN (R: Radial direction; T: Along-track direction; N: Cross-track direction) empirical force model was used (Colombo 1989), and coefficients of one-cycle-per-revolution empirical accelerations were estimated every day in the cross-track. Bias, cosine, and sine terms were computed in the cross-track direction. The initial state vectors were estimated per orbital arc. The MSIS86 model (Hedin 1987) was used and the coefficients C_d were estimated per 6 h. The Box-wing solar radiation pressure model was applied and the coefficients were estimated per 6 h. In the iterative computation, the least square estimation method was applied to estimate the initial parameters of position/velocity of this satellite. The computation continued until the residuals between the derived results of two adjacent iterations were less than the designed tolerance values (Tapley et al. 1994, Kong et al. 2014).

10° and 15° were the cutoff elevation angles of DORIS and SLR data, respectively. During the process of POD, UTC time was used. SLR tracking data time is UTC time, and DORIS tracking data time is TAI time. The constant offset (leap second) of 35 s was introduced when DORIS data were applied.

DORIS format 2.2 range rate data were used to compute orbit for HY-2A. In the observation file, the data having preprocessing indicator marked as “0” are considered to be good data, “1” represents data edited during reprocessing,

“2” represents the data edited during post-processing, “3” indicates that the data are possibly erroneous, and “4” means that the beacon is a 3.0 beacon in the restart mode. During the process of POD, all data indicated as “0” were accepted and the data with other indicators were deleted (Liu 2013). All of the SLR data were used for POD. If the residuals were 3 times larger than sigma, the corresponding data would be deleted. If not, their big errors would pollute the combined database and degrade the accuracy of POD. The DORIS and SLR observations and perturbation models used in the orbit determination are listed in Table 3.

Usually, for the missions equipped with a DORIS receiver, the computed orbit solutions would be compared with the CNES orbits,³ which were solved using combined DORIS, GPS, and SLR data (Peng et al. 2012; Cerri et al. 2013; Couhert et al. 2015; Gao et al. 2015) and achieved about 1.1 cm radial accuracy (Gao et al. 2015; Zhang and Vincent 2011) for HY-2A satellite. The precise CNES orbit and the one we derived may have used different models, and we can examine the validity of the models we applied by comparing our orbits with the CNES ones. The disadvantage is that we can achieve the relative accuracy and check the orbit consistency with respect to the CNES orbit.

POD using DORIS and SLR data for HY-2A

POD using DORIS-only data for HY-2A

Willis et al. (2013) have computed a set of DORIS beacon coordinates, DPOD2008 as an extension of ITRF2008 for POD of altimetry missions. We applied and fixed these coordinates to compute HY-2A satellite orbits.

³ <ftp://cddis.gsfc.nasa.gov/pub/doris/products/orbits/ssa/h2a/>.

Table 3 Adopted dynamic models and measurements in orbit determination for HY-2A

| Items | Description |
|---|--|
| The coordinates of DORIS beacon stations and DORIS measurements | (Willis et al. 2012, 2013); DORIS 2.2 |
| The coordinates of SLR tracking stations and SLR data | ILRS; SLR normal point data |
| Earth gravity model | EGM2008 (Pavlis et al. 2012), 80×80 |
| Planetary N-Body | JPL DE403 (Standish 1998) |
| Solid Earth tides | IERS2010 (Petit and Luzum 2010) |
| Ocean tides and ocean tides loading | FES2004 (Lyard et al. 2006) |
| Relativistic correction | IERS2003 (McCarthy and Petit 2033) |
| Solar radiation pressure | Box-Wing (Rim 1992) |
| Earth albedo radiation | Knocke–Ries–Tapley (Knocke et al. 1988) |
| Atmospheric drag | MSIS86 (Hedin 1987) |
| Tropospheric model | DORIS: Hopfield (Hopfield 1971; Goad and Goodman 1974); SLR: Mendes-Pavlis (Mendes et al. 2002; Mendes and Pavlis 2004) |
| Satellite attitude | Nominal |
| Cut off angle | DORIS: 10° ; SLR: 15° |

Measurement biases and tropospheric refraction scale bias were computed as the unknown parameters for each tracked DORIS beacon. We set the cutoff elevation angle as 10° . A uniquely valuable measurement of orbit quality is to compare the derived orbits with the orbits produced by the other organization. In this section, we computed the orbits of HY-2A using 105-day DORIS-only data, and Fig. 2 shows the daily orbit difference RMS with respect to CNES results.

On October 26, 2012, only 3 beacon stations tracked this mission, and the orbits could not be estimated, so there is a breakpoint in Fig. 2. The statistics of orbit difference RMS derived from 104 days are listed in Table 4.

Figure 2 and Table 4 show that the radial direction has the smallest RMS of orbit difference with respect to the CNES orbits among these three directions, and the RMS in this direction is 1.81 cm, which meets the radar altimeter requirement for science research and applications. The difference RMS value of the cross-track direction is a little bigger than that of the radial direction, and the biggest difference RMS value is that of the along-track direction, and the RMS in this direction is 6.17 cm. This experiment indicates that high density of beacon stations and highly precise HY-2A DORIS tracking data guarantee DORIS the high POD capability. The test results suggest that DORIS-only POD achieved better than 2 cm level relative accuracy in the radial direction during the observation time, and the relative accuracy is higher than the result achieved by Zhu et al. (2013) using DORIS-only data in 30 days, and is a little lower than the result achieved by Gao et al. (2015) using DORIS-only data in 93 days, and both of which were compared with the CNES orbits.

POD using SLR-only data for HY-2A

To show the performance of SLR data in POD, we also computed the orbits for HY-2A using SLR-only data with the dynamic method. The cutoff angle was set as 15° , and Fig. 3 shows the daily orbit difference RMS with respect to CNES orbits.

There are 11 breakpoints in Fig. 3, and the main reasons are that the numbers of passes on these days are all less than 6, not enough to compute the orbits. The statistics of orbit difference RMS derived from 94 days are listed in Table 5.

Figure 3 and Table 5 indicate that the RMS of differences with respect to the CNES orbits is 3.34 cm in the radial direction, and the RMS in the cross-track direction is 5.69 cm, while in the along-track direction, the RMS is 10.62 cm. From these test results, we know that the highest orbit precision is in the radial direction, mainly because the SLR data measure the absolute range between tracking station and satellite. Generally, the 3D RMS can reach 11.96 cm in position. Though the precision achieved using SLR-only data is lower than that derived from DORIS data, SLR data is of capability in calibration (Kong et al. 2014), validation, and improvement of the DORIS-based POD solutions. We have achieved similar RMS difference level to the results achieved by Zhao et al. (2013) and Wang et al. (2014) in 30-day arcs with respect to CNES in radial direction.

POD using DORIS + SLR data for HY-2A

For the POD using combined DORIS and SLR data, the weights for each kind of data play a crucial role in the

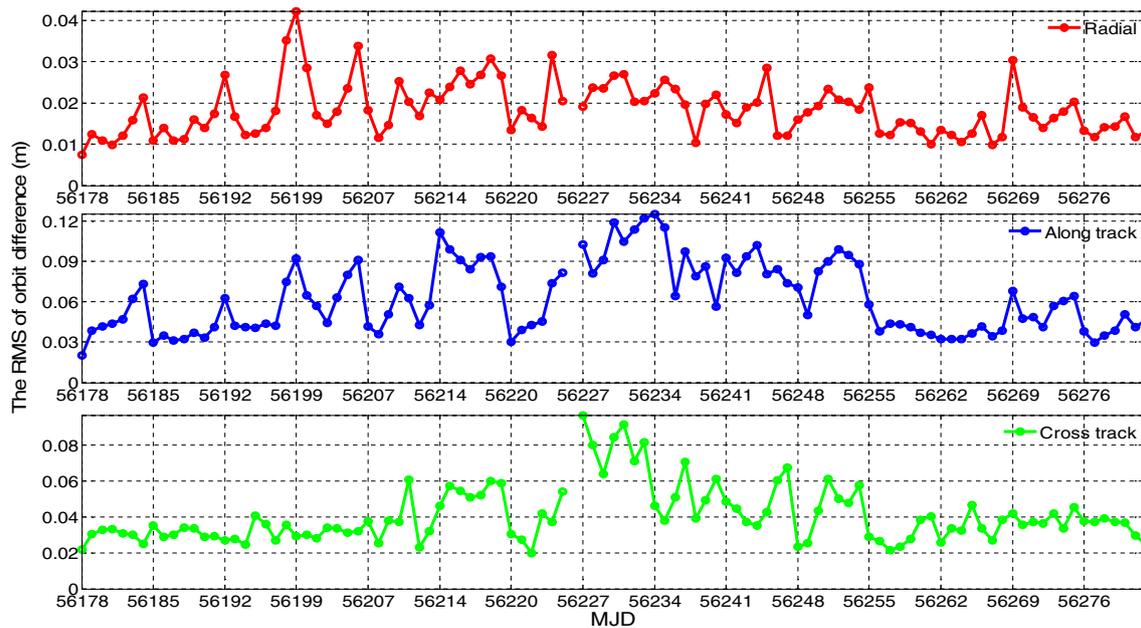


Fig. 2 Daily RMS of orbit difference with respect to CNES orbits in the 105 days using DORIS-only data

Table 4 Statistics of orbit difference RMS with respect to CNES orbits using DORIS-only data (cm)

| Direction | Max | Min | Mean |
|-----------|-------|------|------|
| R | 4.23 | 0.70 | 1.81 |
| T | 12.55 | 2.01 | 6.17 |
| N | 9.69 | 1.99 | 4.04 |
| 3D | 14.85 | 3.06 | 7.69 |

accuracy of POD. Choi (2003) selected 10 cm as the fixed weight for SLR data and 2 mm/s for DORIS data during the process of POD for Jason-1 and then combined with GPS data. Zelensky et al. (2010) adopted 10 cm and 3 mm/s as the weights of SLR and DORIS for Jason-1 and Jason-2 and achieved 1 cm radial accuracy. With the improvement of accuracy of SLR tracking data and DORIS receivers, Peng et al. (2012) selected 5 cm and 0.50 mm/s as weights for SLR and DORIS data to achieve POD for Jason-2 and 1.55 cm radial accuracy has been achieved. Doornbos et al. (2002) adopted 0.45 and 0.55 mm/s as the DORIS data weights and 3 cm as the SLR data weights for Jason-1 and Envisat, respectively.

A realistic way of seeking the optimal relative weights is to set the weights according to the RMS of post-fit residuals of different data types. The RMS of post-fit residuals for DORIS data is about 0.31 mm/s, and that for SLR-only data are about 3.2 cm. These data can serve as the references for the prior standard deviations for these two types of data. The sigmas should be adjusted to determine the best orbit solution, because the amount of DORIS data is large and the DORIS beacon stations are distributed evenly

over the global. Conversely, the amount of SLR data is small and the tracking stations are distributed unevenly. The a priori sigma for DORIS is set around the observation level while that for SLR is set about 3–5 times larger than the post-fit residual RMS of SLR data (Choi 2003).

To test the performance of different weight settings for POD, we designed six candidate combination strategies for weights in the light of earlier studies. In the 105-day orbit arc, 5, 10, and 15 were selected as weights for SLR, and 0.2 and 0.3 mm/s were selected as weights for DORIS. During the process of POD for HY-2A, orbits were computed using the dynamic models and data in Table 4. Again, the dynamic method was applied and the parameters were estimated as the same ones as DORIS-only and SLR-only. We plotted the orbit difference RMS with respect to CNES results in Fig. 4. Table 6 shows the statistics of mean orbit difference RMS with respect to CNES orbits with different weights.

There is a breakpoint in Fig. 4 because we did not compute the orbit on 26 October, 2012 in convenience of comparison between orbits derived using DORIS-only and DORIS + SLR data, and the statistics of the mean orbit difference RMS derived from 104 days are listed in Table 6.

From Fig. 4 we can notice that there are large residuals between 56,208 and 56,255 MJD. According to Mansoori et al. (2015) and Li et al. (2016), there are two geomagnetic storms on days 56208 and 56244. Both geomagnetic storms can produce positive significant effects on the ionospheric delay during the disturbed geomagnetic conditions. So the large residuals between days 56,208 and 56,255 have the close relation to higher ionosphere activity.

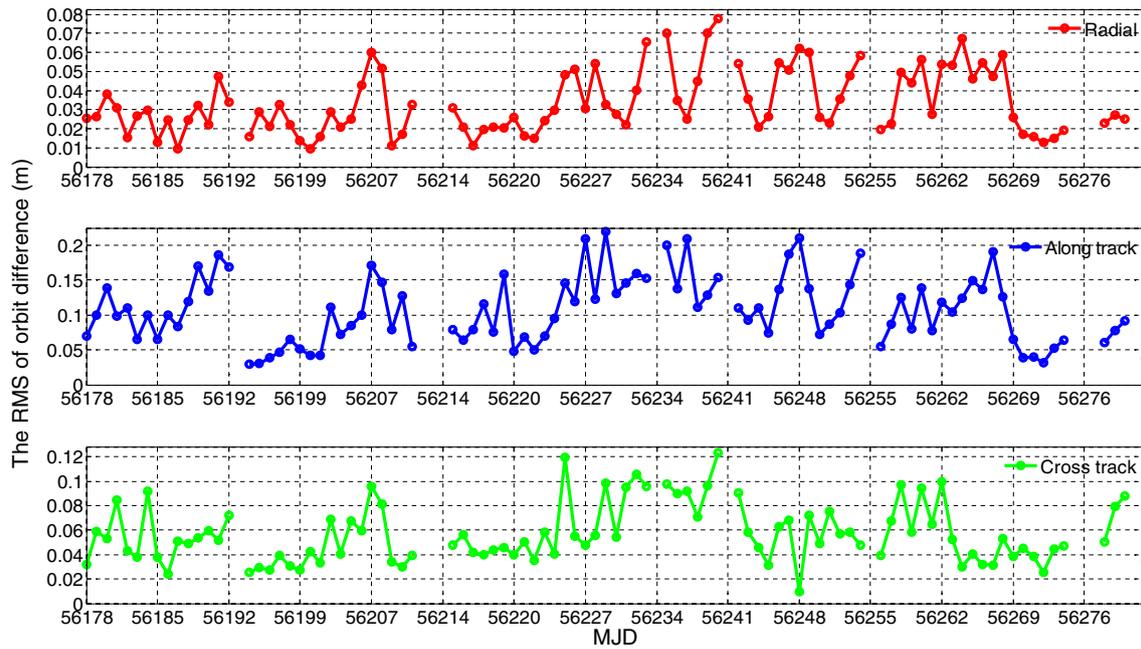


Fig. 3 Daily RMS of orbit difference with respect to CNES orbits in the 105 days using SLR-only data

Table 5 Statistics of orbit difference RMS with respect to CNES orbits using SLR-only data (cm)

| Direction | Max | Min | Mean |
|-----------|-------|------|-------|
| R | 7.76 | 0.93 | 3.34 |
| T | 22.00 | 2.97 | 10.62 |
| N | 12.34 | 0.94 | 5.69 |
| 3D | 24.32 | 4.21 | 11.96 |

Figure 4 and Table 6 demonstrate that the RMS of orbit difference in the radial direction is the smallest among three directions. From Table 6, we can know that giving SLR the fixed weight, the values of RMS difference with respect to CNES orbits in radial direction will increase if we decrease the weight of DORIS, conversely, the RMS values will be decreased with increasing the weight of DORIS data. At the same time, we can see that the mean agreement between the computed orbits and the CNES orbits in the radial direction is about 1.5 cm with every weight setting. The smallest RMS of the computed orbit in the radial direction is 1.37 cm, and the relative weights of DORIS and SLR are 0.2 mm/s and 15.0 cm, respectively. The smallest 3D RMS is 5.87 cm and the relative weights are 0.2 mm/s and 15.0 cm for DORIS and SLR, respectively. Therefore, the optimal relative weight group is 0.2 mm/s and 15.0 cm for DORIS and SLR, respectively. We can also conclude that a large weight should be given to DORIS in the orbit determination with the dynamic strategy using SLR + DORIS data. Table 6 indicates that DORIS plays a dominant role in the precise orbit

determination. The SLR + DORIS solutions rely more heavily on the DORIS data than SLR data since DORIS beacon stations have better temporal and spatial coverage than SLR tracking stations and there are more DORIS data than SLR data. We have achieved higher relative radial accuracy than those achieved using combined DORIS and SLR data in 30 days by Jiang et al. (2014) which were also compared with the CNES orbit.

Comparison between orbit difference RMS using DORIS-only, SLR-only, and DORIS + SLR data

To show the excellent performance of POD using the combined DORIS + SLR data, we plotted the RMS for orbits derived from DORIS-only, SLR-only, and DORIS + SLR data with 6 weight decisions, as shown in Fig. 5. In this figure, we drew the RMS in 3D, radial, along-track and cross-track directions for convenient comparison with every case.

From Tables 4, 5 and Fig. 5, we can see that in the radial direction, the RMS difference between DORIS-only and SLR-only orbits can reach 1.53 cm. The RMS differences are 4.45 and 1.65 cm in the along-track and the cross-track direction, respectively, and the 3D RMS difference is 4.27 cm. The RMS difference of the derived orbit using DORIS-only data is smaller than those from SLR-only data with respect to the CNES orbit, and the RMS in the radial direction is better than 2.0 cm for the DORIS-only orbit which can meet the needs of the altimeter payload on HY-2A. Tables 4, 6 and Fig. 5

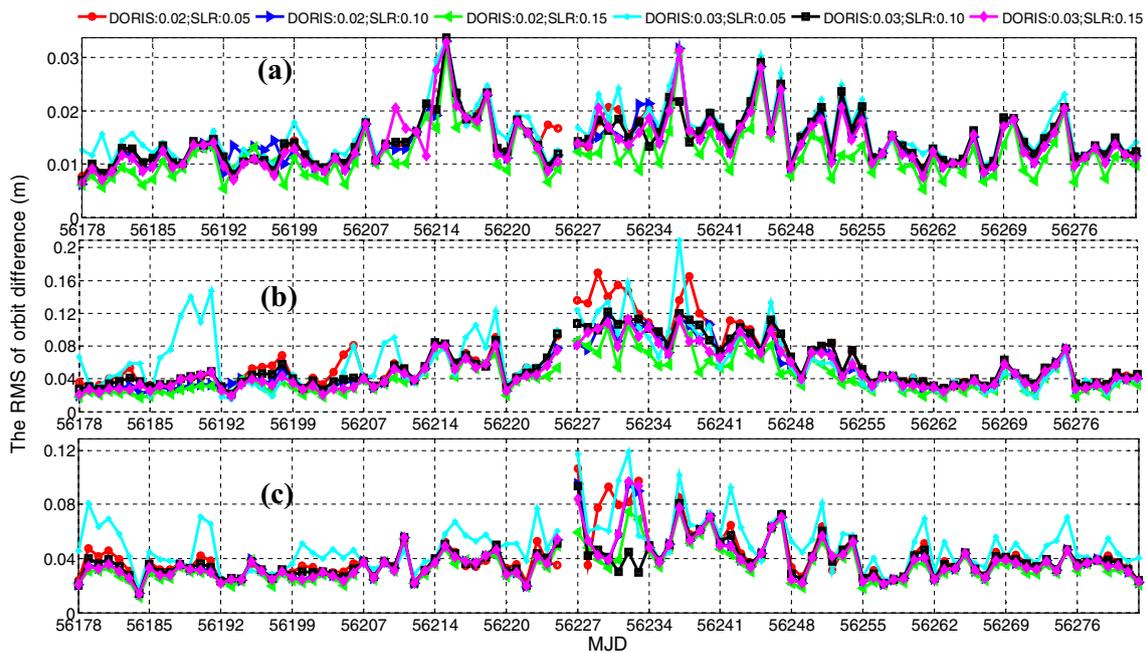


Fig. 4 Daily orbit difference RMS with respect to CNES orbits in the radial (a), along track (b) and cross-track (c) directions in the 105 days using DORIS + SLR data

Table 6 Statistics of the mean orbit difference RMS with respect to CNES orbits using DORIS + SLR data with different relative weights (cm)

| Weights setting | Radial | Along track | Cross track | 3D |
|---|--------|-------------|-------------|------|
| DORIS Weight: 0.2 mm/s; SLR weight: 5.0 cm | 1.57 | 6.34 | 3.69 | 7.35 |
| DORIS Weight: 0.3 mm/s; SLR weight: 5.0 cm | 1.60 | 7.76 | 4.51 | 9.21 |
| DORIS Weight: 0.2 mm/s; SLR weight: 10.0 cm | 1.56 | 5.48 | 3.45 | 6.89 |
| DORIS Weight: 0.3 mm/s; SLR weight: 10.0 cm | 1.49 | 5.63 | 3.56 | 6.95 |
| DORIS Weight: 0.2 mm/s; SLR weight: 15.0 cm | 1.37 | 4.94 | 3.41 | 5.87 |
| DORIS Weight: 0.3 mm/s; SLR weight: 15.0 cm | 1.45 | 5.38 | 3.52 | 6.74 |

indicate the RMS improvements of 4.4 and 18.2 mm from DORIS-only solution to DORIS + SLR solution with weights 0.2 mm/s and 15.0 cm for DORIS and SLR in the radial direction and 3D RMS. Therefore, this analysis demonstrates that SLR data can directly contribute to the overall POD accuracy, and the combined DORIS + SLR technique can produce more accurate solution than DORIS-only and SLR-only. This significant improvement of precision is owed to the high accurate observation of DORIS and SLR, which can significantly improve the data density and then significantly improve the POD capability.

Conclusions

HY-2A is the first ocean dynamic environment monitoring satellite of China following the Jason-1 and Jason-2 altimetry missions. A highly precise radial orbit is the fundamental requirement for accurate monitoring of sea surface topography with the altimetry satellite mission.

Since GPS data are unavailable on internet, we focused on POD using DORIS and SLR tracking data and the combination of these two space geodetic techniques.

To obtain the optimal solution, we estimated orbits using DORIS-only, SLR-only, and DORIS + SLR data, respectively, and analyzed the orbit difference RMS with respect to CNES orbits. We achieved 1.81 cm radial RMS difference and 7.69 cm 3D RMS difference with the dynamic method using DORIS-only data with respect to the CNES orbit, and 3.34 cm radial RMS difference and 11.96 cm 3D RMS difference using SLR-only data with respect to the CNES orbit, respectively. There are mainly three possible reasons for the smaller RMS of orbit difference in radial direction derived from DORIS-only data than those derived from SLR-only data with respect to the CNES orbit: first, the DORIS beacon stations are distributed more evenly than those of SLR tracking stations; second, there are more measurements of DORIS than SLR; finally there are more tracking stations for DORIS than SLR.

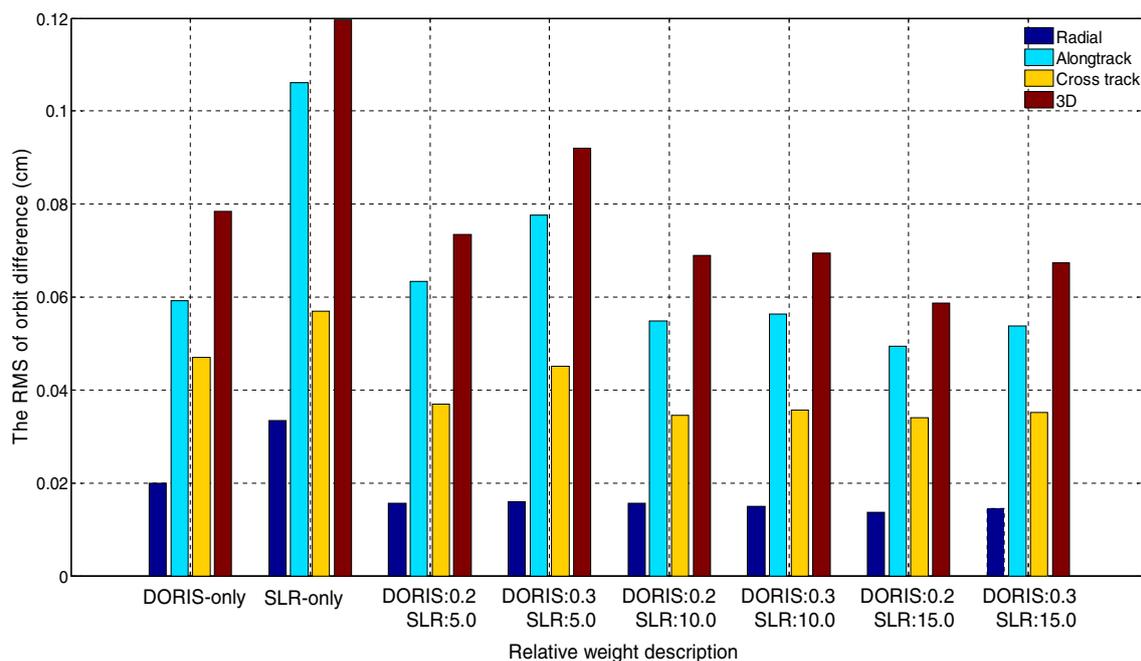


Fig. 5 RMS of the orbit differences with respect to CNES orbits based on DORIS-only, SLR-only and DORIS + SLR data (DORIS weight unit: mm/s, SLR weight unit: cm)

To take advantage of the high accuracy of SLR data, we have set 6 cases of weight groups for DORIS and SLR data, and found that the optimal relative weight group was 0.2 mm/s for DORIS and 15.0 cm for SLR, and 1.37 cm of radial RMS difference and 5.87 cm of 3D RMS difference could be achieved with respect to the CNES orbit. These experiments demonstrated that the orbits determined using DORIS-only, SLR-only, and DORIS + SLR data were all stable and reliable. We also found that DORIS played a dominant role and SLR provided a strong distance constraint in POD using DORIS + SLR data. Figures 4, 5 and Table 6 also indicated that the consistency to the CNES orbit in the radial direction was better than those in the other two directions. There are three possible reasons: first, the accuracy of the force model in the radial direction is more accurate than that in the other two directions; second, the range rate of DORIS Doppler measurements varies mainly in the radial direction; and third, SLR tracks the absolute range between tracking station and satellite.

It should be recognized that SLR data also played a great role in the improvement of the orbit determination, especially for the 3D position, because they enhanced the spatial and temporal distribution of tracking data. In addition, the special attention should be given to the weight setting for different space geodetic data to achieve highly precise orbit. POD from DORIS + SLR data requires further investigation in terms of measurement bias, troposphere bias, and DORIS network time bias. The fruitful

experience of POD for HY-2A will provide a reference for the following HY-2B, HY-2C, and HY-2D satellites.

Acknowledgements We thank anonymous reviewers for their helpful comments. We express our gratitude to CNES for providing HY-2A precise orbit and CDDIS for providing DORIS and SLR data. This work was supported by the National Natural Science Foundation of China (Nos. 41374009 & 41201381), the Public Benefit Scientific Research Project of China (No. 201412001), International Science and Technology Cooperation Program of China (No. 2009DFB00130), the Shandong Natural Science Foundation of China (No. ZR2013DM009), the Basic Science and Technology Research Project of China (Grant No. 2015FY310200), and Open Fund of Engineering Laboratory of Spatial Information Technology of Highway Geological Disaster Early Warning in Hunan Province (Changsha University of Science and Technology) (No. kfj150605).

References

- Auriol A, Tourain C (2010) DORIS system: the new age. *Adv Space Res* 46(12):1484–1496. doi:[10.1016/j.asr.2010.05.015](https://doi.org/10.1016/j.asr.2010.05.015)
- Cerri L, Berthias JP, Bertiger WI, Haines BJ, Lemoine FG, Mercier F, Ries JC, Willis P, Zelensky NP, Ziebart M (2010) Precision orbit determination standards for the Jason series of altimeter missions. *Mar. Geod.* 33(S1):379–418. doi:[10.1080/01490419.2010.488966](https://doi.org/10.1080/01490419.2010.488966)
- Cerri L, Lemoine JM, Mercier F (2013) DORIS-based point mascons for the long term stability of precise orbit solutions. *Adv Space Res* 52(3):466–476. doi:[10.1016/j.asr.2013.03.023](https://doi.org/10.1016/j.asr.2013.03.023)
- Choi KR (2003) Jason-1 precision orbit determination using GPS combined with SLR and DORIS tracking data. Ph.D. Thesis. The University of Texas at Austin, Austin, USA
- Colombo OL (1989) The dynamics of global positioning system orbits and the determination of precise ephemerides. *J Geophys Res* 94:9167–9182

- Couhert A, Cerri L, Legeais JF, Ablain M, Zelensky NP, Haines BJ, Lemoine FG, Bertiger WI, Desai SD, Otten M (2015) Towards the 1 mm/y stability of the radial orbit error at regional scales. *Adv Space Res* 55(1):2–3. doi:[10.1016/j.asr.2014.06.041](https://doi.org/10.1016/j.asr.2014.06.041)
- Doornbos E, Scharroo R, Klinkrad H, Zandbergen R, Fritsche B (2002) Improved modelling of surface forces in the orbit determination of ERS and Envisat. *Can. J. Remote Sens.* 28(4):535–543
- Gao F, Peng B, Zhang Y, Evariste NH, Liu J, Wang X, Zhong M, Lin M, Wang N, Chen R, Xu H (2015) Analysis of HY2A precise orbit determination using DORIS. *Adv Space Res* 55(5):1394–1404. doi:[10.1016/j.asr.2014.11.032](https://doi.org/10.1016/j.asr.2014.11.032)
- Goad CC, Goodman L (1974) A modified Hopfield tropospheric refraction correction model. In: American Geophysical Union Annual Fall Meeting. San Francisco, USA
- Guo JY, Qin J, Kong QL, Li GW (2012) On simulation of precise orbit determination of HY-2 with centimeter precision based on satellite-borne GPS technique. *Appl Geophys* 9(1):95–107. doi:[10.1007/s11770-012-0319-3](https://doi.org/10.1007/s11770-012-0319-3)
- Guo J, Kong Q, Qin J, Sun Y (2013a) On precise orbit determination of HY-2 with space geodetic techniques. *Acta Geophys* 61(3):752–772. doi:[10.2478/s11600-012-0095-8](https://doi.org/10.2478/s11600-012-0095-8)
- Guo J, Zhao QL, Li M, Hu ZG (2013b) Centimeter level orbit determination for HY2A using GPS data (in Chinese). *Geomatics and Information Science of Wuhan University* 38(1):52–55
- Gurtner W, Noomen R, Pearlman MR (2005) The international laser ranging service: current status and future developments. *Adv Space Res* 36(3):327–332. doi:[10.1016/j.asr.2004.12.012](https://doi.org/10.1016/j.asr.2004.12.012)
- Hackel S, Steigenberger P, Hugentobler U, Uhlemann M, Montenbruck O (2015) Galileo orbit determination using combined GNSS and SLR observations. *GPS Solutions* 19(1):15–25. doi:[10.1007/s10291-013-0361-5](https://doi.org/10.1007/s10291-013-0361-5)
- Hedin AE (1987) MSIS-86 thermospheric model. *J Geophys Res* 92(A5):4649–4662
- Hopfield H (1971) Tropospheric effect on electromagnetically measured range: prediction from surface weather data. *Radio Sci* 6(3):357–367. doi:[10.1029/RS006i003p00357](https://doi.org/10.1029/RS006i003p00357)
- Jiang XW, Wang XH, Peng HL, Zhong M, Zhong SM (2014) Precise orbit determination technology based on DORIS solution for HY-2satellite (in Chinese). *Eng Sci* 16(6):83–89
- Kang Z, Tapley B, Bettadpur S, Ries J, Nagel P, Pastor R (2006) Precise orbit determination for the GRACE mission using only GPS data. *J. Geod.* 80:322–331. doi:[10.1007/s00190-006-0073-5](https://doi.org/10.1007/s00190-006-0073-5)
- Knocke PC, Ries JC, Tapley BD (1988) Earth radiation pressure effects on satellites. Proceedings of the AIAA/AAS Astrodynamics Conference, August 15–17. Minneapolis, Minnesota, pp 577–586
- Kong QL, Guo JY, Qin J, Sun Y (2013) Simulation of centimeter-level precise orbit determination for the HY-2 satellite using DORIS and SLR (in Chinese). *Geomatics and Information Science of Wuhan University* 38(6):694–699
- Kong Q, Guo J, Hwang C, Gao F, Lin SF, Zhao C (2014) Precise orbit determination and accuracy analysis of HY-2A satellite using DORIS Doppler data. *Acta Geod. Geophys.* 49:455–470. doi:[10.1007/s40328-014-0066-4](https://doi.org/10.1007/s40328-014-0066-4)
- Lemoine FG, Zelensky NP, Chin DS, Pavlis DE, Rowlands DD, Beckley BD, Luthcke SB, Willis P, Ziebart M, Sibthorpe A, Boy JP, Luceri V (2010) Towards development of a consistent orbit series for TOPEX Jason-1, and Jason-2. *Adv Space Res* 46(12):1513–1540. doi:[10.1016/j.asr.2010.05.007](https://doi.org/10.1016/j.asr.2010.05.007)
- Li S, Galas R, Ewert D, Peng J (2016) An empirical model for the ionospheric global electron content storm-time response. *Acta Geophys* 64(1):253–269. doi:[10.1515/acgeo-2015-0067](https://doi.org/10.1515/acgeo-2015-0067)
- Lin MS, Wang XH, Peng HL, Zhao QL, Li M (2014) Precise orbit determination technology based on dual-frequency GPS solution for HY-2 satellite (in Chinese). *Eng Sci* 16(6):97–101
- Liu JH (2013) Satellite orbit determination method with DORIS system (in Chinese). Master Thesis. National University of Defense Technology, Changsha, China
- Luthcke SB, Rowlands DD, Lemoine FG (2002) The SLR contribution to precise orbit determination in the GPS era. In: 13th International Workshop on Laser Ranging “Toward Millimeter Accuracy”. Washington D.C., USA
- Luthcke SB, Zelensky NP, Rowlands DD, Lemoine FG, Williams TA (2003) The 1-centimeter orbit: Jason-1 precision orbit determination using GPS, SLR, DORIS, and Altimeter Data. *Mar Geod* 26:399–421
- Lyard F, Lefevre F, Letellier T, Francis O (2006) Modelling the global ocean tides: modern insights from FES2004. *Ocean Dynam.* 56(5–6):394–415. doi:[10.1007/s10236-006-0086-x](https://doi.org/10.1007/s10236-006-0086-x)
- Mansoori AA, Khan PA, Atulkar R, Purohit PK, Gwal AK (2015) Ionospheric influences on GPS signals in terms of range delay. *Russian J Earth Sci* 15(3):1–9. doi:[10.2205/2015ES000555](https://doi.org/10.2205/2015ES000555)
- Marini JW, Murray CW (1973), Correction of laser range tracking data for atmospheric refraction at elevations above 10 degrees. GSFC Report X-591-73-351, Goddard Space Flight Center, Greenbelt, USA
- McCarthy DD, Petit G (2003) Equations of motion for an artificial earth satellite. In: IERS Conventions 2003. IERS Technical Note 32, 106–107, Paris, France
- Melachroinos SA, Lemoine FG, Zelensky NP, Rowlands DD, Deng C, Pavlis D (2011) Status of precise orbit determination for JASON-2 using GPS, SLR and DORIS data at NASA/GSFC. *Geophys Res Abstract* 13:EGU2011-10409
- Mendes VB, Pavlis EC (2004) High-accuracy zenith delay prediction at optical wavelengths. *Geophys Res Lett* 31(L14):602. doi:[10.1029/2004GL020308](https://doi.org/10.1029/2004GL020308)
- Mendes VB, Prates G, Pavlis EC, Langley RB (2002) Improved mapping functions for atmospheric refraction correction in SLR. *Geophys Res Lett* 29(10):53-1–53-4. doi:[10.1029/2001GL014394](https://doi.org/10.1029/2001GL014394)
- Mercier F, Cerri L, Berthias JP (2010) Jason-2 DORIS phase measurement processing. *Adv Space Res* 45:1441–1454. doi:[10.1016/j.asr.2009.12.002](https://doi.org/10.1016/j.asr.2009.12.002)
- Pavlis NK, Holmes SA, Kenyon SC, Factor JK (2012) The development and evaluation of the Earth Gravitational Model 2008 (EGM2008). *J Geophys Res* 117:B04406. doi:[10.1029/2011JB008916](https://doi.org/10.1029/2011JB008916)
- Pearlman MR, Degan JJ, Bosworth JM (2002) The international laser ranging service. *Adv Space Res* 30(2):135–143. doi:[10.1016/S0273-1177\(02\)00277-6](https://doi.org/10.1016/S0273-1177(02)00277-6)
- Peng DJ, Wu B, Qu WJ (2012) Jason-2 precise orbit determination with DORIS/SLR tracking data (in Chinese). *J Astronaut* 33(10):1391–1400
- Petit G, Luzum B (2010) IERS conventions (2010). Bureau International des Poids et Mesures, Sevres (France)
- Rim HJ (1992) TOPEX orbit determination using GPS tracking system. Ph.D. Thesis, University of Texas at Austin, Austin, USA
- Rudenko S, Otten M, Visser P, Scharroo R, Schoene T, Esselborn S (2012) New improved orbit solutions for the ERS-1 and ERS-2 satellites. *Adv Space Res* 49:1229–1244. doi:[10.1016/j.asr.2012.01.021](https://doi.org/10.1016/j.asr.2012.01.021)
- Seeber G (1993) *Satellite Geodesy*. Walter de Gruyter, Berlin
- Sheng CZ, Gan VJ, Zhao CM, Zhang XQ, Sun BQ, Chen WT (2014) Assessment of precise orbit determination of Jason-2 satellite using different observation technologies (in Chinese). *Acta Geodaetica et Cartographica* 43(8):796–802. doi:[10.13485/j.cnki.11-2089.2014.0133](https://doi.org/10.13485/j.cnki.11-2089.2014.0133)
- Sośnica K, Thaller D, Dach R, Steigenberger P, Beutler G, Arnold D, Jäggi A (2015) Satellite laser ranging to GPS and GLONASS. *J. Geod.* 89(7):725–743. doi:[10.1007/s00190-015-0810-8](https://doi.org/10.1007/s00190-015-0810-8)

- Standish EM (1998) JPL Planetary and Lunar ephemerides, DE405/LE405. Jet Propulsion Laboratory, Inter Office Memorandum, IOM, 312. F-98-048, 1-18
- Švehla D, Rothacher M (2003) Kinematic and reduced-dynamic precise orbit determination of low Earth orbiters. *Adv. Geosci.* 1:47–56
- Tapley BD, Ries JC, Davis GW, Eanes RJ, Schutz BE, Shum CK, Watkins MM, Marshall JA, Nerem RS, Putney BH, Klosko SM, Luthcke SB, Pavlis D, Williamson RG, Zelensky NP (1994) Precision orbit determination for TOPEX/POSEIDON. *J Geophys Res* 99:24383–24404
- Tavernier G, Fagard H, Feissel-Vernier M, Bail KL, Lemoine F, Noll C, Noomen R, Ries JC, Soudarin L, Valette JJ, Willis P (2006) The international DORIS service: genesis and early achievements. *J. Geod.* 80:403–417. doi:[10.1007/s00190-006-0082-4](https://doi.org/10.1007/s00190-006-0082-4)
- Thaller D, Sošnica K, Dach R, Jäggi A, Beutler G, Mareyen M, Richter B (2013), Geocenter coordinates from GNSS and combined GNSS-SLR solutions using satellite co-locations. In: Rizos C, Willis P (eds) *Earth on the edge: science for a sustainable planet* 139:129–134. Springer, Berlin Heidelberg, doi:[10.1007/978-3-642-37222-3_16](https://doi.org/10.1007/978-3-642-37222-3_16)
- Tseng TP, Hwang CW, Yang SK (2012) Assessing attitude error of FORMOSAT-3/COSMIC satellites and its impact on orbit determination. *Adv Space Res* 49:1301–1312. doi:[10.1016/j.asr.2012.02.007](https://doi.org/10.1016/j.asr.2012.02.007)
- Urschl C, Gurtner W, Hugentobler U, Schaer S, Beutler G (2005) Validation of GNSS orbits using SLR observations. *Adv Space Res* 36:412–417. doi:[10.1016/j.asr.2005.03.021](https://doi.org/10.1016/j.asr.2005.03.021)
- Wang XH, Peng HL, Lin MS, Zhou XH (2014) Precise orbit determination technology based on SLR solution for HY-2 satellite (in Chinese). *Eng Sci* 16(6):90–96
- Willis P (1995) DORIS coordinating center report. In: *IERS Annual Report 1995*. IERS, Paris, France
- Willis P, Gobinddass ML, Garayt B, Fagard H (2012) Recent improvements in DORIS data processing at IGN in view of ITRF2008, the ignwd08 solution. *Geodesy for Planet Earth International Association of Geodesy Symposia* 136:43–49. doi:[10.1007/978-3-642-20338-1_6](https://doi.org/10.1007/978-3-642-20338-1_6)
- Willis P, Zelensky NP, Ries JC, Soudarin L, Cerri L, Moreaux G, Lemoine FG, Otten M, Argus DF, Heflin MB (2013), DPOD2008, A DORIS-oriented terrestrial reference frame for precise orbit determination. In: Rizos C, Willis P (eds) *International Association of Geodesy Symposia*, 143, Springer, Switzerland
- Wu B, Lin MS, Zhang ZP (2011) Global SLR tracking support for HY-2 satellite precise orbit determination. Proceedings of the 17th International Workshop on Laser Ranging, Bad Koetzing, Germany (online at http://cddis.gsfc.nasa.gov/lw17/docs/papers/posters/34-Shanghai_Paper_4.pdf)
- Zelensky NP, Lemoine FG, Ziebart M, Sibthorped A, Willis P, Beckley BD, Klosko SM, Chinn DS, Rowlands DD, Luthcke SB, Pavlis DE, Luceri V (2010) DORIS/SLR POD modeling improvements for Jason-1 and Jason-2. *Adv Space Res* 46:1541–1558. doi:[10.1016/j.asr.2010.05.008](https://doi.org/10.1016/j.asr.2010.05.008)
- Zhang Y, Vincent T (2011), HY2-ICD-0-00009-CNES technical. Document for HY2A Ground Segment
- Zhang Q, Liao XH, Huang C (2000) An investigation of the precise orbit determination by combination of two kinds of measurements (in Chinese). *Acta Astronomica Sinica* 41(4):347–354
- Zhao G, Zhou XH, Wu B (2013) Precise orbit determination of Haiyang-2 using satellite laser ranging. *Chinese Sci Bull* 58(6):589–597. doi:[10.1007/s11434-021-5564-6](https://doi.org/10.1007/s11434-021-5564-6)
- Zhu J, Wang JS, Chen JR, He YF (2013) Centimeter precise orbit determination for HY-2 via DORIS (in Chinese). *J Astronaut* 34(2):163–169. doi:[10.3873/j.issn.1000-1328.2013.02.003](https://doi.org/10.3873/j.issn.1000-1328.2013.02.003)

Sibsonian and non-Sibsonian natural neighbour interpolation of the total electron content value

Kacper Kotulak¹ · Adam Fron¹ · Andrzej Krankowski¹ · German Olivares Pulido² · Manuel Henrandez-Pajares³

Received: 3 January 2017 / Accepted: 4 January 2017 / Published online: 21 January 2017
© The Author(s) 2017. This article is published with open access at Springerlink.com

Abstract In radioastronomy the interferometric measurement between radiotelescopes located relatively close to each other helps removing ionospheric effects. Unfortunately, in case of networks such as LOW Frequency ARray (LOFAR), due to long baselines (currently up to 1500 km), interferometric methods fail to provide sufficiently accurate ionosphere delay corrections. Practically it means that systems such as LOFAR need external ionosphere information, coming from Global or Regional Ionospheric Maps (GIMs or RIMs, respectively). Thanks to the technology based on Global Navigation Satellite Systems (GNSS), the scientific community is provided with ionosphere sounding virtually worldwide. In this paper we compare several interpolation methods for RIMs computation based on scattered Vertical Total Electron Content measurements located on one thin ionospheric layer (Ionospheric Pierce Points—IPPs). The results of this work show that methods

that take into account the topology of the data distribution (e.g., natural neighbour interpolation) perform better than those based on geometric computation only (e.g., distance-weighted methods).

Keywords TEC · Interpolation · Natural neighbour interpolation

Introduction

Information regarding the state of the ionosphere is paramount for the performance of scientific and technological infrastructures such as, for example, power grid networks, pipelines and satellite-based navigation services (e.g., GPS, GLONASS, and soon, Galileo and Beidou too).

It is well known that at short baselines (e.g., less than ~50 km), most of the ionospheric effects may be removed by means of interferometric techniques. Otherwise, for longer baselines, the ionospheric variability must be taken into account.

Interferometric techniques are also widely used in radioastronomy. Nevertheless, radio-telescope networks that operate along baselines longer than 50 km may not be able to remove the ionospheric delay of the signal by means of interferometry between nearby stations. That is the case of LOFAR project (<http://www.lofar.org/>), a network of radio-telescopes in Europe that operates at the low frequency band of the electromagnetic spectrum (10–240 MHz) with baselines as long as 1500 km. Consequently, LOFAR network requires ionospheric modelling in order to minimize the degradation of the observed signal.

Among several experiments and goals of LOFAR, the ionosphere may affect the observational data that are used

✉ Kacper Kotulak
kacper.kotulak@uwm.edu.pl

Adam Fron
adam.fron@uwm.edu.pl

Andrzej Krankowski
kand@uwm.edu.pl

German Olivares Pulido
g.olivares@bom.gov.au

Manuel Henrandez-Pajares
manuel.hernandez@upc.edu

¹ Space Radio-Diagnostics Research Centre, University of Warmia and Mazury, Olsztyn, Poland

² CRC for Spatial Information, Victoria, Australia

³ Department of Applied Mathematics IV, Ionospheric Determination and Navigation Based on Satellite and Terrestrial Systems Research Team, Barcelona, Spain

to constrain theoretical models of the epoch of the reionization, solar science, space weather and cosmic magnetism (see www.lofar.org for further details). One of the Polish LOFAR part tasks is high-resolution monitoring of radio sources in cooperation with GNSS to study ionosphere influence on low-frequency radiowaves. Possibilities of ionospheric, space weather and solar studies were described by Dąbrowski et al. (2016).

The incapability of interferometric techniques of minimizing the impact of the ionospheric delay on the signal might be overcome by using either GIMs or RIMs which provide insight into the state of the ionosphere. Indeed, the International GNSS Service (IGS) provides GIMs through four IGS Ionosphere Associate Analysis Centres (IAACs); CODE, ESA (European Space Agency), JPL (Jet Propulsion Laboratory) and UPC (Universitat Politècnica de Catalunya) (Komjathy et al. 2010; Schaer 1999).

These global ionospheric maps are computed using Total Electron Content (TEC) data provided by GNSS-based ground receivers (Hernández-Pajares et al. 2011). Due to the lack of homogeneity of the GNSS networks worldwide, interpolation techniques are used for filling the gaps (Orus et al. 2005 and citations therein). However, GIMs do not provide the highest resolution for local areas. In such case, RIMs should be used instead (Jakowski et al. 2005).

There are few GNSS ground receiver networks that may provide high-resolution RIMs, such as, for example, GEONET (<http://www.gsi.go.jp/>), in Japan and CRTN in California (<http://sopac.ucsd.edu/>). In Europe, EUREF permanent network (<http://www.epncb.oma.be/>) has longer inter-station distance than the abovementioned networks.

The accuracy of GIMs and RIMs depends on the accuracy of TEC data and ionospheric modelling. For example, GIMs provided by the UPC are computed using tomographic Slant TEC (STEC) data (Hernandez-Pajares et al. 1999) and the use of fast and accurate ways of solving the associated topological problem has also been studied (see Hernández-Pajares et al. 1997). Then, gaps (due to the lack of receivers in the Southern Hemisphere and the non-homogeneous distribution of GNSS ground receivers) are filled by estimating the correlation between the estimated uncertainties with the kriging method (Orus et al. 2005).

Other ionospheric modelling methods are data-assimilative models (GAIM, at the Jet Propulsion Laboratory—Schunk 2002; Pi et al. 2003, 2004, 2009; Hajj 2004; Wang 2004; Mandrake et al. 2005; Scherliess et al. 2006; Komjathy et al. 2010), harmonic spherical functions (Schaer 1999; Choi et al. 2010; Han et al. 2013) and empirical models (e.g., the International Reference Ionosphere project—IRI, <http://iri.gsfc.nasa.gov/>) (Strangeways 2009; Buresova et al. 2009).

It is also important to take into account the time scale of the ionospheric variability. The time-scale spectrum spans from seconds (e.g., scintillations) to years (Solar cycle), thus making the ionospheric temporal variability modelling a real challenge (Krakowski et al. 2007).

In general, the criteria to choose amongst several ionospheric modelling methods may be based on their use. For example, for radio-telescope networks, it is important to provide ionospheric information in real time within the area of coverage. In this regard, tomographic models can estimate the state of the ionosphere in real time (Hernández-Pajares et al. 2000, 2011).

As for the spatial resolution, any model or interpolation method degrades the uncertainty of the raw STEC data, no matter how good it may be. Consequently, for the computation of RIMs, it is also of great importance to find the best spatial modelling available, that is to say, the model that yields the minimum discrepancy with real data.

In this regard, this work presents a comparison of several interpolation methods used for computing RIMs over Europe using GPS-ground receiver stations from the EUREF network in Poland. All these methods use STEC data computed by the UPC TOMographic Model of the IONosphere software (UPC TOMION).

This work is organized as follows: the second section presents a description of the interpolation methods; the third section presents the STEC data set and the results; the fourth section contains the solutions to some interpolation problems; finally, the fifth section summarizes the conclusions.

Description of chosen interpolation methods

Inverse distance weighting (IDW)

Inverse distance weighting can be considered as one of the most intuitive and easiest methods of evaluating the influence of values of certain phenomena measured in less or more scattered data points with unknown values at any arbitrary point.

The general idea of inverse distance weighting is the lowering of influence with increasing distance between the points. The result of the interpolation at each point is a weighted average of the values of the sample or data set, where the weights are decreasing proportionally to a power function of increasing distance between the points (Dumitru 2013). As a degree of IDW we consider an exponentiation of X lying in the denominator (meaning x^{-1} as the first degree, x^{-2} as the second, and so on). For our comparison we have used the second degree of IDW, which is also the most commonly used one.

Two scenarios of IDW method implementation can be considered—global and local. The global IDW interpolation assumes that all the data points have an influence on the interpolated points. In our comparison we have used a kind of local approach, where only an arbitrary number of points (7—which was the average number of neighbours in our natural interpolation) selected based on the closest distance, is considered to influence the value of the interpolated point. In our comparison the values at each point are evaluated using the following formula:

$$F(X, Y) = \sum_{i=1}^n w_i f_i, \tag{1}$$

where n is the number of neighbouring points, f_i is the function value at each neighbouring point and w_i is a normalized weight:

$$w_i = \frac{d_i^{-2}}{\sum_{i=1}^n w_i f_i}, \tag{2}$$

where d stands for the distance between each data point and the interpolated point.

Polynomial interpolation

Polynomial fitting differs significantly from the other methods. Indeed, the computed surface does not fit the data points, but instead it lies within a certain distance from them.

For the approximation of the polynomial coefficients, a_1, a_2, \dots, a_n , we have used the Moore–Penrose pseudoinverse (see Moore 1920) based on the whole set of data points for each epoch. Similarly to IDW, polynomials can be evaluated to a certain degree, resulting in a variety of computed surfaces (e.g., a plane, a bilinear surface, a cubic area, etc.). In our comparison we have used a force method to evaluate an optimal polynomial degree for each epoch (of course in a certain range, as the pseudoinverse gives the possibility of creating polynomials of a degree greatly exceeding the number of data points).

Figure 1 shows the number of coefficients of the polynomial equation in terms of its order. For the selected (6th) order, the polynomial equation can be expressed as follows:

Fig. 1 The polynomial parameters for 6th order polynomial equation in regard to two variables (coordinates)

| | x^40 | x^41 | x^42 | x^43 | x^44 | x^45 | x^46 |
|--------|--------|--------|--------|--------|--------|--------|--------|
| y^40 | a1 | a3 | a6 | a10 | a15 | a21 | a28 |
| y^41 | a2 | a5 | a9 | a14 | a20 | a27 | 0 |
| y^42 | a4 | a8 | a13 | a19 | a26 | 0 | 0 |
| y^43 | a7 | a12 | a18 | a25 | 0 | 0 | 0 |
| y^44 | a11 | a17 | a24 | 0 | 0 | 0 | 0 |
| y^45 | a16 | a23 | 0 | 0 | 0 | 0 | 0 |
| y^46 | a22 | 0 | 0 | 0 | 0 | 0 | 0 |

$$f(x, y) = \sum_{i=0}^6 \left(\sum_{j=0}^{6-i} a_{ij} x^i y^j \right), \tag{3}$$

where a_{ij} are the polynomial coefficients and x, y are the coordinates of each point.

For low degree (up to 10), the polynomial method returns a surface which approximates data smoothly, and it is easily computed, resulting in a continuous surface that easily provides a value for any arbitrary point.

Voronoi diagram and Delaunay triangulation

Natural neighbour interpolation is based on the Voronoi diagram and Delaunay triangulation, which defines the topology in terms of the distribution of nearest neighbours around each point (Sukumar et al. 2001).

To illustrate the idea of Voronoi diagram (often called Dirichlet tessellation or Thiessen diagram) we will consider dataset that consisted of scattered points, called Voronoi sites.

Each site corresponds to a certain Voronoi cell (also called Voronoi face), which defines the range of points that are closer to the chosen site than to other sites. The Voronoi diagram is simply the sum of cells corresponding to all Voronoi sites (Fortune 1995). That leads to the conclusion that a Voronoi diagram is actually a realization of the nearest neighbour interpolation (see Fig. 2).

Voronoi diagram is strictly related to the Delaunay triangulation. Delaunay triangulation is the realization of a set of triangles created in such way that each triangle circumscribed does not have any Voronoi site in its interior (Fortune 1995).

Voronoi edge is a mid-perpendicular of a Delaunay triangle edge. In other words, the Delaunay edge existence determines the existence of Voronoi edge and vice versa.

Voronoi diagrams determine natural neighbourhood—two Voronoi sites are each other’s natural neighbour if there exists a Voronoi edge between them. The same rule can be applied to the Delaunay triangulation—two points are their own natural neighbours if they are linked by a Delaunay edge (see Fig. 3).

Both Voronoi diagram and Delaunay triangulation can be used in interpolation—but only Delaunay triangulation can be used separately, as the easiest way to compute the Voronoi diagram requires also a Delaunay triangulation.

Fig. 2 Graphical comparison between Voronoi diagram (a) and discrete nearest neighbour interpolation of regular grid set of 10,000 points (b)

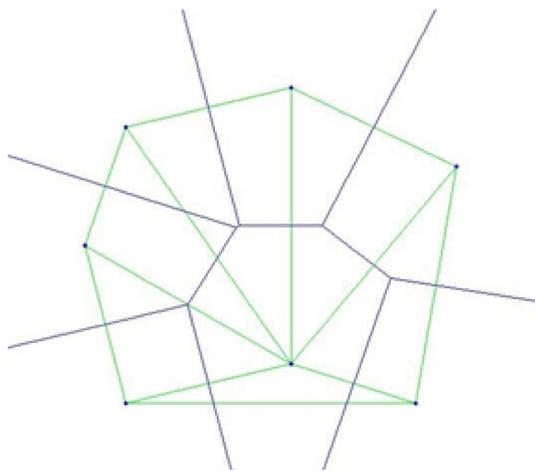
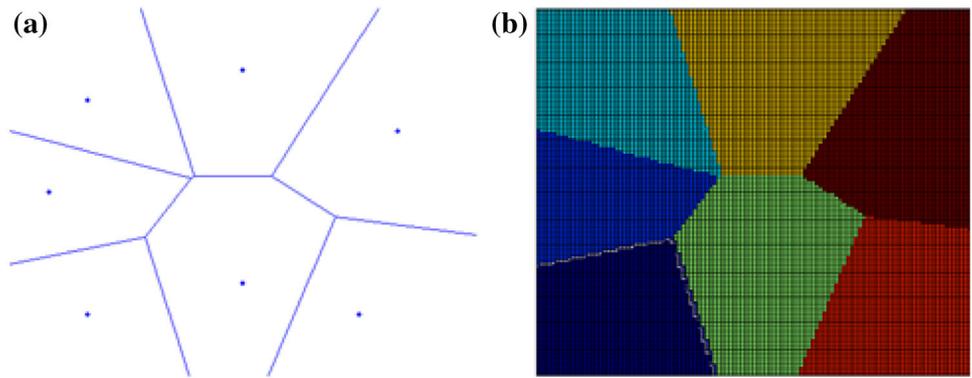


Fig. 3 Voronoi diagram and Delaunay triangulation put together (Voronoi edges are Delaunay edges' mid-perpendiculars, but not the reverse). Not every Voronoi edge can be seen

Natural neighbour interpolation

As previously said, natural neighbour interpolation is a method based on the Voronoi diagram. In calculations we consider two orders of the diagram—the first containing only scattered known data points, and the second order, which is based on the first, but with one of the interpolated points inserted into the data (Fig. 4).

The insertion of a new point into an existing Voronoi diagram results in the creation of a new cell, the area of which 'steals' parts of the neighbouring cells, while the whole rest of the diagram remains unchanged.

The weights of all neighbours of the inserted point are then evaluated as a ratio of the stolen (overlapping) area to the area of the whole new cell, as shown by Eqs. 4 and 5 below (Sibson 1981):

$$f_x = \sum_{m=1}^M \alpha_m f_m, \quad (4)$$

where f_x is the value for the interpolated point, M is the number of natural neighbours (number of neighbouring

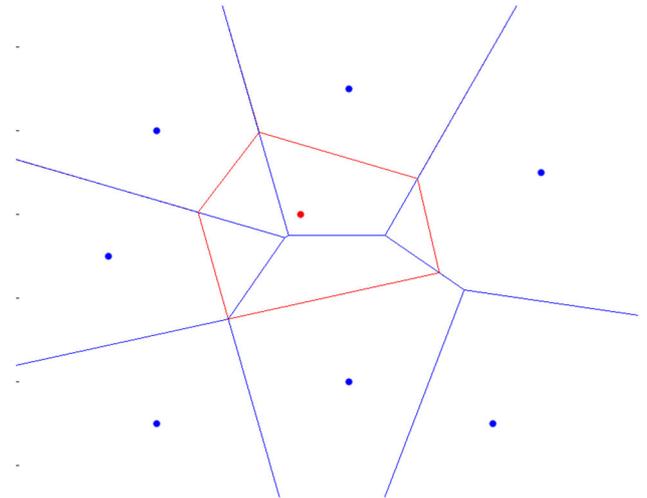


Fig. 4 Two orders of Voronoi diagram—the first (blue) consisting of known points and their cells, and an overlapping cell of the inserted (and then interpolated) point (red) of the second order

Voronoi cells), f_m describes the value of the m -th neighbour and α_m is a weight coefficient computed with Eq. 5 as follows:

$$\alpha_m = \frac{s_m}{s_0}, \quad (5)$$

where s_m is the stolen area of m -th cell and s_0 is the whole area of the new cell (see Fig. 5).

Overlapping areas can be computed in two different ways. In the first scenario, the overlapping area is acquired by subtracting the neighbouring cells area before entering the new point and the same cell after the introduction of such point (Harman 2008). The second scenario is to compute the overlapping area as the area of the intersection of the new cell and the neighbouring cell.

The first scenario is easier to calculate, but much more sensitive and exposed to errors caused by open Voronoi cells. Those errors will be further discussed in later sections of this work.

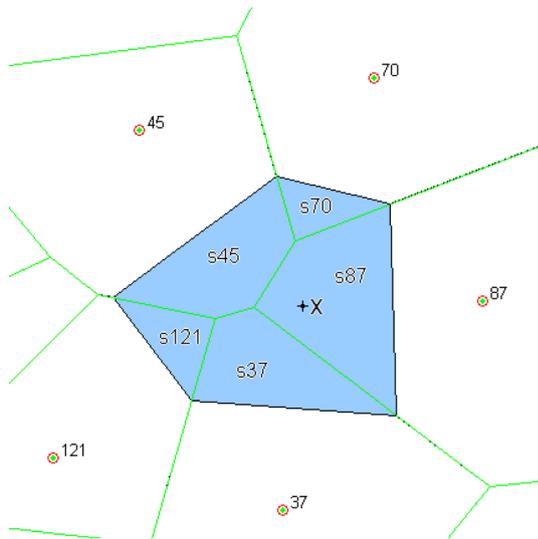


Fig. 5 The area of the new cell divided into sections ‘stolen’ from each neighbour

Non-Sibsonian interpolation

The non-Sibsonian interpolation method is actually a natural neighbour method; therefore, it is also based on Voronoi diagram computation. Nevertheless, unlike the Sibsonian method, this method uses linear quantities instead of areas. This approach solves some of the problems the Sibsonian method faces (e.g., open Voronoi cells).

According to Belikov et al. (1997), the interpolated point can be computed as a linear combination of the neighbours as follows:

$$f_x = \sum_{m=1}^M \alpha_m f_m, \quad (6)$$

where f_x is the value of the interpolated point x , M is the number of natural neighbours (i.e., number of Voronoi cell edges), f_m is the value of the m -th neighbour and α_m is a weight coefficient described by the equation given below:

$$\alpha_m = \frac{d_m}{h_m} \quad (7)$$

with d_m the Voronoi edge length and h_m standing for half the distance between the interpolated point x and the Voronoi edge m (see Fig. 6).

Belikov et al. (1997) note that such a way of computing weights is far easier and more efficient than the Sibsonian approach, which is based on polygons’ areas. Moreover, the non-Sibsonian method does not require building two-order Voronoi diagrams. Finally, it is also free from some area calculation problems, which are present in the Sibsonian natural neighbour method and will be further discussed in posterior sections.

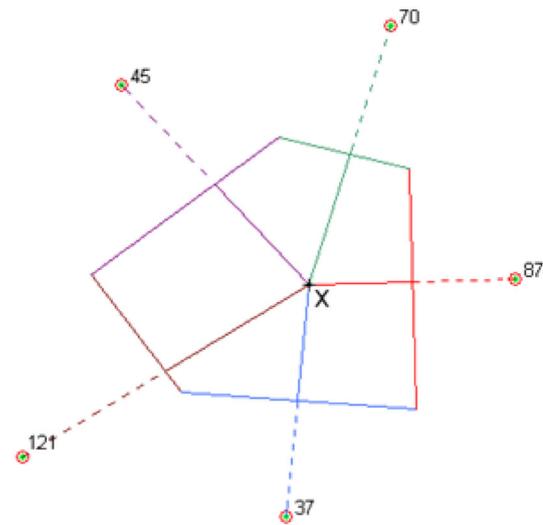


Fig. 6 Non-Sibsonian interpolation mechanism. Edges and distances of the same colour correspond to pairs of d_m and h_m , respectively, used to compute α_m

Quasi-natural interpolation

Quasi-natural method is a composition of the inverse distance weighting interpolation method and Delaunay triangulation. We use prefix ‘quasi’, as the method is not actually based on Voronoi diagram and its features, like classical natural methods. Delaunay triangulation is used to find the target set of neighbour, which will substitute the ‘random set of points’ described by the Shepard method (Dumitru 2013). Natural neighbours chosen this way will gain weights equal to those in Shepard method described with the following equation:

$$w_m = \frac{1}{d_{x,m}^p}, \quad (8)$$

where w_m is the weight corresponding to the neighbour m , $d_{x,m}$ is the distance between the target point x and its neighbour m and p is a power factor. We will use $p = 2$ as it is one of the most commonly used (Dumitru 2013).

A very important issue is the differentiation between natural and nearest neighbours, as even in the local solution the nearest neighbours could be not the same as natural ones (see Figs. 7, 8).

Differentiation between those two groups of neighbours is extremely important for irregularly scattered data.

Comparison between methods

Dataset

Exemplary computations were based on the dataset that consisted of TEC value observations from 19 EUREF

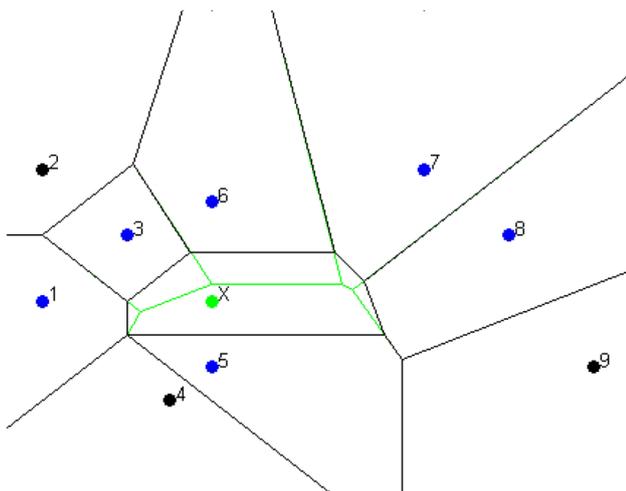


Fig. 7 Natural neighbours (blue) of a certain point X

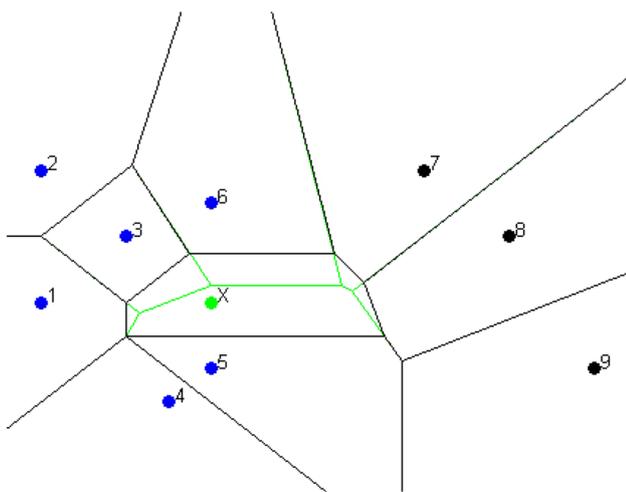


Fig. 8 Nearest neighbours (blue) of a certain point X

Permanent Network (EPN) stations located within and nearby the area of Poland (see Fig. 9). The data set provided by the UPC included Slant TEC (STEC) values for each IPP. The STEC values were observed alongside the lines of sight between the EPN station and every satellite in view. Then it was computed into Vertical TEC (VTEC) related to the proper IPP assuming a 450-km thin-shell height as Fig. 10 shows. The mapping function used for the computation of the VTEC is as follows:

$$\text{VTEC} = \text{STEC} * \sqrt{1 - \left(\frac{R_e}{R_e + h_{\text{ion}}} \cos \varepsilon \right)^2} \quad (9)$$

where ε is the elevation between the receiver and satellite, h_{ion} is the height of ionospheric thin layer and R_e is the Earth's radius.

The IPPs of three stations: Borowiec (bor1), Bydgoszcz (bydg) and Lodz (lodz) were taken as unknown for the

interpolation. Then their known VTEC values were used to assess the performance of the interpolation method.

Those three stations have been chosen due to their locations. They are indeed surrounded by other stations from all sides, thus avoiding extrapolation scenarios, which are out of the scope of this work. As we show later, natural neighbour method does not work well for extrapolation. Although inner stations may have IPPs surrounded by open Voronoi cells, the likelihood of such scenario happening is smaller than for any other station closer to the border of the network.

Moreover, the interpolation was performed on an epoch-basis and every unknown point was interpolated separately and never added to the known dataset. This procedure was aimed at avoiding potential error.

Considering the fact that the dataset contains 30-s interval data for the whole day (24 h observations from 15 June 2015, during quiet geomagnetic conditions after 2013 solar activity peak), the observations of 16 stations from the EUREF network provided 74400 interpolated points divided into 2880 one-epoch subsets, which is a good statistical sample for assessing the performance of the proposed methods.

Results

For each point in each interpolation method we established a level of confidence, including relative errors, standard deviations (as we have the „true” values for the interpolated variables) and root means squares for every whole epoch.

Results are presented below (see Tables 1, 2, 3; Figs. 11, 12, 13). Computation time was also taken into account, as we will try to evaluate the possibility of using natural methods to compute rapid, near-real time local TEC values. Root mean squares, relative errors and standard deviations are shown only for the second scenario of natural neighbour interpolation, as the results in both scenarios are the same, but they are more distorted in the first one by lost points (point-losing problem will be further analysed in other sections).

Also, as it was said earlier, we have taken computation time into account. Results are shown below (see Table 4).

The fastest method is undoubtedly the quasi-natural neighbour method. The slowest methods are the Sibsonian natural neighbour methods. Indeed, as they build two Voronoi diagrams for each point they require longer computational time. The difference of both Sibsonian natural neighbour methods in terms of computational time stems from the fact that the second scenario requires the computation of two polygons intersection. Consequently, it takes longer than simple subtraction of two numbers representing areas of the corresponding Voronoi cells, which is the method used in the first scenario. The non-Sibsonian natural neighbour method requires building only one

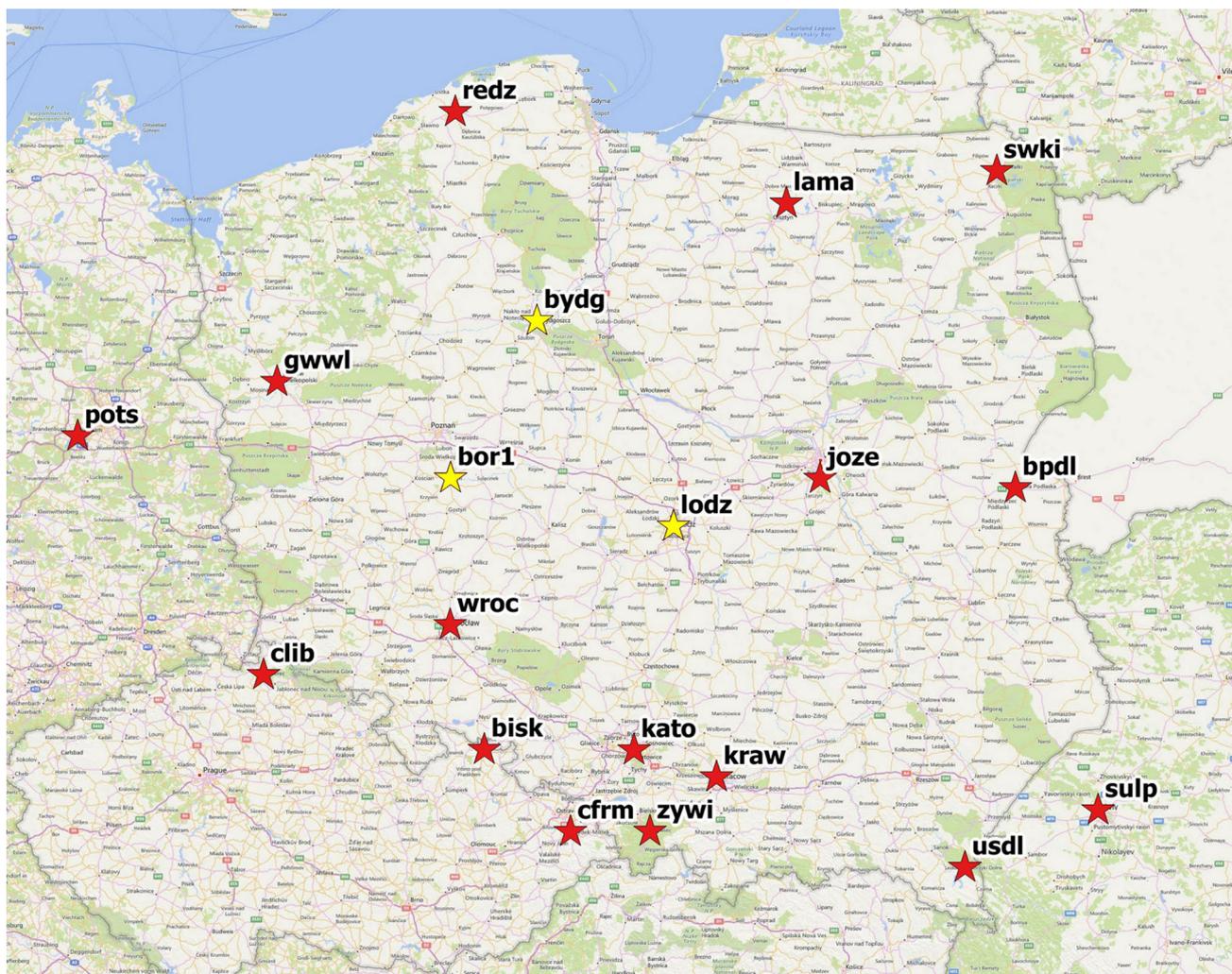


Fig. 9 Location of analysed EPN stations (stations taken as unknowns are marked with *yellow symbol*)

Voronoi diagram and does not require any area computation at all. Weights are based on simple linear quantities (distance between points and length of the Voronoi edge). The fastest quasi-natural neighbour method does not build any Voronoi diagrams at all as it is based on the easier (and quicker) Delaunay triangulation method.

As shown above, we received some statistical information about efficiency of each presented interpolation method, but before we set final efficiency marks, we should turn the attention to some problems we faced while conducting interpolation using natural neighbour methods.

Troubleshooting

Losing points

The first problem with natural neighbour methods is that not every point can be interpolated. For example, during

the process of computation some of them (see Table 5) have not received any value.

The method based on the first scenario lost over 2500 times more points than the second one. Such huge difference between both Sibsonian methods suggests that the computational method systematically fails under some circumstances for the first scenario. In further analysis we will show that the second scenario method loses far more points than it was presented in Table 5, i.e., the losses are simply hidden.

Six lost points of the natural neighbour method (second scenario) are actually located outside the network (we based our computations on IPPs, which are dependent on rapidly changing position of satellite on the sky)—hence the interpolation method failure.

The first scenario method fails to interpolate a point when at least one of the neighbours corresponds with an open Voronoi cell. Indeed, in such a case the area of the Voronoi cell is actually infinite (see Fig. 14), which corresponds to a not-a-number computational result.

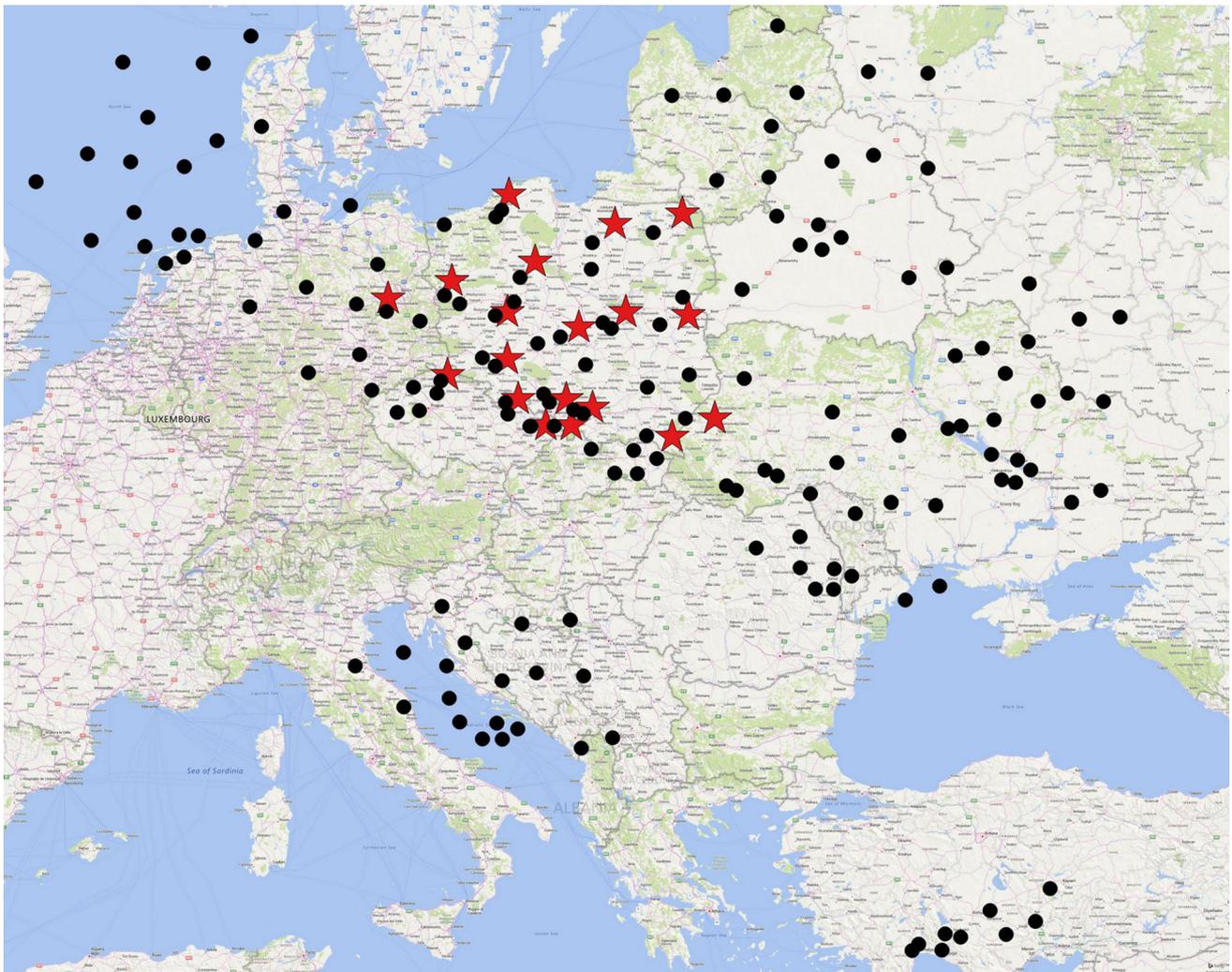


Fig. 10 Exemplary location of observed IPPs (*black dots*) for chosen EPN stations (*red stars*) for epoch 1440 (12:00 UT)

Table 1 Averages of root mean squares for each analysed method (TECU)

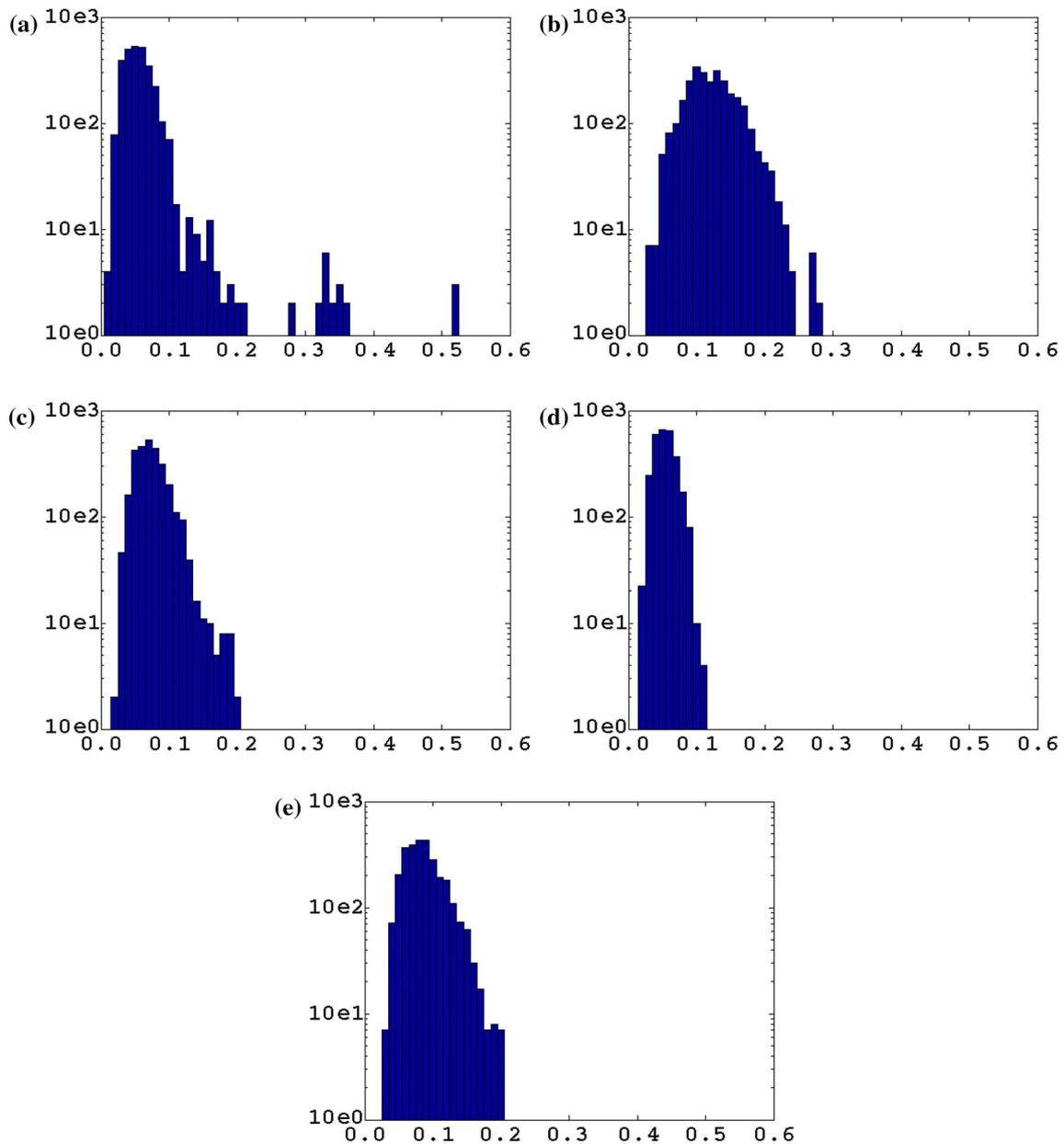
| Interpolation method | Mean | Min | 1st Quartile | Median | 3rd Quartile | Max |
|----------------------|-------|-------|--------------|--------|--------------|-------|
| Natural | 0.063 | 0.013 | 0.040 | 0.055 | 0.069 | 0.789 |
| IDW | 0.123 | 0.033 | 0.097 | 0.122 | 0.148 | 0.279 |
| Quasi-natural | 0.075 | 0.024 | 0.057 | 0.072 | 0.088 | 0.196 |
| Non-Sibsonian | 0.054 | 0.017 | 0.042 | 0.053 | 0.064 | 0.113 |
| Polynomial | 0.088 | 0.032 | 0.066 | 0.084 | 0.103 | 0.200 |

Table 2 Averages of relative errors for each analysed method (%)

| Interpolation method | Mean | Min | 1st Quartile | Median | 3rd Quartile | Max |
|----------------------|-------|-------|--------------|--------|--------------|--------|
| Natural | 0.239 | 0.000 | 0.046 | 0.127 | 0.301 | 24.822 |
| IDW | 0.538 | 0.000 | 0.167 | 0.379 | 0.745 | 7.909 |
| Quasi-natural | 0.327 | 0.000 | 0.105 | 0.233 | 0.433 | 7.263 |
| Non-Sibsonian | 0.242 | 0.000 | 0.066 | 0.156 | 0.318 | 2.950 |
| Polynomial | 0.434 | 0.000 | 0.144 | 0.321 | 0.604 | 4.214 |

Table 3 Averages of standard deviations for each analysed method (TECU)

| Interpolation method | Mean | Min | 1st Quartile | Median | 3rd Quartile | Max |
|----------------------|-------|-------|--------------|--------|--------------|-------|
| Natural | 0.039 | 0.000 | 0.008 | 0.021 | 0.050 | 3.937 |
| IDW | 0.092 | 0.000 | 0.027 | 0.063 | 0.127 | 1.020 |
| Quasi-natural | 0.055 | 0.000 | 0.017 | 0.039 | 0.074 | 0.865 |
| Non-Sibsonian | 0.039 | 0.000 | 0.011 | 0.027 | 0.054 | 0.391 |
| Polynomial | 0.071 | 0.000 | 0.024 | 0.054 | 0.099 | 0.606 |

**Fig. 11** Histograms of root mean squares for **a** natural Sibsonian, **b** IDW, **c** quasi-natural, **d** non-Sibsonian and **e** polynomial methods (with logarithmic scale on the y (vertical) axis)

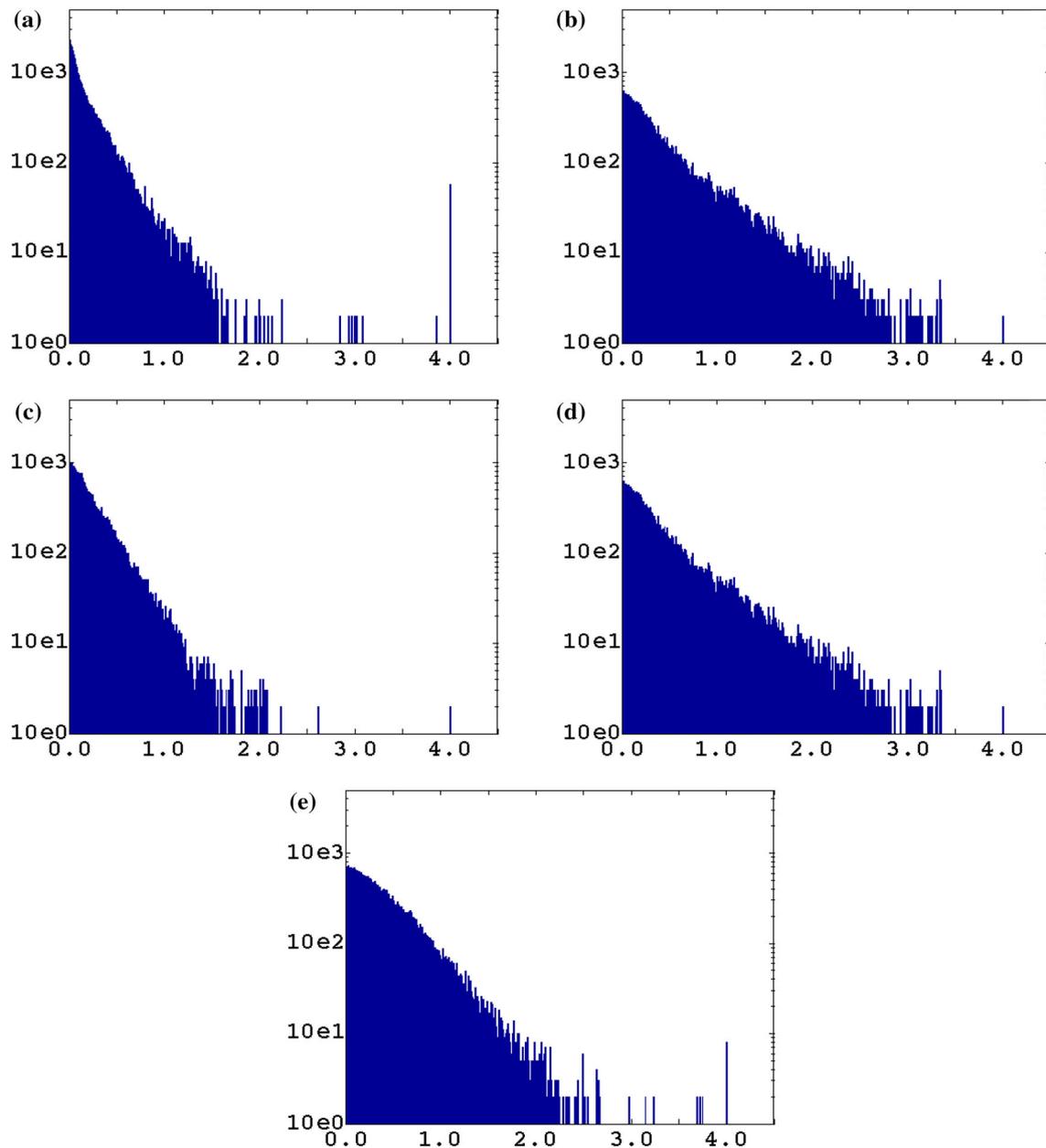


Fig. 12 Histograms of relative errors for **a** natural Sibsonian, **b** IDW, **c** quasi-natural, **d** non-Sibsonian and **e** polynomial methods (with logarithmic scale on the *y* (vertical) axis)

Similarly to the second-scenario Sibsonian method, the non-Sibsonian natural method fails not only in case of an extrapolation situation, but also when the target point is too close to the network border. In this situation the Voronoi cell corresponding to the interpolated point is open (see Fig. 15).

The interpolated point (marked with red boldface) is not actually located outside the network, so one would expect the method would be able to compute it. Nevertheless, the open area of the Voronoi cell is infinite, thus preventing the

computation of the weight coefficients (based on the length of those edges).

Efficiency drops

Another problem with the natural methods is that occasionally the efficiency drops. This affects only the second scenario of the Sibsonian method and is highly correlated with hidden point losses, which were pointed in the previous section.

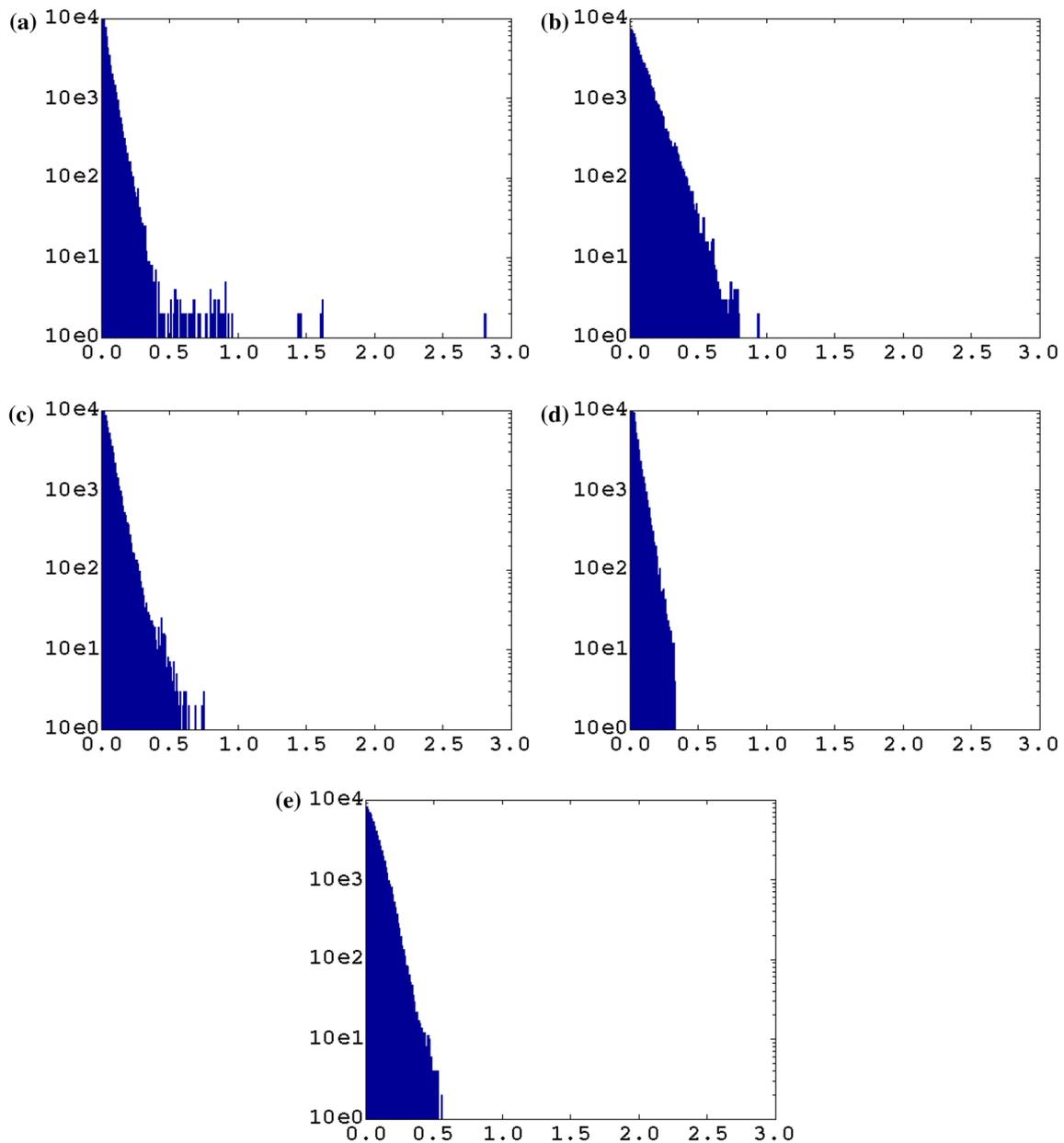


Fig. 13 Histograms of standard deviations for **a** natural Sibsonian, **b** IDW **c** quasi-natural, **d** non-Sibsonian and **e** polynomial methods (with logarithmic scale on the y (vertical) axis)

Table 4 Computation time (in seconds) of each analysed method for the full set of 74,400 points, with the first epoch containing 23 points and one single IPP (Intel i7, 4 GB RAM, 500 GB HDD)

| Interpolation method | Full set (74,400 points) | One epoch (23 points) | One point |
|------------------------|--------------------------|-----------------------|-----------|
| Natural (1st scenario) | 581.2 | 0.1566 | 0.0075 |
| Natural (2nd scenario) | 2020.44 | 0.4773 | 0.0195 |
| IDW | 67.94 | 0.0311 | 0.0208 |
| Quasi-natural | 36.22 | 0.0174 | 0.0008 |
| Non-Sibsonian | 585.79 | 0.1564 | 0.0076 |
| Polynomial | 96.32 | 0.0442 | 0.0192 |

Table 5 Lost points at each epoch (absolute and relative numbers in whole set of 74,400 points)

| Interpolation method | Lost points (nominally) | Lost points (%) |
|------------------------|-------------------------|-----------------|
| Natural (1st scenario) | 16,017 | 21.53 |
| Natural (2nd scenario) | 6 | 0.01 |
| IDW | 0 | 0.00 |
| Quasi-natural | 0 | 0.00 |
| Non-Sibsonian | 72 | 0.10 |
| Polynomial | 0 | 0.00 |

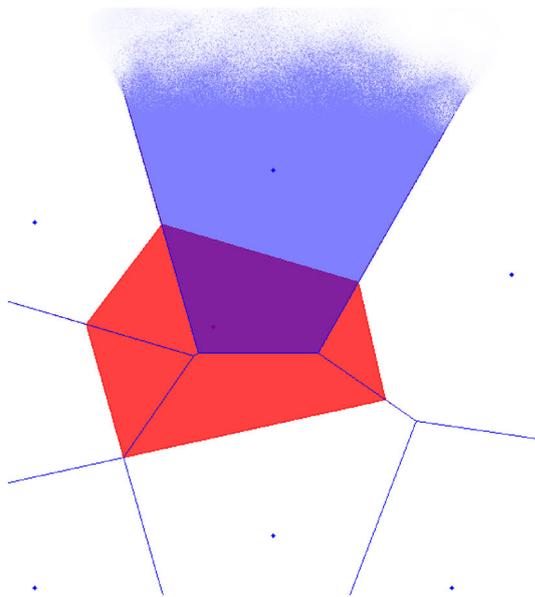


Fig. 14 Open neighbouring cell problem in Sibsonian natural method. Searched overlapping area (*purple*) is computed as the difference of the whole cell area (before putting *red* cell into diagram) and the diminished cell area (*light-blue*)

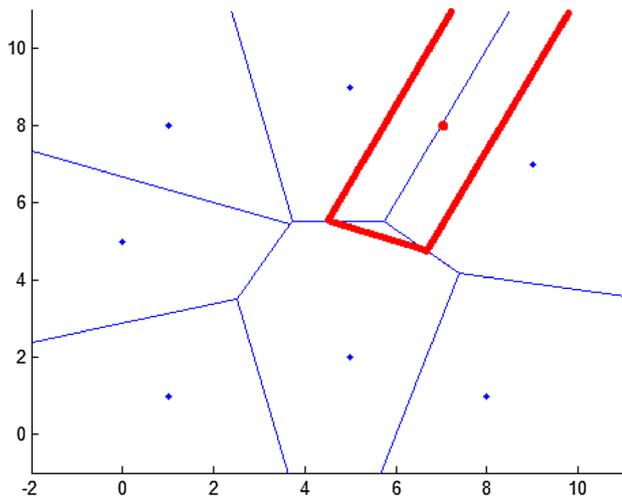


Fig. 15 Open searched cell problem in non-Sibsonian natural method

The same problems affecting those previous two methods also affect the second scenario natural method. In this case, the points are not lost but they provide wrong values. Relative errors and standard deviation values (see Figs. 11, 12, 13) can be clearly seen on histograms of RMS. Accidentally, there occur larger values than in other methods.

First, we should check the problem associated with an open cell for the searched point when using the non-Sibsonian method. Figure 16 contains standard deviations for the 72 points lost in non-Sibsonian natural neighbour interpolation.

As shown, standard deviations (with the maximum value of 3.937 TECU units $-1 \text{ TECU} = 10^{16} \text{ electrons/m}^2$) can reach almost 3 TECU. For some points the standard deviation gets a smaller value, but as their behaviour is impossible to predict, it is much safer to discard the open cell points from the solution (there are very few of them—less than 0.1% of all points) or at last flag them as ‘possible to be wrong’. Detecting such points is very simple to perform and will be shown afterwards.

The second problem—open neighbouring cell problem—is far more complex. To understand its nature, we have to look closer at the geometry of Voronoi cells. Overlapping areas should look like in Fig. 17.

When a neighbouring cell is open (i.e., with at least one of its vertices laying to infinity), the geometry of overlapping areas is changed. The infinity-point is excluded by the algorithm and the polygons are closed thereby ignoring it. This causes the loss of one of the target cell vertices and, consequently, it changes the geometry of the overlapping areas. Such situation is presented in Fig. 18.

This leads to the undesirable reduction of one or more weights. The easiest way to detect that is to compute the relative difference between the sum of overlapping areas

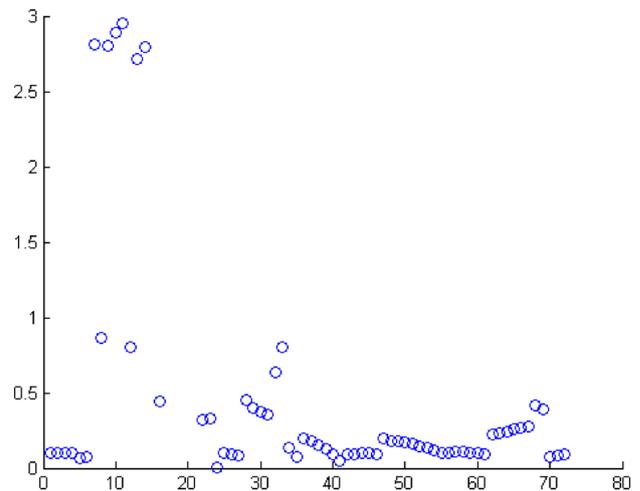


Fig. 16 Standard deviations (in TECU) distribution for 72 open searched cells

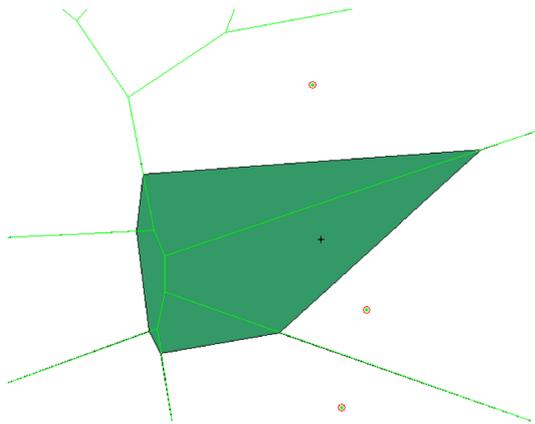


Fig. 17 Properly located overlapping areas

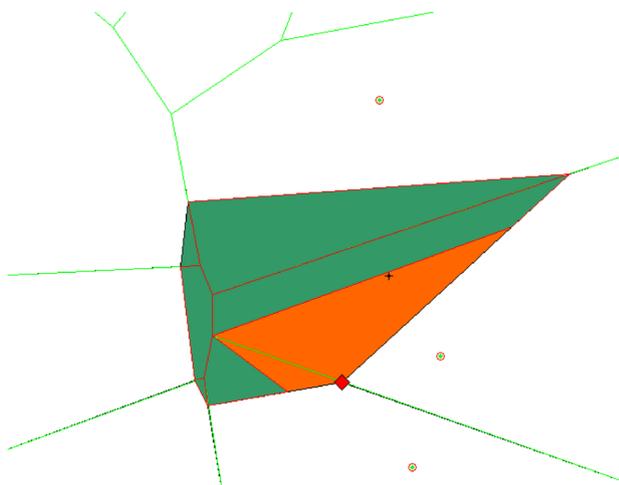


Fig. 18 Deformed overlapping area geometry caused by infinity-vertex problem. Overlapping areas are computed as the *green areas*. Orange ones are lost. Point marked with *red diamond* is excluded because it lies on the infinity-heading line

and the expected value equal to the Voronoi cell of interpolating point, as in Eq. 10.

$$D = (S_x - \Sigma S_i) / S_x \tag{10}$$

where D is the deformation coefficient, S_x is the expected Voronoi cell area and ΣS_i is the sum of overlapping areas (the same method allows to detect points convicted by the previous problem as in such case the deform coefficient will be unable to be computed). Figure 19 illustrates the dependence of standard deviations on the level of deformation.

Table 6 presents the total number of cells with deformed area.

We received over 4600 points (over 6%) with their corresponding deformed cell. Table 7 shows how the statistical distribution of standard deviation changes after excluding all points with uncertain geometry.

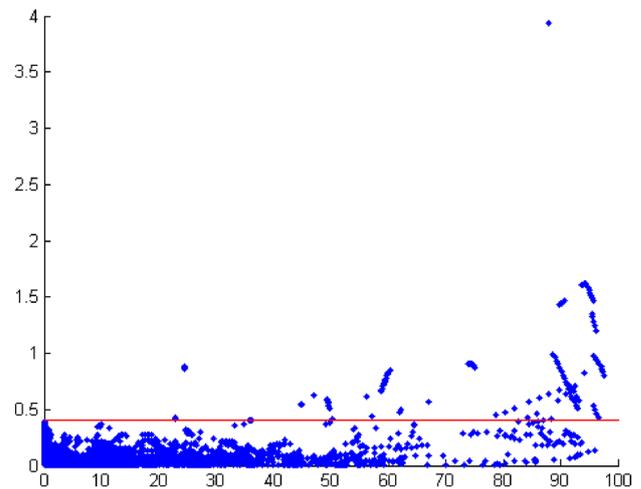


Fig. 19 Dependence of standard deviations (*upwards*; in TECU) on level of deformation (*horizontally*; in percent). The *red line* marks standard deviation level of 0.4 TECU

Table 6 Numbers of points with cells deformed on certain level

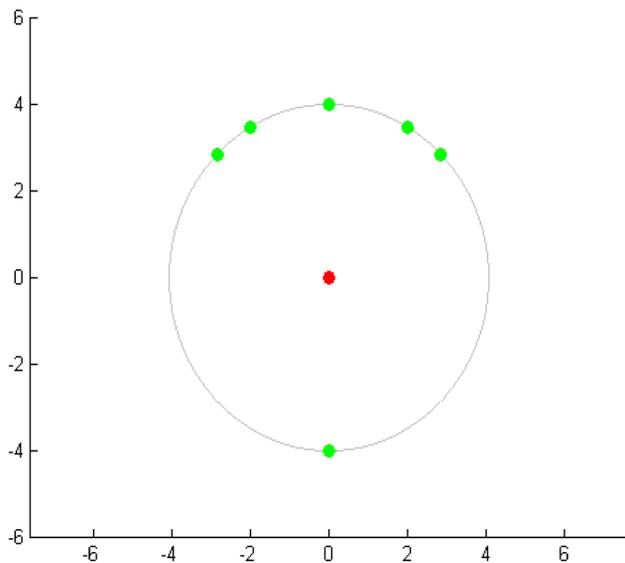
| Deformation | Number of cells | Percentage |
|-------------|-----------------|------------|
| (0.9:1) | 72 | 0.10 |
| (0.8:0.9) | 54 | 0.07 |
| (0.7:0.8) | 22 | 0.03 |
| (0.6:0.7) | 42 | 0.06 |
| (0.5:0.6) | 133 | 0.18 |
| (0.4:0.5) | 146 | 0.20 |
| (0.3:0.4) | 281 | 0.38 |
| (0.2:0.3) | 528 | 0.71 |
| (0.1:0.2) | 1042 | 1.40 |
| (0:0.1) | 2302 | 3.09 |
| Total | 4622 | 6.21 |

It is clearly seen that excluding all cell-deformed points solves the problem with large errors. However, such operations require excluding over 6% of total amount of points. Figure 20 shows that the vast majority of points with standard deviation over 0.4 TECU have cells deformed for more than 40%. As can be seen in Table 8, there are 469 of such points, which is only 10% of all excluded points. Table 8 also presents how the standard deviation distribution would look like if we exclude only those points with deformation coefficient above 40%.

The maximum value of standard deviation is not as low as when all cell-deformed points were excluded, but still is reduced from almost 4 TECU to less than 1 TECU and less than 1% of points had to be excluded. This shows that not only flagging suspected points, but also giving them some kind of risk factor may be a good idea.

Table 7 Natural neighbour method standard deviations distribution after excluding all deformed and open cell points

| Mean | Min | 1st quartile | Median | 2nd quartile | Max | Lost points | Open cells points | Deformed cells points | Excluded points total | Percentage |
|-------|-------|--------------|--------|--------------|-------|-------------|-------------------|-----------------------|-----------------------|------------|
| 0.036 | 0.000 | 0.008 | 0.021 | 0.047 | 0.396 | 6 | 72 | 4622 | 4694 | 6.31 |

**Fig. 20** Exemplary set of equally distant neighbours

It should also be mentioned that this problem is not mathematical. It is caused by limits of computers capabilities—overlapping areas are deformed due to the infinity-computing trouble. This problem could be also solved by including boundaries and computing Voronoi diagrams inside the closed area, but, despite this solution, it requires using a special function, which is not so straightforward and, what is more important, takes over seven times longer to compute (the classic Voronoi based natural neighbour method took about 30 min; see Table 4). Regarding the fact that such a problem occurs rather rarely and only in situations when the dataset is not sufficient, the use of Voronoi limiting function should be considered for each interpolation case.

Final discussion and conclusion

We have presented a comparison of several interpolation methods, namely two methods based on distance-weights (IDW and Quasi-natural), two Voronoi-based

methods (natural and non-Sibsonian) and a polynomial method.

For the comparison we have used STEC data computed with UPC TOMION software from GPS ground-receivers from EUREF network over Poland. The 30-s interval dataset spans a period of 24 h from 15 June 2015, during quiet geomagnetic conditions.

The natural neighbour method, in both Sibsonian and non-Sibsonian approaches, provides relative errors over two times smaller than the IDW method (mean value of 0.239, 0.242 and 0.538%, respectively) and almost two times smaller than the outcomes of the polynomial interpolation with mean value of 0.434% (see Table 2). Similar improvement can be seen in standard deviations—0.029 TECU for natural neighbour versus 0.092 TECU for IDW and 0.071 TECU for polynomial interpolation (see Table 3). Highly topology-dependent natural neighbour methods provide the best accuracy as the TEC value depends on its own topology. However, the Sibsonian method in its both scenarios has to cope with many problems and point losses. Moreover, the more effective second scenario-based method requires relatively high computation time (especially when opposed to ultra-rapid quasi natural neighbour method). The solution to the trade-off between accuracy and computational time might be the non-Sibsonian method, which is also very topology-dependent, but least sensitive to the open-cell problem. Another solution may be the quasi-natural method, which provides slightly worse, but still quite promising results (see Tables 1, 2, 3). This method is only partially dependent on the topology, which leads to the drop of accuracy, but, on the other hand, it remains almost unaffected by all topology-caused problems.

Also, a comparison between natural neighbour interpolation and Kriging was considered and will be discussed in future works.

When considering accuracy, the key issue is the way interpolation methods tackle different topologies. To illustrate this problem, let us consider the situation of a

Table 8 Natural neighbour method standard deviations distribution after excluding open cell points and those with corresponding cells deformed for more than 40%

| Mean | Min | 1st quartile | Median | 2nd quartile | Max | Lost points | Open cells points | Deformed cells points | Excluded points total | Percentage |
|-------|-------|--------------|--------|--------------|-------|-------------|-------------------|-----------------------|-----------------------|------------|
| 0.036 | 0.000 | 0.008 | 0.021 | 0.048 | 0.888 | 6 | 72 | 469 | 541 | 0.73 |

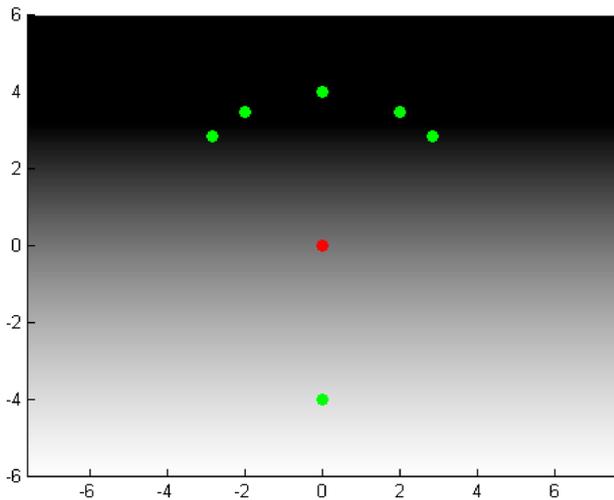


Fig. 21 Exemplary values gradient. Units are arbitrary

point surrounded by six neighbours located in exactly the same distance from the point, like in Fig. 20.

First, we should consider how the weights would look like in the pure distance-based methods (IDW and quasi-natural methods). In quasi natural method all the weights will be equal to 0.16, as all six points are natural neighbours and they are at the same distance of the interpolated point. The IDW method—which depends on the defined number of neighbours—may behave like quasi-natural method but only if the number of neighbours is equal to six. In other situations (neighbours number different than six), IDW method will discard randomly one or more neighbours or, if defined neighbours number is too large, take some extra, non-neighbouring but nearby points into calculation.

To illustrate the influence of such weight distribution, we should assign some values to the points. Let us consider the example shown in Fig. 21, where there are six values distributed with a linear gradient.

The interpolated point is right in the middle, so it should get a value approximately equal to the mean value of the points at the top and that at the bottom. Nevertheless, in quasi-natural and IDW methods all weights have the same value, thereby biasing the result in favour of the most populated cluster of points at the top.

Figure 22 shows the Voronoi diagram for those points in Fig. 21. The shape of the Voronoi cell (for the non-Sibsonian method) and overlapping areas structure (for the Sibsonian method) display how the influence of the points at the top over the interpolated point is diminished with respect to the one at the bottom. Since the weight coefficients are related to the size of the Voronoi cell and overlapping areas (Eqs. 5 and 7, respectively), the smaller the areas and Voronoi cells in the cluster of points, the lower the values of the weight coefficients are.

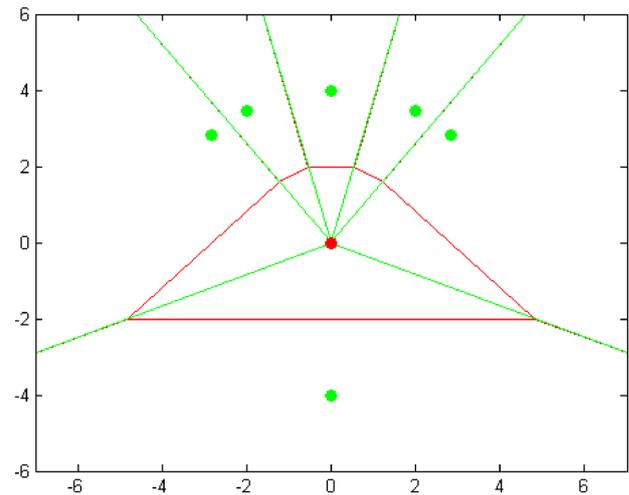


Fig. 22 Two-ordered Voronoi diagram for the set of points. Units are arbitrary

To summarize, by taking into account the topology, the result is more precise. This simple example helps to explain why natural neighbour methods perform better with non-homogeneously distributed datasets than methods solely based on relative distances.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Belikov VV, Ivanov VD, Kontorovich VK, Korytnik SA, Semenov AYU (1997) The non-Sibsonian interpolation: a new method of interpolation of the value of a function on an arbitrary set of points. *Comput Math Math Phys* 37(1):9–15
- Buresova D, Nava B, Galkin I, Angling M, Stankov SM, Coisson P (2009) Data ingestion and assimilation in ionospheric models. *Ann Geophys* 52(3/4):235–253
- Choi B-K, Lee W-K, Cho S-K, Park J-U, Park P-H (2010) Global GPS ionospheric modelling using spherical harmonic expansion approach. *J Astron Space Sci* 27(4):359–366. doi:10.5140/JASS.2010.27.4.359
- Dąbrowski BP, Krankowski A, Błaszkiwicz L, Rothkaehl H (2016) Prospects for solar and space weather research with polish part of the LOFAR telescope. *Acta Geophys* 64(3):825–840. doi:10.1515/acgeo-2016-0028
- Dumitru PD, Plopeanu M, Badea D (2013) Comparative study regarding the methods of interpolation. In: 1st European conference geodesy and geomatics engineering GENG'13 “Recent Advances in Geodesy and Geomatics Engineering”, 8–10 October, 2013, Antalya, Turkey, pp 45–52
- Fortune S (1995) Voronoi diagrams and Delaunay triangulations. In: Toth CD, O’Rourke J, Goodman JE (eds) *Handbook of discrete and computational geometry*. CRC Press Inc, Boca Raton, pp 377–388

- Hajj GA, Wilson BD, Wang C, Pi X, Rosen IG (2004) Data assimilation of ground GPS total electron content into a physics-based ionospheric model by use of the Kalman filter. *Radio Sci* 39:RS1S05. doi:[10.1029/2002RS002859](https://doi.org/10.1029/2002RS002859)
- Han D, Yun H, Kee C (2013) Modeling of ionospheric delay for SBAS using spherical harmonics functions. *Trans Nav Int J Marine Nav Safety Sea Transport* 7(2):205–209. doi:[10.12716/1001.07.02.07](https://doi.org/10.12716/1001.07.02.07)
- Harman C, Johns M (2008) Voronoi natural neighbors interpolation. In: *Proceeding of class of 2008 senior conference on computational geometry*, Swarthmore College, Swarthmore, USA, 49–53
- Hernandez-Pajares M, Juan JM, Sanz J (1999) New approaches in global ionospheric determination using ground GPS data. *J Atmos Sol -Terr Phys* 61:1237–1247
- Hernández-Pajares M, Juan JM, Sanz J (1997) Neural network modeling of the ionospheric electron content at global scale using GPS data. *Radiol Sci* 32(3):1081–1089
- Hernández-Pajares M, Juan JM, Sanz J, Colombo OL (2000) Application of ionospheric tomography to real-time GPS carrier-phase ambiguities resolution at scales of 400–1000 km and with high geomagnetic activity. *Geophys Res Lett* 27(13):2009–2012
- Hernández-Pajares M, Juan JM, Sanz J, Aragón-Ángel A, García-Rigo A, Salazar D, Escudero M (2011) The ionosphere: effects, GPS modeling and the benefits for space geodetic techniques. *J Geodyn* 85(12):887–907. doi:[10.1007/s00190-011-0508-5](https://doi.org/10.1007/s00190-011-0508-5)
- Jakowski N, Wilken V, Schlueter S, Stankov S, Heise S (2005) Ionospheric space weather effects monitored by simultaneous ground and space based GNSS signals. *J Atmos Sol-Terr Phys* 67(12):1074–1084
- Komjathy A, Wilson B, Pi X, Akopian Y, Dumett D, Iijima B, Verkhoglyya-dova O, Mannucci AJ (2010) JPL/USC GAIM: on the impact of using COSMIC and ground-based GPS measurements to estimate ionospheric parameters. *J Geophys Res* 115:A02307. doi:[10.1029/2009JA014420](https://doi.org/10.1029/2009JA014420)
- Krankowski A, Shagimuratov II, Baran LW, Yakimova G (2007) The structure of the mid- and high-latitude ionosphere during the November 2004 storm event obtained from GPS observations. *Acta Geophys* 55(4):490–508. doi:[10.2478/s11600-007-0033-3](https://doi.org/10.2478/s11600-007-0033-3)
- Mandrake L, Wilson B, Wang C, Hajj G, Mannucci A, Pi X (2005) A performance evaluation of the operational Jet Propulsion Laboratory/University of Southern California Global Assimilation Ionospheric Model (JPL/USC GAIM). *J Geophys Res* 110:A12306. doi:[10.1029/2005JA011170](https://doi.org/10.1029/2005JA011170)
- Moore EH (1920) On the reciprocal of the general algebraic matrix. *Bull Am Math Soc* 26(9):394–395
- Orus R, Hernandez-Pajares M, Juan JM, Sanz J (2005) Improvement of global ionospheric VTEC maps by using kriging interpolation technique. *J Atmos Sol-Terr Phys* 67:1598–1609
- Pi X, Wang C, Hajj GA, Rosen G, Wilson BD, Bailey GJ (2003) Estimation of $E \times B$ drift using a global assimilative ionospheric model: an observation system simulation experiment. *J Geophys Res* 108(A2):1075. doi:[10.1029/2001JA009235](https://doi.org/10.1029/2001JA009235)
- Pi X, Had GA, Wilson BD, Mannucci AJ, Komjathy A, Mandrake L, Wang C (2004). 3-Dimensional assimilative ionospheric modelling for regions of large TEC gradient. In: *Proceedings of ION 2004 national technical meeting*, Satellite Division of the Institute of Navigation, 26–28 January 2004, San Diego, USA
- Pi X, Mannucci AJ, Iijima BA, Wilson BD, Komjathy A, Runge TF, Akopian V (2009) Assimilative modeling of ionospheric disturbances with FORMOSAT-3/COSMIC and ground-based GPS measurements. *Terr Atmos Ocean Sci* 20(1):273–285. doi:[10.3319/TAO.2008.01.04.01\(F3C\)](https://doi.org/10.3319/TAO.2008.01.04.01(F3C))
- Schaer S (1999) Mapping and predicting the earth's ionosphere using the global positioning system. Ph.D. thesis, Astronomical Institute, University of Bern, Switzerland
- Scherliess L, Schunk RW, Sojka JJ, Thompson DC, Zhu L (2006) Utah State University global assimilation of ionospheric measurements Gauss-Markov Kalman filter model of the ionosphere: model description and validation. *J Geophys Res* 111:A11315. doi:[10.1029/2006JA011712](https://doi.org/10.1029/2006JA011712)
- Schunk RW (2002) Global assimilation of ionospheric measurements (GAIM). In: *Proceedings of international ionospheric effects symposium*, 7–9 May 2002, Alexandria, USA
- Sibson R (1981) A brief description of natural neighbor interpolation. In: Barnett V (ed), *Interpreting multivariate data*, Wiley, Chichester, pp 21–36
- Strangeways HJ, Kutiev I, Cander LR, Kouris S, Gherm V, Marin D, De La Morena B, Pryse SE, Perrone L, M. Pietrella, S. Stankov, L. Tomasik, Tulunay E, Tulunay Y, Zernov N, Zolesi B (2009) Near-Earth space plasma modelling and forecasting. *Ann Geophys* 52(3/4):255–271. doi:[10.4401/ag-4579](https://doi.org/10.4401/ag-4579)
- Sukumar N, Moran B, Semenov AY, Belikov VV (2001) Natural neighbour Galerkin methods. *Int J Numer Meth Eng* 50:1–27
- Wang C, Hajj G, Pi X, Rosen IG, Wilson B (2004) Development of the global assimilative ionospheric model. *Radio Sci* 39:RS1S06. doi:[10.1029/2002RS002854](https://doi.org/10.1029/2002RS002854)

Interference of an upstream pier on local scour at downstream piers

Shivakumar Khaple¹ · Prashanth Reddy Hanmaiahgari¹ · Roberto Gaudio² · Subhasish Dey¹

Received: 10 January 2017 / Accepted: 16 January 2017 / Published online: 31 January 2017
© Institute of Geophysics, Polish Academy of Sciences & Polish Academy of Sciences 2017

Abstract In this study, three kinds of pier arrangements were tested. They are (i) two piers in tandem, (ii) two piers in staggered arrangement, and (iii) three piers in symmetrically staggered arrangements. In the arrangement of two piers in tandem, the equilibrium scour depth at downstream pier decreases with an increase in downstream distance up to approximately eight times pier diameter and then increases with further increase in downstream distance. However, the scour depth at downstream pier is always smaller than that at upstream pier. In the arrangement of two staggered piers, the scour depth at the downstream pier for $L/b = 4$, where L is the offset distance and b is the pier diameter, is the same as that of the upstream pier at $S = 8b$, where S is the streamwise spacing or distance between piers. Further, for three piers in staggered arrangement, as the lateral spacing between downstream piers increases, the equilibrium scour depth at downstream pier decreases.

Keywords Bridge piers · Erosion · Local scour · Pier interference · Scour

List of symbols

| | |
|-----------|--|
| B | Channel width |
| b | Pier diameter |
| d_s | Instantaneous scour depth |
| d_{se} | Equilibrium scour depth at an isolated single pier |
| d_{s1} | Equilibrium scour depth at upstream pier |
| d_{s2} | Average value of equilibrium scour depth at downstream piers |
| d_{s2a} | Equilibrium scour depth at downstream pier (right pier) |
| d_{s2b} | Equilibrium scour depth at downstream pier (left pier) |
| d_{s2c} | Computed scour depth at downstream pier |
| d_{s2m} | Measured scour depth at downstream pier |
| d_{16} | 16% finer sediment size |
| d_{50} | Median sediment size |
| d_{84} | 84% finer sediment size |
| F | Flow Froude number |
| g | Gravitational acceleration |
| h | Flow depth |
| k_e | Equivalent roughness height |
| K_T | Spacing factor in tandem arrangement |
| L | Lateral offset distance |
| R_p | Pier Reynolds number |
| S | Streamwise spacing |
| s | Relative density of sediments |
| T | Time scale |
| T^* | $T(\Delta g d_{50}^3)^{0.5}/b^2$ |
| t | Time |
| U | Approaching flow velocity |
| U_c | Critical flow velocity |
| U_{*c} | Critical shear velocity |
| u | Time-averaged streamwise velocity |
| x | Streamwise distance |

✉ Subhasish Dey
sdey@iitkgp.ac.in

Shivakumar Khaple
shivabku@gmail.com

Prashanth Reddy Hanmaiahgari
hpr@civil.iitkgp.ernet.in

Roberto Gaudio
gaudio@unical.it

¹ Department of Civil Engineering, Indian Institute of Technology Kharagpur, Kharagpur, West Bengal, India

² Dipartimento di Ingegneria Civile, Università della Calabria, Rende, CS, Italy

| | |
|------------|---|
| \hat{x} | Nondimensional streamwise distance ($=x/b$) |
| z | Vertical flow depth |
| \hat{z} | Nondimensional vertical distance ($=z/b$) |
| Δ | $s - 1$ |
| Θ_c | Threshold shields parameter |
| θ | Angle between upstream and downstream piers |
| ρ | Mass density of water |
| σ_g | Geometric standard deviation of sediment size |
| τ_c | Threshold bed shear stress |
| ϕ | Angle of repose of sediment |

Introduction

The obstruction to the flowing stream by a bridge pier causes a three-dimensional separation of approaching flow forming a vortical flow field around the pier, resulting in local scour at pier (Dey et al. 1995). Pier scour is a challenging problem for the hydraulic, civil, bridge engineers, etc. Failure of bridges due to scour at pier foundations is a common occurrence. Therefore, it becomes a topic of continued interest, especially for the hydraulic engineers. As a result, many researchers worked on this problem and suggested a sizable number of pier scour prediction formulae or design equations. Among them, Laursen and Toch (1956), Breusers et al. (1977), Melville (1997), Cardoso and Bettess (1999), Melville and Coleman (2000), and Dey and Raikar (2007) and so on, conducted experiments and put forward different relations for the estimation of the equilibrium scour depth at various shaped piers. Unfortunately, scour depth predictions at piers from the available design equations are widely varying, when they are applied to various prototypical sites.

Based on the mode of sediment transport by the approaching flow, Chabert and Engeldinger (1956) classified local scour into two categories, namely, clear water scour and live bed scour. Clear water scour takes place in absence of sediment transport by the approaching flow into the scour hole. On the other hand, live bed scour occurs when the scour hole is continuously fed with the sediments by the approaching flow. It was observed that the maximum clear water scour depth is by approximately 10% greater than the live bed scour depth. Regarding the mechanism of pier scour, the components of the near-pier flow field are the downflow, horseshoe vortex, wake vortices and bow wave, as shown in Fig. 1. The approaching stream flow gradually becomes zero at the vertical line in the upstream face of the pier.

The approaching velocity profile along the vertical depth decreases downward, has a maximum at the free surface and zero at the bed. It causes a decrease in the stagnation pressure in the downward direction at the upstream face of the pier and consequently the downflow is developed. The

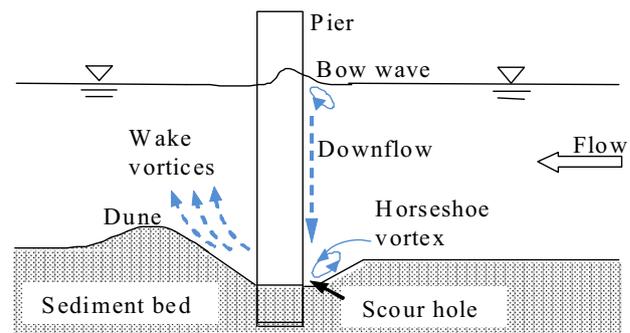


Fig. 1 Definition sketch of pier scour showing the flow field

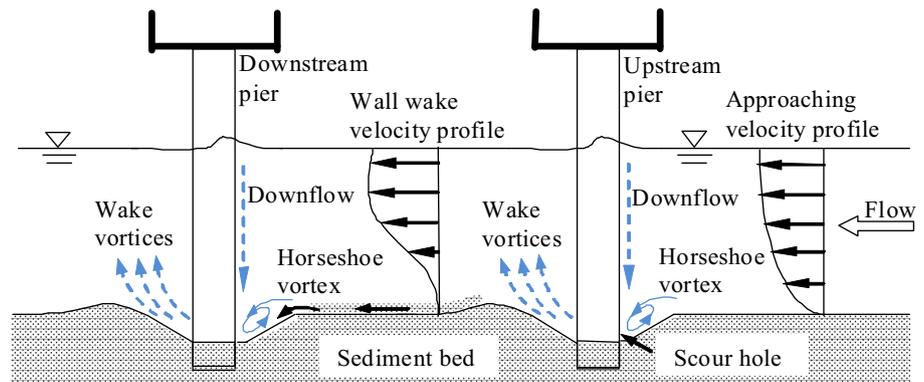
horseshoe vortex is developed owing to the separation of flow that takes place at the upstream face of the pier and it travels by the sides of the pier (Raudkivi 1986). As a result, a scour hole is developed at pier owing to the removal of bed sediments. According to Dey (1991) and Dey et al. (1995), the horseshoe vortex is formed owing to the diving mode of approaching flow into the scour hole to cause scour. A bow wave is formed due to obstruction to the flowing stream at the upstream free surface adjacent to the pier and it rotates in the opposite direction of the horseshoe vortex. The bow wave is rather innocent, not participating as a scour agent.

Temporal variation of scour depth at a single pier was studied extensively by several investigators (Yanmaz and Altinbilek 1991; Sumer et al. 1993; Dey 1999; Mia and Nago 2003; Oliveto and Hager 2005). The time required to reach an equilibrium clear water scour depth was found to be extensively long; it can even reach a month or more, because the equilibrium scour depth is achieved asymptotically with time (Breusers et al. 1977; Melville and Chiew 1999).

Some researchers, however, studied the temporal variation of scour at compound bridge piers (Ettema 1980; Melville and Raudkivi 1996; Melville and Chiew 1999; Oliveto and Hager 2002; Kumar et al. 2003; Chang et al. 2004; Yanmaz 2006; Lu et al. 2011; Kothiyari and Kumar 2012; Ferraro et al. 2013; Moreno et al. 2015). Melville and Raudkivi (1996) developed design relationships and introduced the concept of an effective pier size for the estimation of maximum scour depth at nonuniform piers. Lu et al. (2011) proposed a semiempirical model to compute the temporal variation of scour depth at nonuniform circular piers with unexposed foundations. Kothiyari and Kumar (2012) introduced a mathematical model to compute the temporal variation of scour depth at circular compound bridge piers for all possible cases of footing positions with respect to the general level of the bed.

Besides, rapid urbanization and increased traffic volume cause a need for construction of new bridges next to the existing ones, thereby creating problems of pier interference. In addition, for economic, design and safety reasons,

Fig. 2 Schematic of scour holes at upstream and downstream piers showing the possible flow field



twin piers are favored; therefore, safe and economic design of bridge pier requires an accurate prediction of design scour depth at them, accounting also for the upstream pier interference on scour at the downstream pier.

This topic has, however, been not studied extensively. Moreover, the existence of parallel railway-bridge and road-bridge or a newly constructed bridge by the side of an old bridge gives rise to a situation where interference of the presence of an upstream pier and scour at it on scour at a downstream pier is prevalent (Fig. 2). Upstream pier could primarily influence the scour at a downstream pier in two ways: (i) upstream pier modifies the flow field in the wake flow region that acts as an approaching flow for the downstream pier to scour, and (ii) the sediment that is scoured from the upstream pier is being fed to the scour hole at the downstream pier. As a result, equilibrium scour depth at the downstream pier is less than that of the upstream pier.

In two inline bridge pier interaction, the presence of the two piers can generate a complex interaction in the hydrodynamic characteristics of the flow field; hence, the scour mechanism around twin piers is different from that around a single pier. Scour at twin piers were experimentally studied by Hannah (1978), Elliott and Baker (1985), Beg and Beg (2015), and Wang et al. (2016).

Elliott and Baker (1985) studied the effect of pier spacing on scour at bridge piers and concluded that the design equation given by Breusers et al. (1977) was derived only for a particular set of parameters and its use in other conditions can only be carried out with caution. Beg and Beg (2015) studied the effects of the interference of two bridge piers with unequal size in tandem arrangement by varying the pier spacing in clear water condition. They found that the scour depth at downstream pier is minimum at a spacing of $35b$ (where $b = 33$ mm). Wang et al. (2016) experimentally investigated the characteristics of scour at twin circular piers, by varying the spacing between the piers and the flow depth. They observed that the scour depth at the downstream pier was lesser than that at the upstream pier.

The literature review reveals that significantly minimal research has been undertaken on the equilibrium scour depth prediction and temporal variation of scour depth for the cases of: (i) two piers in tandem, (ii) two piers in staggered, and (iii) three piers in symmetrically staggered arrangements.

The present study therefore aims at performing a detailed experimental investigation on the influence of an upstream (front) pier on scour at downstream (rear) piers. The scour at downstream piers in the presence of an upstream pier is a function of the alignment (tandem or staggered) of upstream and downstream piers, their relative sizes and spacing between them. The experimental campaign was carried out on scour at two inline piers, two and three piers in staggered arrangement under steady flow, clear water scour condition. The equilibrium scour depth was quantified at upstream and downstream piers for both tandem and staggered arrangements. The formulae for computing the temporal variation of scour depths at downstream piers are empirically developed for different pier arrangements.

Experimental methodology and procedure

Experiments were conducted in a 15.50 m long, 0.91 m wide, and 0.70 m deep flume with a longitudinal slope of 0.02%, located in the Hydraulic and Water Resources Engineering Laboratory at the Department of Civil Engineering, Indian Institute of Technology Kharagpur, India. The test section was located at 8 m from the inlet in order to generate a fully developed turbulent flow (as verified with velocity measurements). The sidewalls of the test section were made of 4.5 m long transparent glass to visualize the flow and the scour process. At the inlet, a concrete stilling basin, consisting of a perforated baffle wall and two vertical steel screens was present, covering the entire cross section to damp the flow disturbance. At the outlet, an adjustable tailgate was installed in order to regulate the flow depth. The flume was connected to a re-

circulating water supply system in the laboratory. To minimize the amount of sediment required for the tests, a 0.2 m high false floor was constructed upstream of the test section, forming a 2.5 m long and 0.2 m deep sediment recess, which was filled with uniform sediment. The same type of sediment used in test section was glued onto the top surface of the false floor. The piers were embedded in the middle of the sediment recess box. Three small holes were provided at the bottom of the downstream wall of the sediment recess to drain out the water for drying the sediment bed at the end of each test. A sediment trap was constructed at the downstream wall of the sediment recess box, having a clear length of 2.5 m to arrest the scoured sediments. Figure 3 shows the schematic of the experimental setup.

The discharge was measured with a calibrated V-notch weir at the inlet tank. A Vernier point gauge was used to measure the water surface level in the flume. All the experiments started with a perfectly leveled sediment bed. The minimum duration of the each experimental run was 36 h to ensure reaching the equilibrium scour depth, when the rate of increase in scour depth did not exceed 5% of the pier diameter in 24 h (Melville and Chiew 1999). Equilibrium scour depth was measured by the point gauge in the drained condition of the sediment bed at the end of the experiment. However, temporal variations of scour depths were recorded during the experimental run over a period of time.

Two different uniform sediments, with median sizes $d_{50} = 0.96$ mm and 1.8 mm, respectively, were used in the experiments. The properties of the uniform sediments used in the experiments are given in Table 1. The geometric standard deviation σ_g , given by $(d_{84}/d_{16})^{0.5}$, was less than 1.4, indicating uniform sediments (Dey et al. 1995).

The piers were fixed on the base of the flume at desired locations. The upstream pier was located at 8.5 m from the inlet, whereas the downstream pier/piers was/were located downstream of the upstream pier at different spacing $S = 2 - 12b$. In this study, two identical circular piers

were considered for a test. Two different pier diameters (6 and 8.2 cm) were tested.

The experiments were conducted for scour at a single pier. The results are presented in Table 2 with various hydraulic conditions and dimensionless groups, where h is the flow depth and B is the channel width. Figure 4 shows the photographs of the experimental setup during and after the experimental run.

In Fig. 4a, the light traces the separation of flow due to the obstruction to the flow from the upstream pier as it travels by the sides of the pier to the downstream. Upstream pier modifies the flow field downstream of it, forming the wake flow, which acts as an approaching flow for the inline downstream pier to scour. Initially, the rate of scour evolution at upstream and downstream piers were almost same and, after a certain time interval, the scour hole at the downstream pier is started feeding the sediment that is scoured from the upstream pier. The approaching flow velocity at the downstream pier is less as compared to the approaching flow velocity at the upstream pier; hence, the downstream pier scour depth is less as compared to upstream scour depth. Figure 4b shows the photograph of the contours of equilibrium scour at upstream and downstream piers with spacing between them.

The flow velocity was set as $U = 0.9U_c$ (clear water condition), where U is the approaching flow velocity and U_c is the critical approaching flow velocity for the inception of sediment motion. The critical shear velocities U_{*c}

Table 1 Characteristics of sediments used in experiments

| d_{50} (mm) | s | σ_g | ϕ (°) |
|---------------|------|------------|------------|
| 0.96 | 2.65 | 1.21 | 27 |
| 1.80 | 2.65 | 1.18 | 30 |

d_{50} median diameter of sediments, s relative density of sediments, σ_g geometric standard deviation of particle size distribution, and ϕ angle of repose of sediments

Fig. 3 Schematic of experimental setup in a laboratory flume

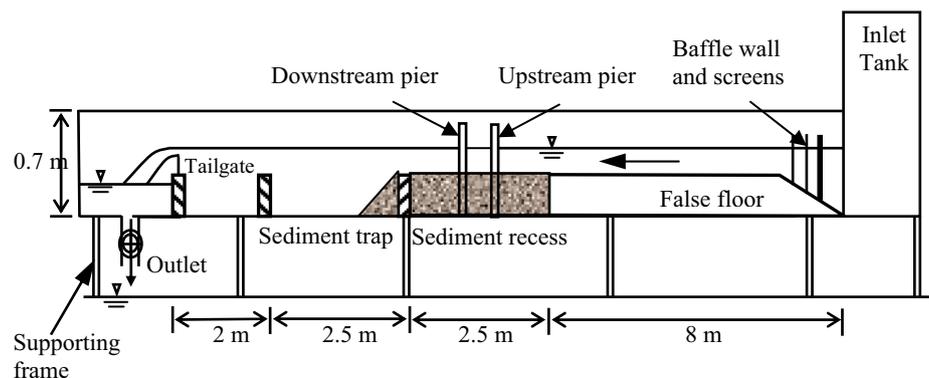


Table 2 Experimental conditions and dimensionless groups

| Run | b (cm) | d_{50} (mm) | h (cm) | U (m/s) | U_c (m/s) | U_{*c} (m/s) | b/d_{50} | B/b | B/h | F | h/d_{50} | R_p | d_{se} (cm) |
|-----|----------|---------------|----------|-----------|-------------|----------------|------------|-------|-------|------|------------|-------|---------------|
| A01 | 6.0 | 0.96 | 16 | 0.346 | 0.384 | 0.0203 | 62.5 | 15.2 | 5.7 | 0.27 | 166.7 | 20760 | 12.85 |
| A02 | 8.2 | 0.96 | 16 | 0.346 | 0.384 | 0.0203 | 85.4 | 11.1 | 5.7 | 0.27 | 166.7 | 28372 | 15.10 |
| B01 | 6.0 | 1.80 | 16 | 0.472 | 0.524 | 0.0305 | 33.2 | 15.2 | 5.7 | 0.37 | 88.4 | 38320 | 13.20 |
| B02 | 8.2 | 1.80 | 16 | 0.472 | 0.524 | 0.0305 | 45.3 | 11.1 | 5.7 | 0.37 | 88.4 | 38704 | 15.30 |

b pier diameter, d_{50} median diameter of sediments, h flow depth, U approaching flow velocity, U_{*c} critical shear velocity, U_c critical flow velocity, B channel width, F flow Froude number, R_p pier Reynolds number, d_{se} equilibrium scour depth at an isolated single pier

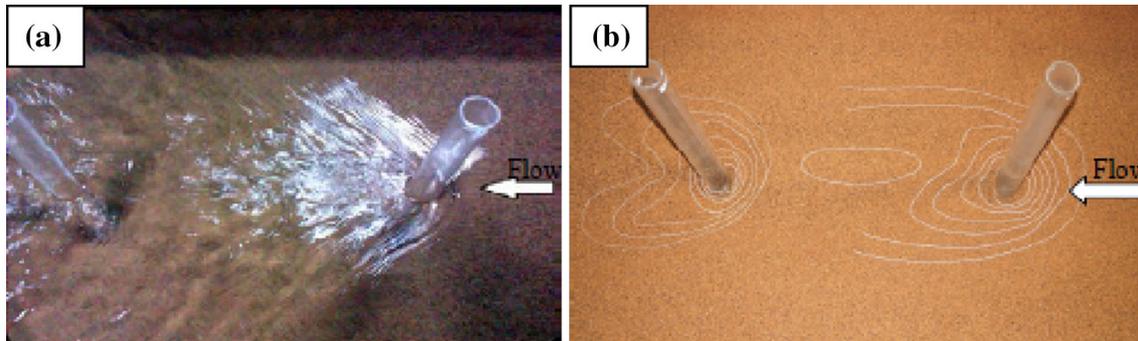


Fig. 4 Photographs of the experimental setup of two piers in tandem arrangement: **a** during and **b** after the experimental run (flow direction from right to left)

for sediments used in this study are furnished in Table 2. The values of U_{*c} were determined from the Shields diagram. The corresponding critical velocities U_c were calculated using the logarithmic equation of average velocity for a rough bed, as was used by Lauchlan and Melville (2001). The logarithmic equation of average velocity is as follows:

$$\frac{U_c}{U_{*c}} = 5.75 \log \left(\frac{h}{k_e} \right) + 6, \quad (1)$$

where h is the flow depth and k_e is the equivalent roughness height ($=2d_{50}$).

Temporal scour depths during the experimental runs were measured at the upstream nose of the piers at a regular interval of time with the help of a point gauge. On the completion of the experimental run, the supply of water to the flume was gradually stopped and water from the flume was drained out carefully so that the scour hole and pattern at the piers remained undisturbed. The experimental conditions shown in Table 3, developed by several researchers, were satisfied.

In addition, the velocities at the upstream and the downstream locations of a circular pier of diameter $b = 6$ cm over a rigid bed was measured by an Acoustic Doppler Velocimeter (ADV) to have an understanding about the wall wake flow behind a circular pier. The flow condition was similar to Run A01. It may be noted that this

wall wake flow became an approaching flow for a downstream pier in a tandem arrangement.

Results and discussion

Two piers in tandem arrangements

Visualization of scour process and wall-wake velocity

To visualize scour process in tandem arrangement, equivalent run for tandem arrangement Run A01 was chosen. In this experiment, initially, water was filled slowly to avoid sudden erosion of loose sediment at piers. After filling the flume completely, approaching flow velocity was gradually increased from very low to the required average velocity. As the flow velocity reached about 50% of the critical velocity U_c , dislodgement of sediment particles initiated at the edge of the front pier at an angle of $\pm 50^\circ$ to 60° . However, scour at downstream pier was initiated with a further increase in velocity. The initial dislodgement of particle motion at the pier flanks was due to acceleration of tangential flow past the pier. At a required approaching flow velocity ($U = 0.9U_c$), the downflow grew at the upstream face of the pier and acted as an impinging jet. Due to the erosive action of downflow, scour continued and the scour hole developed with time. The horseshoe vortex

Table 3 Experimental conditions

| No. | Condition | Effect | References |
|-----|---------------------------------------|---|----------------------------------|
| 1 | $B/b \geq 10$ | Negligible side wall (or blockage) effect | Chiew and Melville (1987) |
| 2 | $b/d_{50} \approx 25$ to 130 | Negligible sediment size effect | Tafarojnoruz et al. (2010, 2012) |
| 3 | $\sigma_g < 1.4$ | Uniform sediment | Dey et al. (1995) |
| 4 | $R_p > 7000$ | Negligible flow viscosity | Monti (1994) |
| 5 | Variation of $d_{se} < 0.05b$ in 24 h | Achieved equilibrium condition | Melville and Chiew (1999) |

B channel width, b pier diameter, d_{50} median diameter of sediments, h flow depth, σ_g geometric standard deviation of particle size distribution, R_p pier Reynolds number, d_{se} equilibrium scour depth at an isolated single pier

developed owing to the diving mode of flow into the scour hole, which was visually observed from the trajectory motion of dye injected at the upstream edge of scour hole. In the upstream part of the scour hole the sediment particles were mainly dislodged by the action of downflow and velocity of horseshoe vortex. The process of removal of sediment particles by the action of downflow and horseshoe vortex within the scour hole continued until an equilibrium state was reached. The equilibrium scour depth was attained, when there is no further erosion by the impact of the downflow and horseshoe vortex within the scour hole zone. In this study, 36 h duration was found to be sufficient to ensure the achievement of an equilibrium scour depth. The aforementioned scour process was quite similar for both upstream and downstream piers.

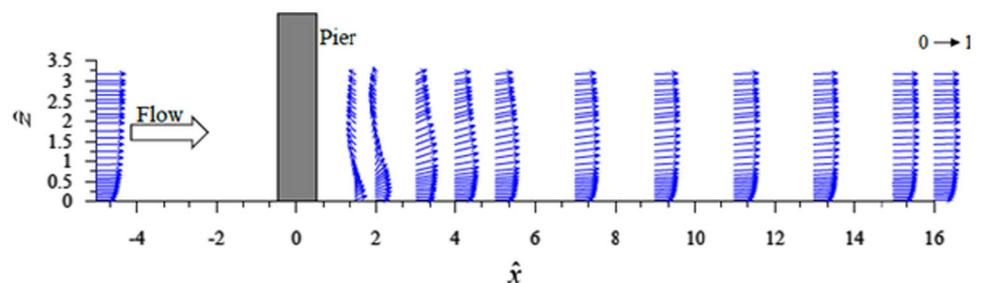
Further, sediment particles were lifted off spontaneously from the scour hole into the flow. The transported sediments from the scour hole at front pier were deposited downstream of the pier to form a dune shaped (sand heap) structure. The dune and the superficial sand migrated to further downstream to feed into the scour hole formed at the downstream piers. Initially, for about 2–3 h, the rate of scour and dune migration was relatively high and, thereafter, gradually decreased with time. It was visually observed that a drop in water surface elevation above the dune took place as compared to the uniform flow depth upstream of pier arrangement.

Figure 5 shows nondimensional velocity vector plot in $\hat{x}\hat{z}$ -plane at upstream and downstream of piers placed on a

rough wall. The streamwise distance x and vertical flow depth z are made nondimensional by using pier diameter b as $\hat{x} = x/b$ and $\hat{z} = z/b$. In near wall-wake region, flow exhibits reversal velocity vectors immediately behind the pier combined with a negative pressure gradient along the flow depth. Further downstream of the pier, velocity profiles in far wall-wake flow became positive with a velocity deficit and tended to attain the upstream velocity profile with an increase in distance. In near-wake region, sediment particles were rapidly picked up from the bed by the induced pressure gradient. Initially, the mode of sediment transport towards downstream was in the form of suspended load owing to wake vortices. However, in the developed stage, transportation mode was bed load. It may be noted that the far wall-wake flow with a velocity deficit is the cause for receiving a less approaching flow by the downstream pier.

Scour at a single pier is influenced by various interdependent parameters, such as flow velocity, flow depth, sediment size, sediment gradation, pier shape, and time of scour. An extensive study was carried out by several researchers on scour at a single pier, and a large number of design formulae was proposed. However, this study suggests an additional factor that affects scour at a downstream pier in the presence of an upstream pier in a tandem arrangement. Therefore, in this case, the spacing between the upstream and downstream piers is the parameter to introduce in the pier scour formulation.

Fig. 5 Velocity vector plot at upstream and downstream of the pier. Vector $0 \rightarrow 1$ refers to a scale of $u/U = 0-1$



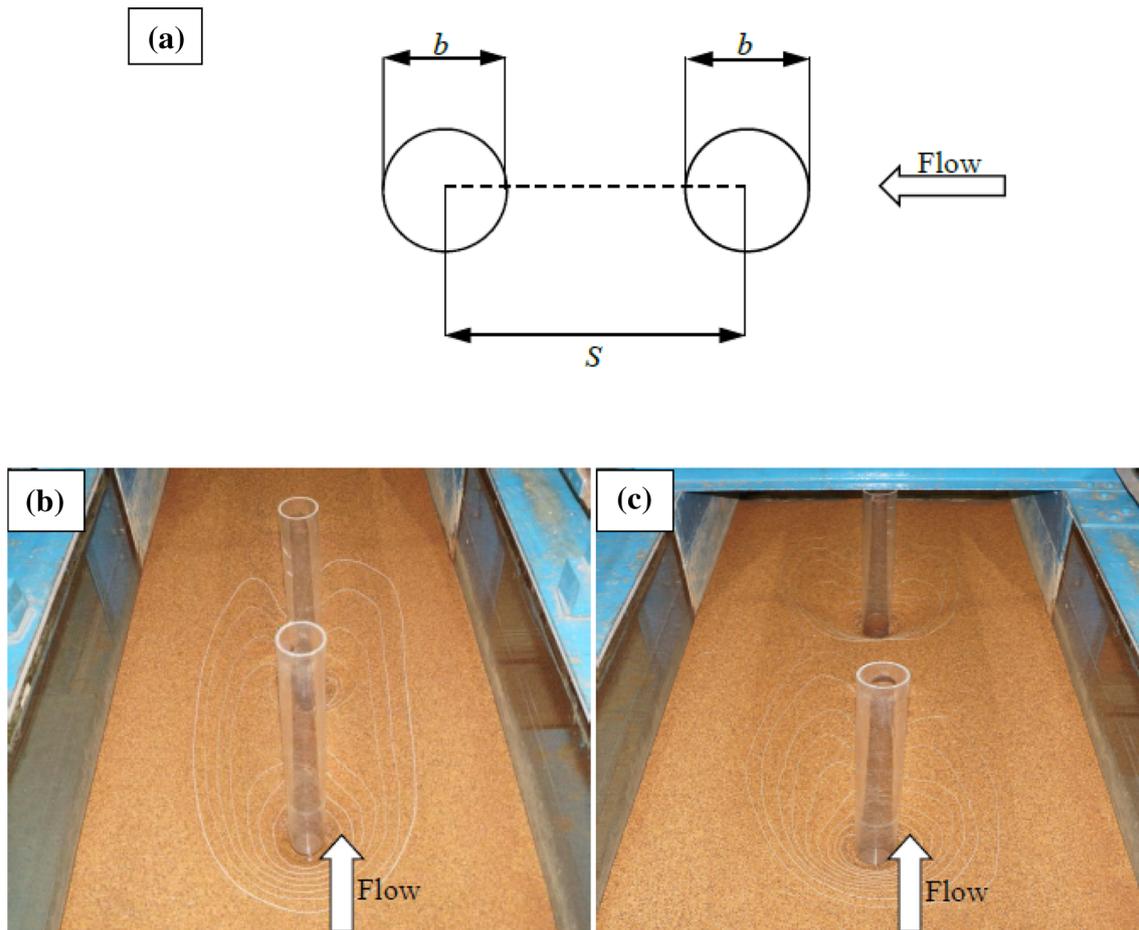


Fig. 6 a Schematic of piers in a tandem arrangement, and photographs of equilibrium scour hole found at identical piers in tandem arrangement for: **b** $S/b = 4$ and **c** $S/b = 10$

Figure 6 shows the experimental photographs of identical piers in a tandem arrangement. Two identical piers were embedded in the sediment bed along the midsection for different pier center-to-center spacings $S = 2b, 2.5b, 3b, 4b, 5b, 6b, 7b, 8b, 9b, 10b, 11b,$ and $12b$, as shown in Fig. 6a. For all the experiments, the upstream pier was fixed at a distance of 8.5 m from the inlet of the flume and only the downstream pier was moved such that a desired spacing S was achieved in different experiments. The experiments on piers in tandem were conducted in two sediment beds having median sizes of 0.96 mm and 1.8 mm with two identical circular piers of diameters 6 cm and 8.2 cm. Table 4 presents the experimental results. Figure 6 shows an equilibrium scour hole developed at two piers in tandem arrangement for two spacings. If the spacing between the piers is smaller, a combined elliptical shaped scour hole was observed (Fig. 6b), whereas, with an increase in pier spacing, two separate circular shaped scour holes at upstream and downstream piers were observed (Fig. 6c).

It was observed that the upstream pier primarily influences the scour at a downstream pier in two ways, as

follows: (i) the upstream pier modifies the flow field in the wake flow region in the form of deficit in velocity that acts as an approaching flow for the downstream pier, causing a decrease in scour depth at downstream pier. (ii) The bed sediment that is scoured at the upstream pier is being fed to the scour hole formed at the downstream pier. As a result, a reduction of scour depth at the downstream pier as compared to that at the upstream pier is usually observed.

Equilibrium scour depth at a downstream pier

To assess the effects of spacing S between two piers on equilibrium scour depth d_{s2} at downstream pier in a tandem arrangement, we can introduce a parameter $K_T (=d_{s2}/d_{se})$ in Melville and Coleman's (2000) equation of equilibrium scour depth containing K factors.

Figure 7 shows the variation of nondimensional equilibrium scour depth $K_T (=d_{s2}/d_{se})$ at downstream pier for different median diameter of sediments d_{50} and pier diameters b in tandem arrangements. The K_T is called spacing factor in tandem arrangement. It is observed that

Table 4 Scour depths measured at identical piers in tandem arrangement for uniform sediments

| Run | S (cm) | b (cm) | d_{s1} (cm) | d_{s2} (cm) | S/b | d_{s1}/d_{se} | d_{s2}/d_{se} |
|--|----------|----------|---------------|---------------|-------|-----------------|-----------------|
| Median sediment size, $d_{50} = 0.96$ mm (series I) | | | | | | | |
| R1 | 12.00 | 6.00 | 13.00 | 10.80 | 2.00 | 1.01 | 0.84 |
| R2 | 15.00 | 6.00 | 12.90 | 10.90 | 2.50 | 1.00 | 0.84 |
| R3 | 18.00 | 6.00 | 12.90 | 10.90 | 3.00 | 1.00 | 0.84 |
| R4 | 24.00 | 6.00 | 12.80 | 10.80 | 4.00 | 0.99 | 0.84 |
| R5 | 30.00 | 6.00 | 12.90 | 10.70 | 5.00 | 1.00 | 0.83 |
| R6 | 36.00 | 6.00 | 12.90 | 10.60 | 6.00 | 1.00 | 0.82 |
| R7 | 48.00 | 6.00 | 12.90 | 10.20 | 8.00 | 1.00 | 0.79 |
| R8 | 60.00 | 6.00 | 12.80 | 10.70 | 10.00 | 0.99 | 0.83 |
| R9 | 72.00 | 6.00 | 12.90 | 10.70 | 12.00 | 1.00 | 0.83 |
| Median sediment size, $d_{50} = 0.96$ mm (series II) | | | | | | | |
| R10 | 16.40 | 8.20 | 15.10 | 13.40 | 2.00 | 1.00 | 0.88 |
| R11 | 24.60 | 8.20 | 15.10 | 13.40 | 3.00 | 1.00 | 0.88 |
| R12 | 32.80 | 8.20 | 15.10 | 13.20 | 4.00 | 1.00 | 0.87 |
| R13 | 41.00 | 8.20 | 15.10 | 12.70 | 5.00 | 1.00 | 0.84 |
| R14 | 57.40 | 8.20 | 15.10 | 12.30 | 7.00 | 1.00 | 0.81 |
| R15 | 65.60 | 8.20 | 15.10 | 12.20 | 8.00 | 1.00 | 0.80 |
| R16 | 73.80 | 8.20 | 15.10 | 12.60 | 9.00 | 1.00 | 0.83 |
| R17 | 90.20 | 8.20 | 15.10 | 12.60 | 11.00 | 1.00 | 0.83 |
| Median sediment size, $d_{50} = 1.8$ mm (series III) | | | | | | | |
| R18 | 12.00 | 6.00 | 13.40 | 11.10 | 2.00 | 1.02 | 0.84 |
| R19 | 24.00 | 6.00 | 13.30 | 10.70 | 4.00 | 1.01 | 0.81 |
| R20 | 48.00 | 6.00 | 13.10 | 10.20 | 8.00 | 1.00 | 0.77 |
| R21 | 60.00 | 6.00 | 13.10 | 10.90 | 10.00 | 1.00 | 0.83 |
| R22 | 72.00 | 6.00 | 13.10 | 10.90 | 12.00 | 1.00 | 0.83 |
| Median sediment size, $d_{50} = 1.8$ mm (series IV) | | | | | | | |
| R23 | 24.60 | 8.20 | 15.50 | 13.60 | 3.00 | 1.01 | 0.88 |
| R24 | 41.00 | 8.20 | 15.30 | 13.10 | 5.00 | 1.00 | 0.85 |
| R25 | 57.40 | 8.20 | 15.30 | 12.90 | 7.00 | 1.00 | 0.84 |
| R26 | 73.80 | 8.20 | 15.20 | 13.40 | 9.00 | 0.99 | 0.87 |
| R27 | 90.20 | 8.20 | 15.30 | 13.50 | 11.00 | 1.00 | 0.88 |

S spacing between the two identical piers, b pier diameter, d_{s1} equilibrium scour depth at upstream pier, d_{s2} equilibrium scour depth at downstream pier, d_{se} equilibrium scour depth at an isolated single pier

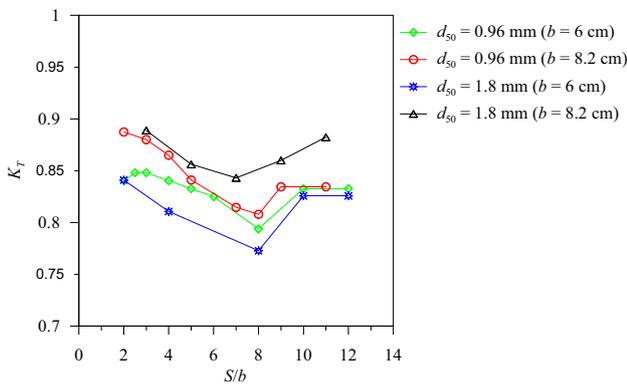


Fig. 7 K_T as a function of S/b for different d_{50} and b in tandem arrangements

the equilibrium scour depth d_{s2} at downstream pier decreases with an increase in spacing from $4b$ to $8b$ and minimum value of the equilibrium scour depth was recorded at $8b$ irrespective of the diameter of the identical piers. As the spacing increases from $4b$ to $8b$, wake vortices that shed from the upstream pier are weakened; therefore, sediment eroded from the scour hole at the upstream pier increasingly dumps in the scour hole at the downstream pier, where the equilibrium scour depth decreases as a consequence. The equilibrium scour depth at the downstream pier is found to increase with further increase in spacing from $8b$ to $12b$ (Table 4). Analysis of the experimental data demonstrates that the approaching velocity of the upstream pier on scour at the downstream pier exists within the spacing range $4b$ – $12b$. As an effect of the

upstream pier to reduce the approaching flow velocity at its downstream owing to velocity deficit, the scour depth at the downstream pier reduces as compared to that of an upstream pier. From Fig. 7, it is apparent that the equilibrium scour depth at the downstream pier varies approximately from 0.8 to 0.87 times the equilibrium scour depth at a single pier. The K_T factor is determined from the curves, as shown in Fig. 7. Thus, the proposed K_T factor is adequate from the view point of a safe pier foundation design at downstream pier.

To investigate the influence of pier size b and pier spacing S , two identical piers of varying diameter (6 cm and 8.2 cm) and spacing ($S = 2b - 12b$) were tested, as shown in Table 4. The results showed that as the pier size increases, the equilibrium scour depth increases at the upstream pier as well as at the downstream pier. Shen et al. (1969) and Kothyari et al. (1992) also observed similar findings. Apart from studying the effects of spacing and identical pier diameters, the effect of median sediment size on equilibrium scour depth is also studied. It is found that the increase in median sediment size marginally increases the equilibrium scour depth irrespective of the b/d_{50} ratio.

Temporal variation of scour depth at a downstream pier

The temporal variation of scour depth at a downstream pier aligned in tandem arrangement along the flow direction was investigated. In the experiments, the asymptotic nature of the scour depth development over time was observed. Duration of each experiment was considered as 36 h, when the rate of increase of the scour depth was found to be less than 5% of the pier diameter in last 24 h (Melville and Chiew 1999). A total of 42 experiments was carried out in this study. The details of experimental conditions are presented in Tables 4, 5 and 6. The equilibrium scour depths at both piers were measured at a regular time interval until the equilibrium condition was achieved.

Table 5 Experimental results of scour at identical piers in staggered arrangement at $L/b = 4$

| S (cm) | b (cm) | d_{s1} (cm) | d_{s2} (cm) | S/b | d_{s1}/d_{se} | d_{s2}/d_{se} |
|----------|----------|---------------|---------------|-------|-----------------|-----------------|
| 12 | 6 | 12.1 | 12.0 | 2 | 0.92 | 0.91 |
| 24 | 6 | 12.4 | 12.3 | 4 | 0.94 | 0.93 |
| 36 | 6 | 12.2 | 11.7 | 6 | 0.92 | 0.89 |
| 48 | 6 | 12.4 | 12.4 | 8 | 0.94 | 0.94 |
| 60 | 6 | 12.2 | 13.0 | 10 | 0.92 | 0.98 |

S streamwise spacing, b pier diameter, d_{s1} equilibrium scour depth at upstream pier, d_{s2} equilibrium scour depth at downstream pier, and d_{se} equilibrium scour depth at an isolated single pier

According to Sumer et al. (1993), the scour depth d_s at any instant of time t can approximately be represented in functional form as

$$d_s = d_{se} \left(1 - \exp \left(-\frac{t}{T} \right) \right), \quad (2)$$

where d_s is the temporal scour depth, d_{se} the equilibrium scour depth, t the time, and T the time scale.

The quantity T represents the time period during which substantial scour develops. The value of T can be determined from the d_s versus t plot by estimating the slope of the tangent line to the $d_s(t)$ curve at $t = 0$, as shown schematically at left top in Fig. 8a and b for different median sediment sizes, $d_{50} = 0.96$ mm and 1.8 mm, respectively. The time scale can be represented in nondimensional form as $T_* [= T(\Delta g d_{50}^3)^{0.5}/b^2]$, where $\Delta = s - 1$ and g is the gravitational acceleration. The nondimensional time scale T_* can be written in the following functional form

$$T_* = T_*(S/b, \Theta_c), \quad (3)$$

where Θ_c is the threshold Shields parameter, that is, $\tau_c/(\Delta \rho g d_{50})$, τ_c is the threshold bed shear stress, and ρ is the mass density of water. The variations of nondimensional time scale T_* with pier spacing for different sediment sizes d_{50} and pier diameters b are given in Fig. 8a and b. The nondimensional time scale T_* at downstream pier increases with an increase in spacing from $2b$ to $4b$ and maximum value of T_* is observed at $4b$ irrespective of the diameter of the piers. The time scale T_* at downstream pier is found to be almost constant with further increase in spacing from $4b$ to $12b$ (Fig. 8a and b).

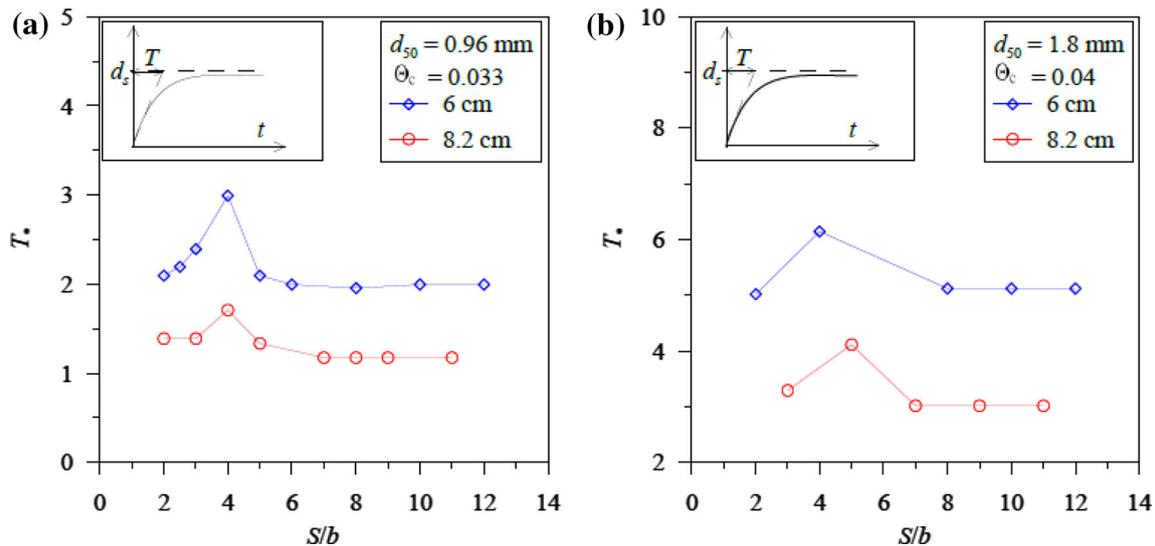
The nondimensional temporal variations of scour depth d_{s2}/d_{se} versus nondimensional time t/T are plotted in Figs. 9 and 10, where T is the time scale of the scouring process, as discussed in Sumer et al. (1993), which is equal to the time period during which substantial scour develops. These figures show that the time series plots fall on each other, irrespective of the spacing between the piers and of the sediment size. Plots also showed that 90% of the scour hole development took place at approximately $t/T = 10$. The following exponential equation, obtained by best fitting the numerical data with experimental data, is found to well represent the time series data. It can easily be used for an accurate estimation of the temporal variation of scour depth as a function of U/U_c , t/T , and S/b as follows:

$$\frac{d_{s2}}{d_{se}} = \exp \left\{ -0.0383 \left[\frac{U}{U_c} \log \left(\frac{1}{43.25} \frac{t}{T} \left(\frac{S}{b} \right)^{-0.25} \right) \right]^2 \right\}. \quad (4)$$

Table 6 Experimental results of scour depths at three identical piers in staggered arrangements

| S (cm) | b (cm) | d_{s1} (cm) | $d_{s2} = (d_{s2a} + d_{s2b})/2$ (cm) | θ (°) | S/b | d_{s1}/d_{se} | d_{s2}/d_{se} |
|--|----------|---------------|---------------------------------------|--------------|-------|-----------------|-----------------|
| Median sediment size, $d_{50} = 1.8$ mm ($L = 2b$) | | | | | | | |
| 12 | 6 | 14.1 | 15.5 | 45.0 | 2 | 1.07 | 1.17 |
| 24 | 6 | 13.2 | 14.6 | 26.6 | 4 | 1.00 | 1.11 |
| 36 | 6 | 13.0 | 14.6 | 18.4 | 6 | 0.98 | 1.11 |
| 48 | 6 | 12.8 | 14.0 | 14.0 | 8 | 0.97 | 1.06 |
| 66 | 6 | 13.1 | 13.1 | 11.3 | 10 | 0.99 | 0.99 |
| Median sediment size, $d_{50} = 1.8$ mm ($L = 3b$) | | | | | | | |
| 18 | 6 | 12.9 | 13.3 | 45.0 | 3 | 0.98 | 1.01 |
| 30 | 6 | 12.9 | 13.1 | 31.0 | 5 | 0.98 | 0.99 |
| 42 | 6 | 12.6 | 12.7 | 23.2 | 7 | 0.95 | 0.96 |
| 54 | 6 | 13.0 | 12.4 | 18.4 | 9 | 0.98 | 0.94 |
| 66 | 6 | 12.9 | 12.6 | 15.2 | 11 | 0.98 | 0.95 |

S streamwise spacing, b pier diameter, d_{s1} equilibrium scour depth at upstream pier, d_{s2a} equilibrium scour depth at downstream pier (right pier), d_{s2b} equilibrium scour depth at downstream pier (left pier), d_{s2} average value of equilibrium scour depth at downstream piers = $(d_{s2a} + d_{s2b})/2$, θ angle between upstream and downstream piers, d_{se} equilibrium scour depth, and L lateral offset distance

**Fig. 8** Variations of T^* with S/b for different median sediment sizes d_{50} and pier diameters b : **a** $d_{50} = 0.96$ mm and **b** $d_{50} = 1.8$ mm

Equation (4) is valid in the following ranges of the independent variables: $S/b = 2$ – 12 and $t/T = 0$ – 40 , with a coefficient of correlation of 0.98.

Equation (4) does not account for the median sediment size. The reason is attributed to the fact that the approaching velocities in experimental runs were close to the critical flow velocities of the bed sediments. From Eq. (4), it is clear that the temporal variation of scour depth at downstream pier is directly proportional to spacing between the piers and inversely proportional to time scale. In the above equation, as t tends to be the equilibrium time, the time dependent scour depth reaches the equilibrium scour depth. The experimental data sets

are found to be well represented by Eq. (4) for downstream pier in tandem arrangement, as shown in Figs. 9 and 10. These plots show that the maximum deviation between computed and measured equilibrium scour depths at downstream piers is $\pm 20\%$. It is understood that the errors in estimated depths at the downstream pier mainly occur in the shortest period of time. The errors in predicted values of temporal scour depths at downstream piers are similar for median sediment sizes $d_{50} = 0.96$ mm and 1.8 mm. Therefore, the scour depths at downstream piers in tandem arrangement can be estimated using Eq. (4), with great advantage in the design of pier foundation when piers are aligned in series.

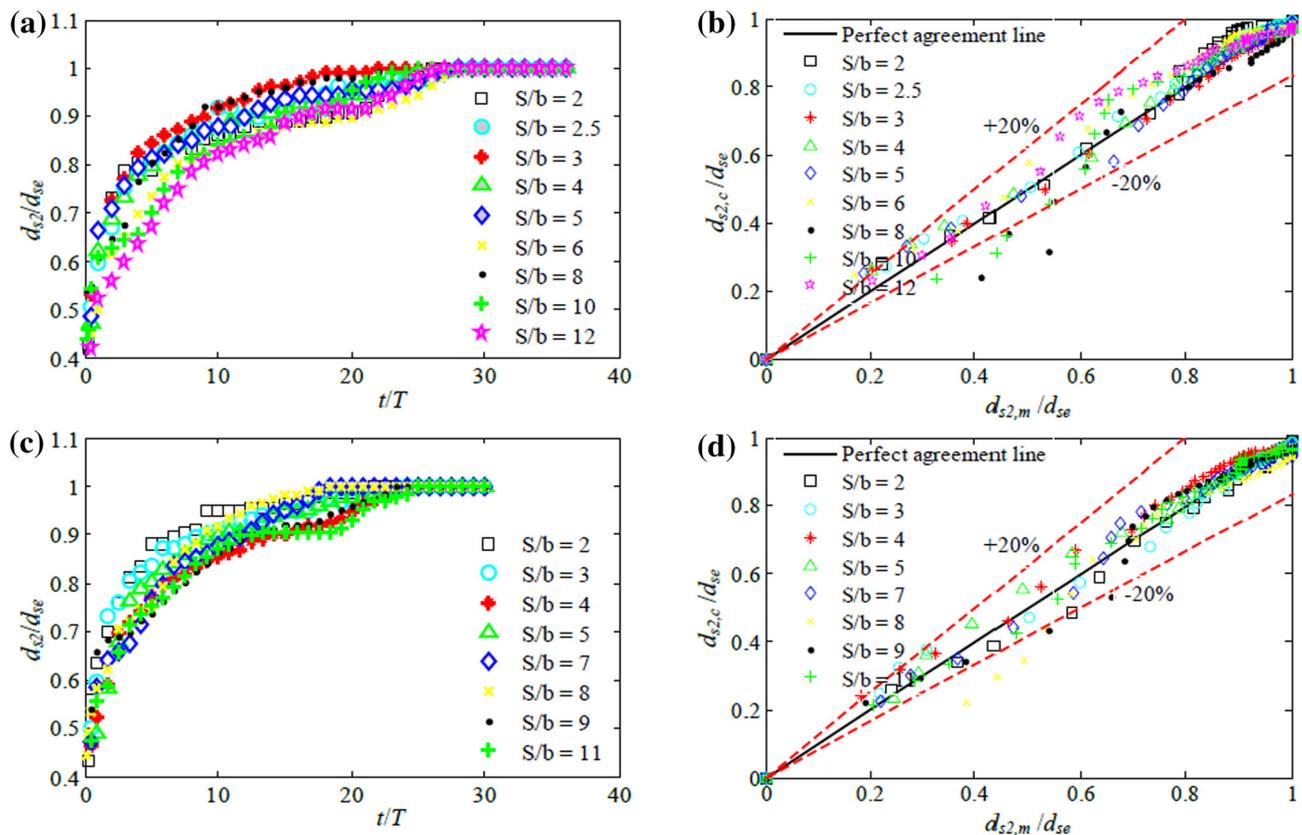


Fig. 9 Temporal variation of scour depth at a downstream pier of identical pier diameters in tandem arrangement in sediment bed with median size $d_{50} = 0.96$ mm: **a** $b = 6$ cm, **b** computed versus

measured scour depth for $b = 6$ cm, **c** $b = 8.2$ cm, and **d** computed versus measured scour depth for $b = 8.2$ cm

Two piers in staggered arrangements

Equilibrium scour depth at a downstream pier

Scour at piers in staggered arrangements is investigated in this section. Two identical piers of 6 cm diameter were embedded in the sediment bed ($d_{50} = 1.8$ mm) at a lateral offset spacing of $L = 4b$, as shown in Fig. 11a. The streamwise spacing S was varied from $2b$ to $10b$. Figure 11b shows the equilibrium scour holes at two piers in a staggered arrangement. In all the experiments, the upstream pier was located at a distance of 8.5 m from the inlet of the flume along the midsection and only the downstream pier was moved streamwise at desired spacing S and lateral offset spacing L in different experiments. Table 5 presents the experimental results of equilibrium scour depths at two piers in staggered arrangements.

Figure 12 shows the variation of nondimensional equilibrium scour depth at downstream pier with different streamwise spacings S in a staggered arrangement. The equilibrium scour depth d_{se} at a single pier was 13.2 cm. It is observed from Table 5 that the equilibrium scour depths measured at the upstream pier are independent of the offset

distance as well as streamwise distance between the two piers. However, the equilibrium scour depth measured at the downstream pier increases with an increase in streamwise spacing S . For a given nondimensional offset distance ($L/b = 4$) and at a streamwise spacing of $2b$, the equilibrium scour depth at downstream pier is equal to that at the upstream pier, because the downstream pier acts as an isolated pier. In this case, the design criterion for the scour depth at the downstream pier is same as that of the upstream pier. The equilibrium scour depth at the downstream pier is the lowest for the streamwise spacing of $6b$, because transported sediment from the scour hole at the upstream pier fills the scour hole at the downstream pier.

The equilibrium scour depth at downstream pier increases with an increase in streamwise spacing for a given offset distance L owing to the increase in exposure to the approaching flow velocity. However, this trend is not followed for $S = 6b$, where the measured scour depth is found to be minimum. It is observed that for streamwise spacing $S > 8b$ and $L/b = 4$, the equilibrium scour depth at a downstream pier is greater than that at an upstream pier, because the downstream pier is on the boundary of the wake region created by the upstream pier and the resulting

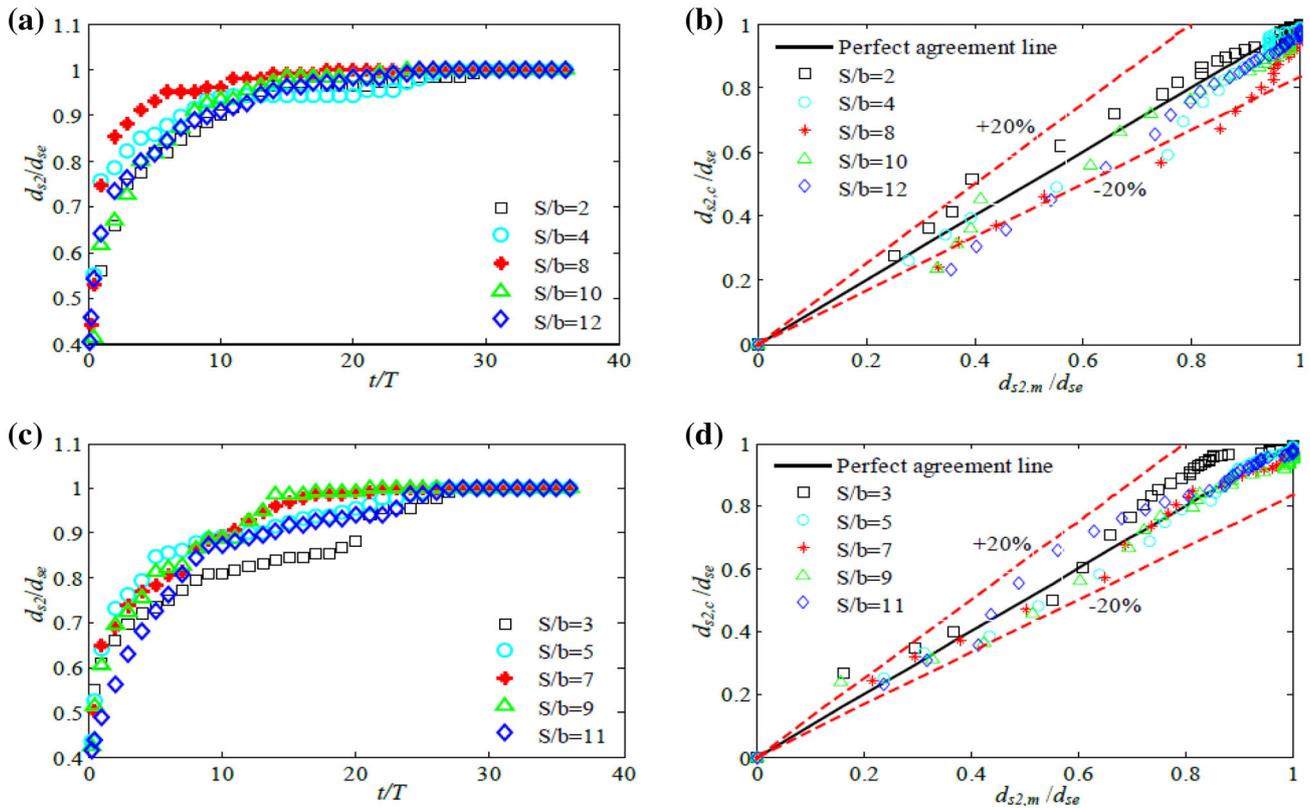


Fig. 10 Temporal variation of scour depth at a downstream pier of pier diameters in tandem arrangement with a sediment bed of median size $d_{50} = 1.8$ mm: **a** $b = 6$ cm, **b** computed versus measured scour

depth for $b = 6$ cm, **c** $b = 8.2$ cm, and **d** computed versus measured scour depth for $b = 8.2$ cm

approaching flow velocity is somewhat greater than the average approaching flow velocity.

Temporal variation of scour depth at a downstream pier

The value of time scale T can be determined from d_s versus t plots, as discussed in the preceding section. The variation of nondimensional time scale T^* with nondimensional spacing S/b in the arrangement of two staggered piers for a median sediment size $d_{50} = 1.8$ mm and a pier diameter $b = 6$ cm is shown in Fig. 13. The nondimensional time scale T^* at downstream pier increases with an increase in spacing from $2b$ to $4b$ and maximum value of T^* is observed at $4b$. The time scale T^* at downstream pier is found to be almost constant with further increase in spacing from $6b$ to $10b$ (Fig. 13).

The nondimensional temporal variation of scour depth d_{s2}/d_{se} at downstream pier in a staggered arrangement against nondimensional time is plotted in Fig. 14a. This figure shows that the time series plots fall on each other irrespective of the spacing between the piers and of the sediment size. Plots also showed that 90% of the scour hole development took place at approximately $t/T = 5$, which is rather faster than that in tandem arrangement. The

following exponential equation, obtained by best fitting the numerical data with experimental data, was found to well represent the time series data. It can easily be used for an estimation of the temporal variation of scour depth as a function of U/U_c , L/b , t/T , and S/b as follows:

$$\frac{d_{s2}}{d_{se}} = \exp \left\{ -0.0155 \left[\frac{U}{U_c} \log \left(\frac{1}{78.68} \left(\frac{L}{b} \right)^{0.25} \frac{t}{T} \left(\frac{S}{b} \right)^{-0.25} \right) \right]^2 \right\} \tag{5}$$

It may be noted that Eq. (5) is valid in the following ranges of the independent variables: $L/b = 4$, $t/T = 0-36$ and $S/b = 2-10$, with a coefficient of correlation of 0.97. The experimental data sets are found to be well represented by Eq. (5). In Eq. (5), nondimensional lateral distance L/b is introduced as an independent variable to determine the temporal scour depth at downstream pier. From Eq. (5), it is clear that the scour depth d_{s2}/d_{se} is inversely proportional to the nondimensional lateral distance L/b . Scatter plots of experimental and computed scour depths are drawn in a staggered arrangement as shown in Fig. 14b. It may be noted that the computed data points fall on the perfect agreement line within the $\pm 20\%$ error band. Therefore, it can be stated that the temporal variation of scour depth estimated at downstream pier by

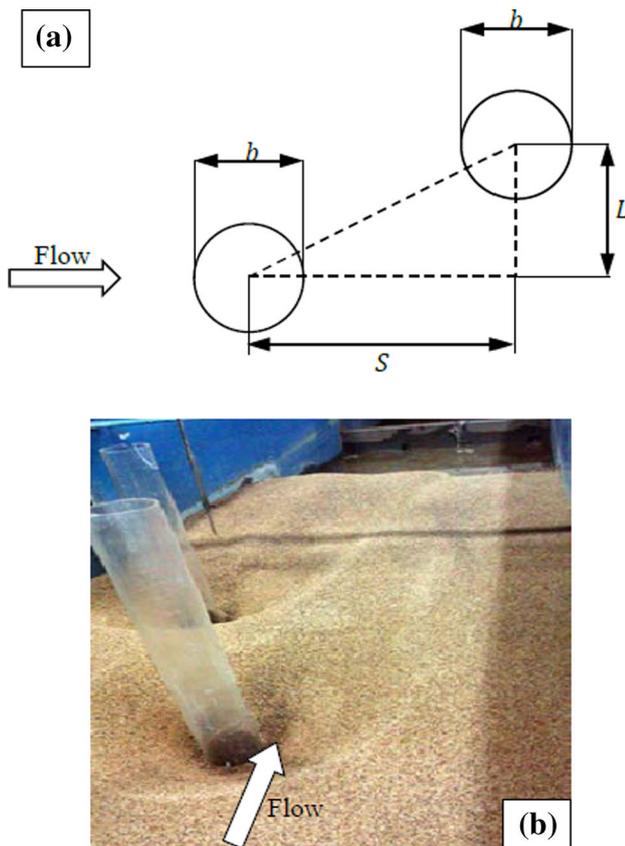


Fig. 11 a Schematic of two piers in a staggered arrangement, and b equilibrium scour hole at two piers in a staggered arrangement at a lateral offset spacing of $L = 4b$

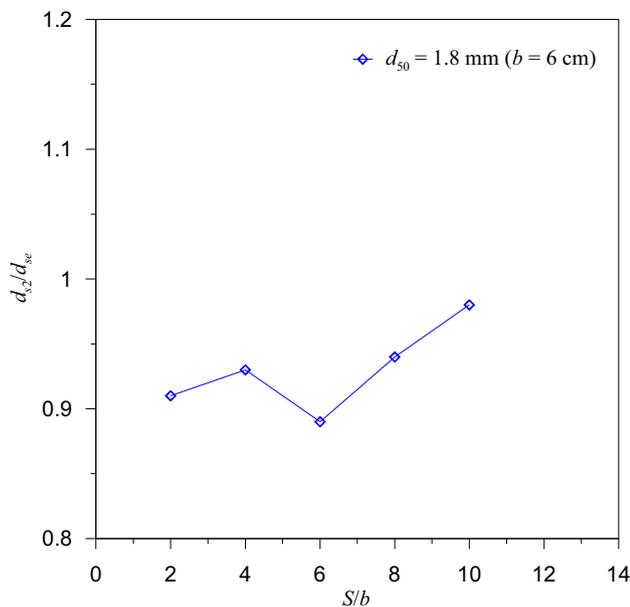


Fig. 12 Variation of nondimensional equilibrium scour depth at a downstream pier with streamwise spacing S for a constant lateral offset spacing ($L = 4b$) in a staggered arrangement

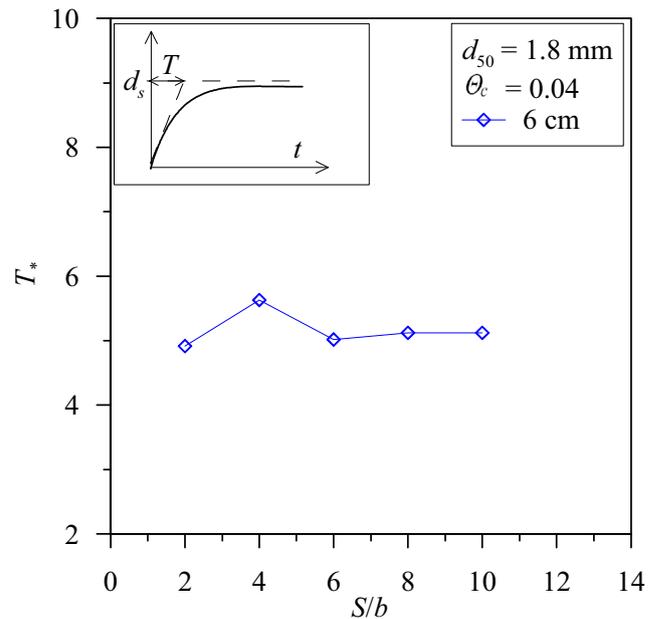


Fig. 13 Variation of T^* with S/b for a median sediment size $d_{50} = 1.8$ mm and pier diameter $b = 6$ cm in the arrangement of two staggered piers

using Eq. (5) would be useful in designing downstream pier foundation in the arrangement of two staggered piers.

Three piers in staggered arrangements

Equilibrium scour depth at downstream piers

The scour at three piers in staggered arrangement was finally observed. It may, however, be noted that the information available on this study from existing literature is very limited. Thus, there is a need to carry out experimental investigation in this regard. A series of laboratory tests was carried out to study the effect of interference of upstream pier on scour depths at two downstream piers in staggered arrangement (Fig. 15a). Three identical piers of 6 cm diameter embedded in a sediment bed of median sediment size $d_{50} = 1.8$ mm were used in these experiments. One pier was installed in the upstream and two piers were installed in the downstream, as shown in Fig. 15a. Offset pier spacing between the two downstream piers was varied as $L = 2b$ and $3b$. Experiments were carried out with the same procedure adopted for other experimental runs with two piers in staggered arrangement. Photographs of the scour holes taken after completion of the experiments with two different offset spacings between the downstream piers are shown in Fig. 15b and c. Table 6 furnishes the experimental conditions and the corresponding measured equilibrium scour depths. Since the downstream piers are symmetrical with respect to the

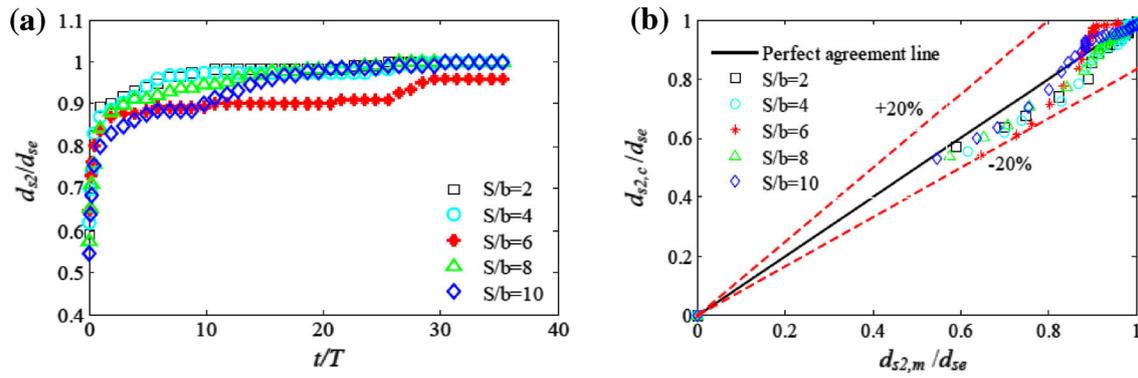


Fig. 14 Temporal variations of scour depth at downstream pier of two identical pier diameters in a staggered arrangement ($L = 4b$) with a sediment bed of $d_{50} = 1.8$ mm: **a** $b = 6$ cm, **b** computed versus measured scour depths for $b = 6$ cm

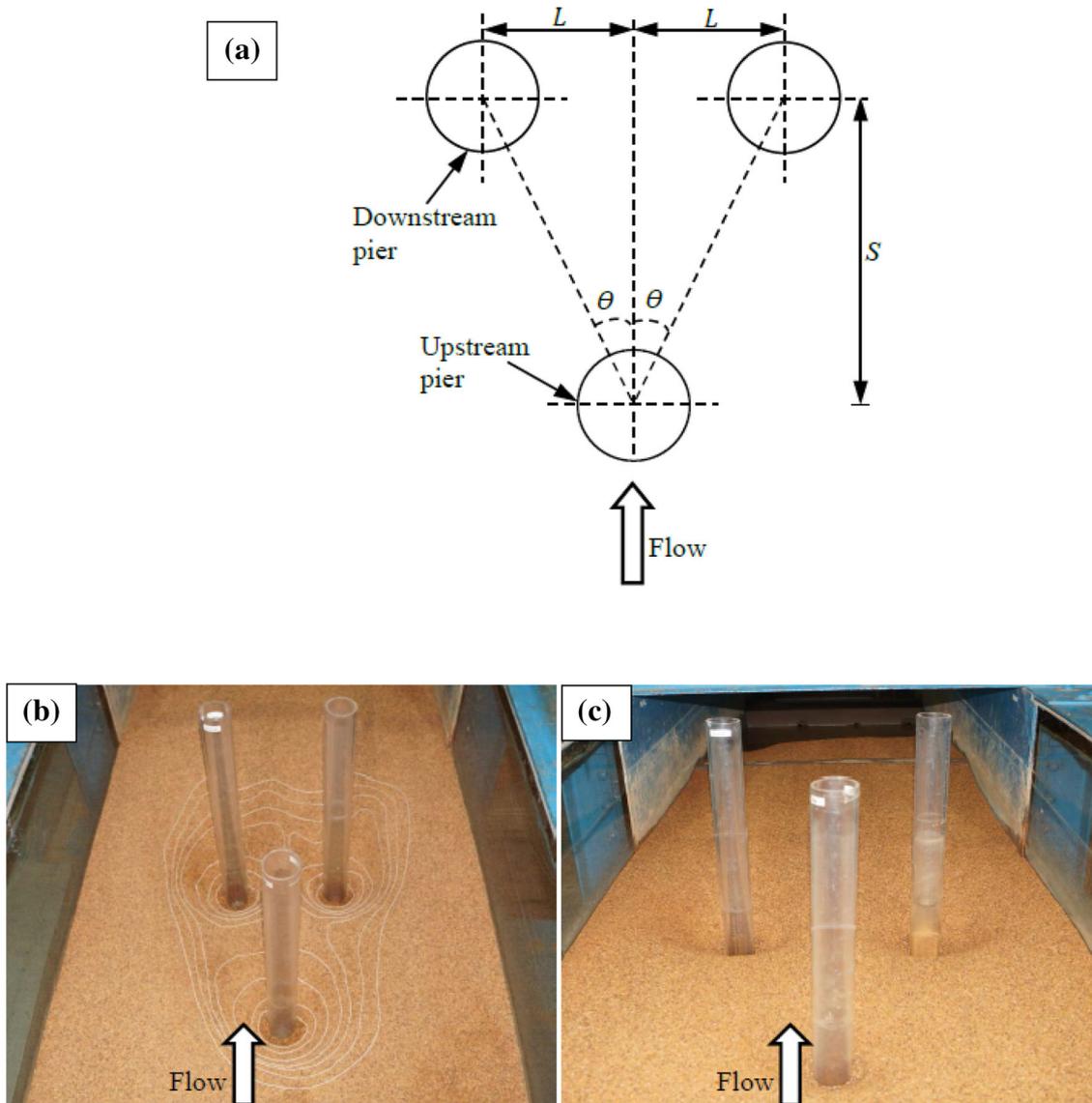


Fig. 15 Scour holes at three identical piers in staggered arrangement: **a** schematic of three piers in staggered arrangement, **b** photographs of offset spacing $L = 2b$, and **c** offset spacing $L = 3b$

midsection, the scour depths measured at the downstream piers are quite close and hence the average value of equilibrium scour depth at downstream piers is considered in the analysis.

In the series of experiments on piers in staggered arrangement, the scour depths at downstream piers are affected by the wake vortices shed from the upstream pier depending on the angle and radial distance between the piers. Two sets of experiments were carried out with offset spacing L between two downstream piers $2b$ and $3b$, respectively. In both cases, streamwise spacing S between the upstream and downstream piers was varied; therefore, the angle θ between the upstream and downstream piers with respect to the centerline of the channel was varied as well as the radial distance between the upstream and downstream piers. The values of spacing and angles are given in Table 6.

It was observed that, as the radial distance between the upstream and downstream piers increases, the scour depths at the downstream piers decrease, because of the weakening of wall wake velocity that acts as an approaching velocity for the downstream piers. As compared to the $L = 3b$ case, the scour depths at the downstream piers are higher than that for the $L = 2b$ case. Therefore, it can be concluded that, as the lateral offset spacing between the downstream piers increases, the equilibrium scour depths at the downstream piers decrease. From Table 6, it is also obvious that for a shorter streamwise spacing $S = 2b$ between the upstream and the downstream piers, the equilibrium scour depth at upstream pier is slightly higher than that of an isolated pier, because of the possible feedback effect of the downstream piers on the scouring strength at the upstream pier. Figure 16a and b show the variations of

nondimensional scour depths at the upstream and downstream piers with streamwise spacing S in a staggered arrangement of three identical piers. From the plots, it is evident that, as the streamwise spacing S between the upstream pier and downstream piers increases, the scour depths at the downstream piers decrease and attain approximately the same value that occurs at the upstream pier at a streamwise distance $S = 10b$.

Temporal variation of scour depth at downstream piers

The variations of nondimensional time scale T_* with nondimensional spacing S/b for a median sediment size d_{50} and three identical pier diameters in staggered arrangements for different lateral offset spacing $L = 2b$ and $3b$ are shown in Fig. 17a and b. The nondimensional time scale T_* at downstream piers with offset spacing $L = 2b$ increases with an increase in streamwise spacing S from $2b$ to $4b$ and maximum value of T_* is observed at $4b$. With further increase in spacing from $4b$ to $10b$ the time scale T_* at downstream piers is found to be almost constant (Fig. 17a). The results for offset spacing of $L = 3b$ are similar to those of $L = 2b$ (Fig. 17b).

The equation to estimate the temporal variation of scour depths at the downstream piers in three piers in staggered arrangements is slightly modified in terms of one coefficient, because of the occurrence of increase in scour depth at the downstream piers in the arrangement of three staggered piers with respect to that in the arrangement of two staggered piers with similar spacing in streamwise and lateral directions. The scour depth at the downstream pier increases, because the two downstream piers together act effectively as a bigger pier. The equation to estimate the

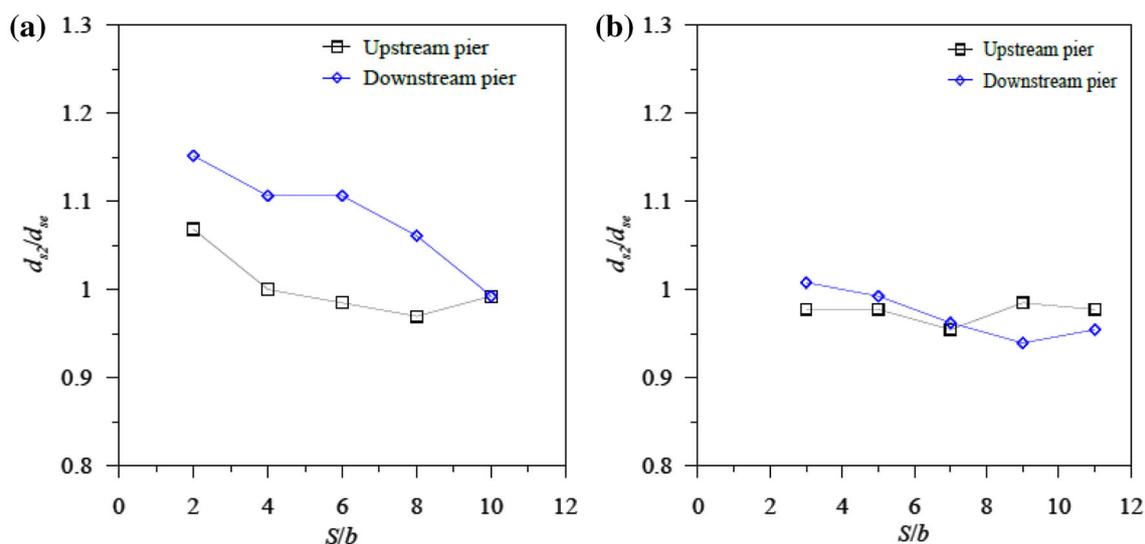


Fig. 16 Variations of nondimensional equilibrium scour depth at three identical pier diameters in staggered arrangements with a sediment bed of $d_{50} = 1.8$ mm: **a** $L = 2b$ and $b = 6$ cm, and **b** $L = 3b$ and $b = 6$ cm

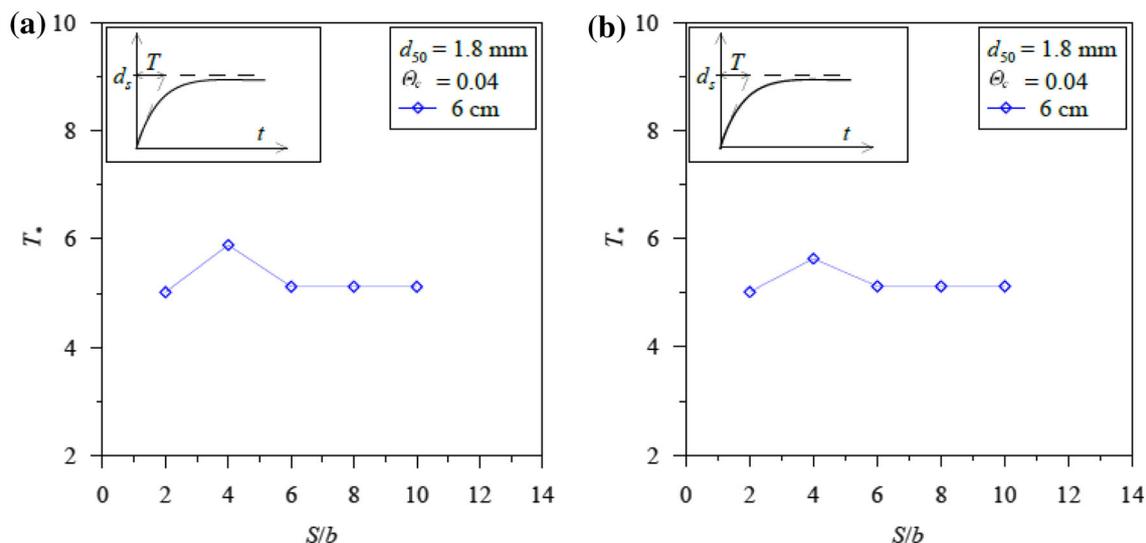


Fig. 17 Variations of T^* with S/b at three identical pier diameters in staggered arrangements with a sediment size of $d_{50} = 1.8$ mm: **a** $L = 2b$ and $b = 6$ cm, and **b** $L = 3b$ and $b = 6$ cm

temporal variation of scour depths at the two downstream piers in the arrangements of three staggered piers is given by:

$$\frac{d_{s2}}{d_{se}} = \exp \left\{ -0.0383 \left[\frac{U}{U_c} \log \left(\frac{1}{78.68} \left(\frac{L}{b} \right)^{0.25} \frac{t}{T} \left(\frac{S}{b} \right)^{-0.25} \right) \right]^2 \right\}. \quad (6)$$

It may be noted that Eq. (6) is valid for the following ranges of the independent variables: $L/b = 2-3$, $t/T = 0-36$, and $S/b = 2-11$, with a coefficient of correlation of 0.97. The experimental data sets of three piers in staggered arrangement are found to be well represented by Eq. (6), as shown in Fig. 18a and c. Scatter plots comparing experimental and computed equilibrium scour depths are displayed in Fig. 18b and d for the experiments with three piers in the staggered arrangement. It may be noted that the computed scour depths at downstream piers with offset spacing $L = 2b$ fall above the perfect agreement line. On the other hand, the computed scour depths at the downstream pier with offset spacing $L = 3b$ fall just below the perfect agreement line. All the data plots are within the $\pm 20\%$ error band. Therefore, it can be concluded that the scour depths estimated at the downstream piers using Eq. (6) would be useful in the arrangement of three staggered piers.

Conclusions

The present study highlights the interference of an upstream pier on clear water scour at downstream piers. Apart from measuring the equilibrium scour depths,

temporal variation of scour depths was measured at downstream piers in: (i) two piers in tandem, (ii) two piers in staggered, and (iii) three piers in staggered arrangements. The primary findings are summarized as follows:

In tandem arrangement, the scour depth at downstream piers varies approximately from 0.8 to 0.87 times the scour depth that occurs at a single pier. Therefore, in a tandem arrangement, no additional design criterion is required for the estimation of scour depths at the downstream piers. The minimum value of the equilibrium scour depth at the downstream pier in tandem arrangement occurs at a spacing of $8b$, which is considered as the best configuration for design. The equilibrium scour depth at the upstream pier in all combinations of tandem and staggered pier arrangements is almost the same as that occurring at a single pier. The scour depth at the downstream pier increases with an increase in pier diameter b and pier spacing S in a tandem arrangement.

For two piers in tandem arrangement, the scour depths at the downstream pier are less than that of the upstream pier. The design criterion for the scour depth at the downstream pier in the staggered arrangement for $L/b = 4$ is the same as that for the scour depth occurring at upstream pier at a streamwise distance $S = 8b$.

For three piers in staggered arrangement, the scour depths at the downstream piers for the streamwise spacing less than $8b$ and for narrow lateral spacing $2b$ are greater than that at the isolated single pier owing to the effect of wake vortices. As the radial distance between the upstream and downstream piers increases, the scour depths at downstream piers decrease. In this arrangement, as the lateral spacing L between the downstream piers increases, the equilibrium scour depths at downstream piers decrease.

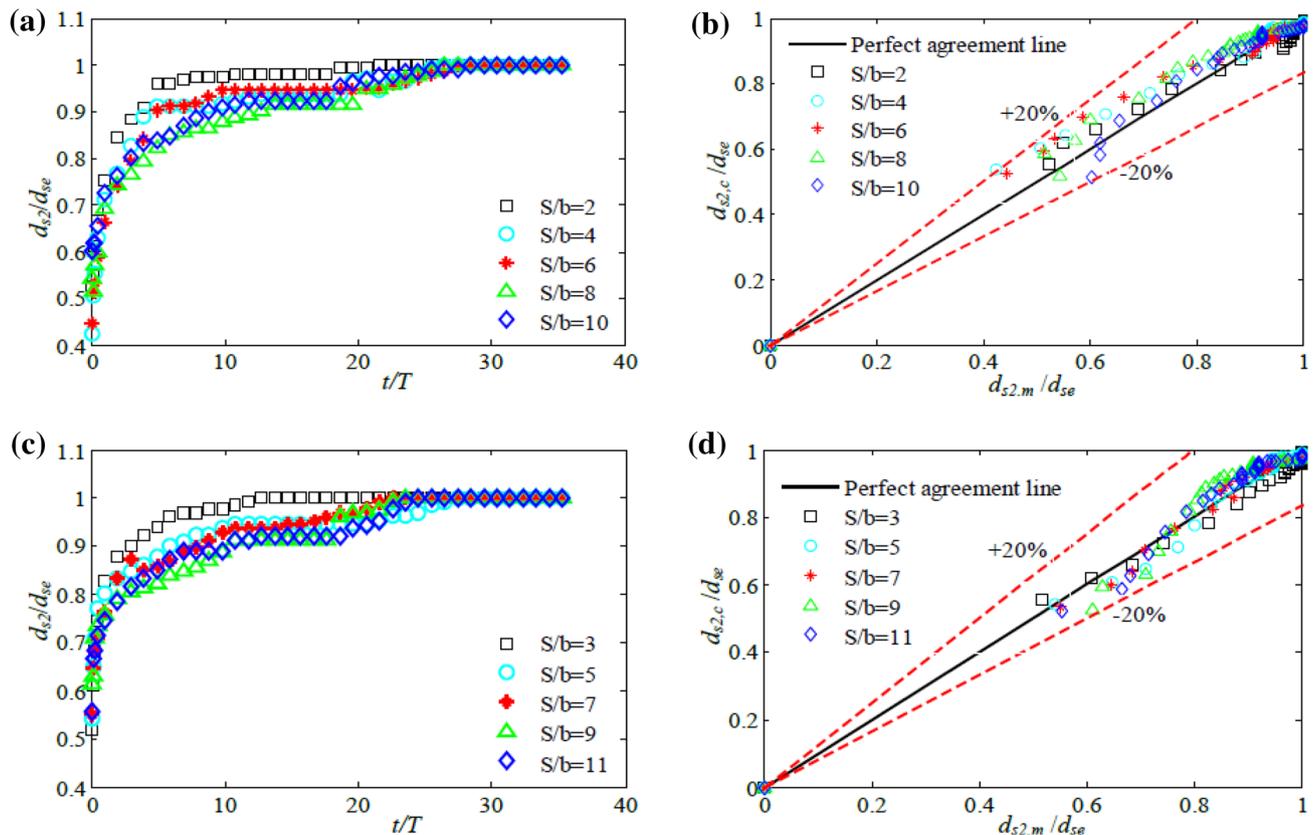


Fig. 18 Temporal variation of scour depth at downstream piers of three pier diameters ($b = 6$ cm) in staggered arrangements with sediment bed of $d_{s0} = 1.8$ mm: **a** time series for $L = 2b$, **b** computed

versus measured scour depths for $L = 2b$, **c** time series for $L = 3b$, and **d** computed versus measured scour depths for $L = 3b$

Further, as the center-to-center streamwise spacing S between upstream and downstream piers increases, the scour depths at the downstream piers decrease and attain approximately the same scour depths occurring at the upstream pier at a streamwise distance $S = 10b$ and lateral spacing $L = 2b$.

Regression equations are developed to estimate the temporal variation of scour depths at a given time at the downstream piers for the various pier arrangements. For all the pier arrangements, the estimated scour depths are within $\pm 20\%$ error band for the downstream piers; hence, a safety factor of 1.2 can be adopted in design calculations.

References

- Beg M, Beg S (2015) Scour hole characteristics of two unequal size bridge piers in tandem arrangement. *ISH J Hydraul Eng* 21(1):85–96. doi:10.1080/09715010.2014.963176
- Breusers HNC, Nicollet G, Shen HW (1977) Local scour around cylindrical piers. *J Hydraul Res* 15(3):211–252. doi:10.1080/00221687709499645
- Cardoso AH, Bettess R (1999) Effects of time and channel geometry on scour at bridge abutments. *J Hydraul Eng* 125(4):388–399. doi:10.1061/(ASCE)0733-9429(1999)125:4(388)

- Chabert J, Engeldinger P (1956) Etude des affouillements autour des piles de ponts. Serie A, Laboratoire National d'Hydraulique, Chatou (in French)
- Chang W, Lai J, Yen C (2004) Evolution of scour depth at circular bridge piers. *J Hydraul Eng* 130(9):905–913. doi:10.1061/(ASCE)0733-9429(2004)130:9(905)
- Chiew YM, Melville BW (1987) Local scour around bridge piers. *J Hydraul Res* 25(1):15–26. doi:10.1080/00221688709499285
- Dey S (1991) Clear water scour around circular bridge piers: a model. Ph.D. thesis, Department of Civil Engineering, Indian Institute of Technology, Kharagpur
- Dey S (1999) Time variation of scour in the vicinity of circular piers. *Proc Inst Civil Eng Water Mar Eng* 136(2):67–75
- Dey S, Raikar RV (2007) Characteristics of horseshoe vortex in developing scour holes at piers. *J Hydraul Eng* 133(4):399–413. doi:10.1061/(ASCE)0733-9429(2007)133:4(399)
- Dey S, Bose SK, Sastry GLN (1995) Clear water scour at circular piers: a model. *J Hydraul Eng* 121(12):869–876. doi:10.1061/(ASCE)0733-9429(1995)121:12(869)
- Elliott KR, Baker CJ (1985) Effect of pier spacing on scour around bridge piers. *J Hydraul Eng* 111(7):1105–1109. doi:10.1061/(ASCE)0733-9429(1985)111:7(1105)
- Ettema R (1980) Scour at bridge piers. Ph.D. thesis, Department of Civil Engineering, University of Auckland, Auckland
- Ferraro D, Tafarojnoruz A, Gaudio R, Cardoso AH (2013) Effects of pile cap thickness on the maximum scour depth at a complex pier. *J Hydraul Eng* 139(5):482–491. doi:10.1061/(ASCE)HY.1943-7900.0000704

- Hannah CR (1978) Scour at pile groups. In: Research report no. 28-3, Civil Engineering Department, University of Canterbury, Christchurch
- Kothyari UC, Kumar A (2012) Temporal variation of scour around circular compound piers. *J Hydraul Eng* 138(11):945–957. doi:[10.1061/\(ASCE\)HY.1943-7900.0000593](https://doi.org/10.1061/(ASCE)HY.1943-7900.0000593)
- Kothyari UC, Ranga Raju KG, Garde RJ (1992) Live-bed scour around cylindrical bridge piers. *J Hydraul Res* 30(5):701–715. doi:[10.1080/00221689209498889](https://doi.org/10.1080/00221689209498889)
- Kumar A, Kothyari UC, Ranga Raju KG (2003) Scour around compound bridge piers. In: Proceedings of the 30th IAHR congress, Thessaloniki, pp 309–316
- Lauchlan CS, Melville BW (2001) Riprap protection at bridge piers. *J Hydraul Eng* 127(5):412–418. doi:[10.1061/\(ASCE\)0733-9429\(2001\)127:5\(412\)](https://doi.org/10.1061/(ASCE)0733-9429(2001)127:5(412))
- Larsen EM, Toch A (1956) Scour around bridge piers and abutments. In: HR-30 and Iowa highway research board bulletin, vol 4
- Lu JY, Shi ZZ, Hong JH, Lee JJ, Raikar RV (2011) Temporal variation of scour depth at nonuniform cylindrical piers. *J Hydraul Eng* 137(1):45–56. doi:[10.1061/\(ASCE\)HY.1943-7900.0000272](https://doi.org/10.1061/(ASCE)HY.1943-7900.0000272)
- Melville BW (1997) Pier and abutment scour: integrated approach. *J Hydraul Eng* 123(2):125–136. doi:[10.1061/\(ASCE\)0733-9429\(1997\)123:2\(125\)](https://doi.org/10.1061/(ASCE)0733-9429(1997)123:2(125))
- Melville BW, Chiew YM (1999) Time scale for local scour at bridge piers. *J Hydraul Eng* 125(1):59–65. doi:[10.1061/\(ASCE\)0733-9429\(1999\)125:1\(59\)](https://doi.org/10.1061/(ASCE)0733-9429(1999)125:1(59))
- Melville BW, Coleman SE (2000) Bridge scour. Water Resources Publications, LLC, Highlands Ranch
- Melville BW, Raudkivi AJ (1996) Effects of foundation geometry on bridge pier scour. *J Hydraul Eng* 122(4):203–209. doi:[10.1061/\(ASCE\)0733-9429\(1996\)122:4\(203\)](https://doi.org/10.1061/(ASCE)0733-9429(1996)122:4(203))
- Mia MF, Nago H (2003) Design method of time dependent local scour at circular bridge pier. *J Hydraul Eng* 129(6):420–427. doi:[10.1061/\(ASCE\)0733-9429\(2003\)129:6\(420\)](https://doi.org/10.1061/(ASCE)0733-9429(2003)129:6(420))
- Monti R (1994) Indagine sperimentale delle caratteristiche fluidodinamiche del campo di moto intorno ad una pila circolare. Ph.D. thesis, Politecnico di Milano, Milan (**in Italian**)
- Moreno M, Maia R, Couto L (2015) Effects of relative column width and pile-cap elevation on local scour depth around complex piers. *J Hydraul Eng* 142(2):04015051. doi:[10.1061/\(ASCE\)HY.1943-7900.0001080](https://doi.org/10.1061/(ASCE)HY.1943-7900.0001080)
- Oliveto G, Hager WH (2002) Temporal evolution of clear-water pier and abutment scour. *J Hydraul Eng* 128(9):811–820. doi:[10.1061/\(ASCE\)0733-9429\(2002\)128:9\(811\)](https://doi.org/10.1061/(ASCE)0733-9429(2002)128:9(811))
- Oliveto G, Hager WH (2005) Further results to time-dependent local scour at bridge elements. *J Hydraul Eng* 131(2):97–105. doi:[10.1061/\(ASCE\)0733-9429\(2005\)131:2\(97\)](https://doi.org/10.1061/(ASCE)0733-9429(2005)131:2(97))
- Raudkivi AJ (1986) Functional trends of scour at bridge piers. *J Hydraul Eng* 112(1):1–13. doi:[10.1061/\(ASCE\)0733-9429](https://doi.org/10.1061/(ASCE)0733-9429)
- Shen HW, Schneider VR, Karaki SS (1969) Local scour around bridge piers. *J Hydraul Div* 95(HY6):1919–1940
- Sumer BM, Christiansen N, Fredsøe J (1993) Influence of cross section on wave scour around piles. *J Waterway Port Coast Ocean Eng* 119(5):477–495. doi:[10.1061/\(ASCE\)0733-950X\(1993\)119:5\(477\)](https://doi.org/10.1061/(ASCE)0733-950X(1993)119:5(477))
- Tafarojnoruz A, Gaudio R, Grimaldi C, Calomino F (2010) Required conditions to achieve the maximum local scour depth at a circular pier. In: Proceedings of the 32 Convegno Nazionale di Idraulica e Costruzioni Idrauliche, Palermo
- Tafarojnoruz A, Gaudio R, Calomino F (2012) Bridge pier scour mitigation under steady and unsteady flow conditions. *Acta Geophys* 60(4):1076–1097. doi:[10.2478/s11600-012-0040-x](https://doi.org/10.2478/s11600-012-0040-x)
- Wang H, Tang H, Liu Q, Wang Y (2016) Local scouring around twin bridge piers in open-channel flows. *J Hydraul Eng* 142(2):04015051. doi:[10.1061/\(ASCE\)HY.1943-7900.0001154](https://doi.org/10.1061/(ASCE)HY.1943-7900.0001154)
- Yanmaz AM (2006) Temporal variation of clear water scour at cylindrical bridge piers. *Can J Civ Eng* 33(8):1098–1102. doi:[10.1139/106-054](https://doi.org/10.1139/106-054)
- Yanmaz AM, Altinbilek HDGA (1991) Study of time-dependent local scour around bridge piers. *J Hydraul Eng* 117(10):1247–1268. doi:[10.1061/\(ASCE\)0733-9429](https://doi.org/10.1061/(ASCE)0733-9429)

Microseismic event denoising via adaptive directional vector median filters

Jing Zheng^{1,2} · Ji-Ren Lu² · Tian-Qi Jiang² · Zhe Liang³

Received: 10 January 2017 / Accepted: 16 January 2017 / Published online: 27 January 2017
© Institute of Geophysics, Polish Academy of Sciences & Polish Academy of Sciences 2017

Abstract We present a novel denoising scheme via Radon transform-based adaptive vector directional median filters named adaptive directional vector median filter (AD-VMF) to suppress noise for microseismic downhole dataset. AD-VMF contains three major steps for microseismic downhole data processing: (i) applying Radon transform on the microseismic data to obtain the parameters of the waves, (ii) performing S-transform to determine the parameters for filters, and (iii) applying the parameters for vector median filter (VMF) to denoise the data. The steps (i) and (ii) can realize the automatic direction detection. The proposed algorithm is tested with synthetic and field datasets that were recorded with a vertical array of receivers. The P-wave and S-wave direct arrivals are properly denoised for poor signal-to-noise ratio (SNR) records. In the simulation case, we also evaluate the performance with mean square error (MSE) in terms of signal-to-noise ratio (SNR). The result shows that the distortion of the proposed method is very low; the SNR is even less than 0 dB.

Keywords Microseismic · Denoising · Adaptive directional vector median filter (AD-VMF) · Vector median filter (VMF)

Introduction

Underground rockburst will induce a microseismic event at an unknown origin time and unknown location (Jones et al. 2014). Locating the microseismic events accurately is of paramount importance for studying the state of stress and fracturing in rock mass (Kwiatek et al. 2010; Usher et al. 2013). Usually, making an estimate of the hypocenter location and origin time needs to observe the arrival times of the P- and S-waves of an event accurately. However, because of the low signal-to-noise ratio (SNR), the P- and S-waves are usually drawn in noise and the arrival times are hard to recognize (Kwietniak 2015; Kushnir et al. 2014). Suppressing the noise and enhancing the signals are applied to improve the accuracy to recognize the P- and S-waves.

Noise-suppression has drawn a lot of attention because of the importance to precondition the microseismic data. Vera Rodriguez et al. (2012) applied a time frequency basis pursuit denoising algorithm on multicomponent microseismic data. Forghani-Arani et al. (2013) introduced τ -p transform to suppress noise for microseismic data. Velis et al. (2015) developed method to denoise microseismic phase arrivals using a reduced-rank approximation of the singular value decomposition of the data along the detected phase arrivals. Sabbione et al. (2015) developed a denoising scheme apex-shifted parabolic Radon transform for microseismic data. Han and van der Baan (2015) proposed a scheme for Microseismic denoising via ensemble empirical mode decomposition and adaptive thresholding.

Microseismic monitoring can be treated as a passive seismic monitoring. The microseismic and seismic data have many similarities. Many denoising methods for seismic datasets can be tried to transplant the application in the microseismic datasets after modifying according to the

✉ Jing Zheng
zhengjing8628@163.com

¹ State Key Laboratory of Coal Resources and Safe Mining, China University of Mining and Technology, Beijing, China

² College of Geoscience and Surveying Engineering, China University of Mining and Technology, Beijing, China

³ Anhui University of Science and Technology, Huainan, China

characteristics of the data. In this paper, we want to develop a novel scheme for microseismic downhole dataset using an adaptive directional vector median filter (AD-VMF). The AD-VMF is designed according to the arrivals relationships between multi-channels for borehole microseismic geophones. The method contains two major parts. First, the possible signal directions are detected by Radon transform (Deans 1984) to seek the arrivals in noisy multichannel data, and signal lengths are detected by S-transform (Stockwell 2007). Then, the microseismic phase arrivals are denoised via VMF, whose parameters are designed by the first part, including the direction and the window size.

The paper is organized as follows. In “Methodologies”, we briefly introduce the basis for microseismic event and modified vector median filter (VMF) processing for microseismic downhole dataset. We divide “Proposed method” into two parts: in the first part, we introduce the technique for AD-VMF according to the characteristics of the microseismic arrivals in details; in the second part, the performance of the AD-VMF is analyzed. Finally, conclusions are provided in “Conclusions”.

Methodologies

Travel-time curves for microseismic events

When monitoring is done in the borehole, each source and the geophones can be treated in a 2D plane. We can assume that the vertical array of j receivers with coordinates (x, z_j) , and the source of microseismic event with coordinates (x_0, z_0) , as described in Fig. 1. Treating the medium as a constant velocity medium, the travel-times for a seismic event occurring at time t_0 can be calculated by the equation as follows:

$$\begin{cases} t_p(z_j) = t_0 + \frac{1}{V_p} \sqrt{(x-x_0)^2 + (z_j-z_0)^2} \\ t_s(z_j) = t_0 + \frac{1}{V_s} \sqrt{(x-x_0)^2 + (z_j-z_0)^2} \end{cases} \quad (1)$$

where V_p and V_s are velocities of P- and S-waves, respectively. According to Sabbione et al. (2015), the arrival times can be approximated by:

$$\begin{cases} t_p(z_j) = \tau_p + q_p(z_j - z_0)^2 \\ t_s(z_j) = \tau_s + q_s(z_j - z_0)^2 \end{cases} \quad (2)$$

where $\tau_p = t_0 + \frac{(x-x_0)}{V_p}$, $\tau_s = t_0 + \frac{(x-x_0)}{V_s}$, $q_p = \frac{1}{2(x-x_0)V_p}$, $q_s = \frac{1}{2(x-x_0)V_s}$.

This means that we can fit the arrival times of the microseismic datasets using parabolic curves. To get the fitting curve, the microseismic events should be first

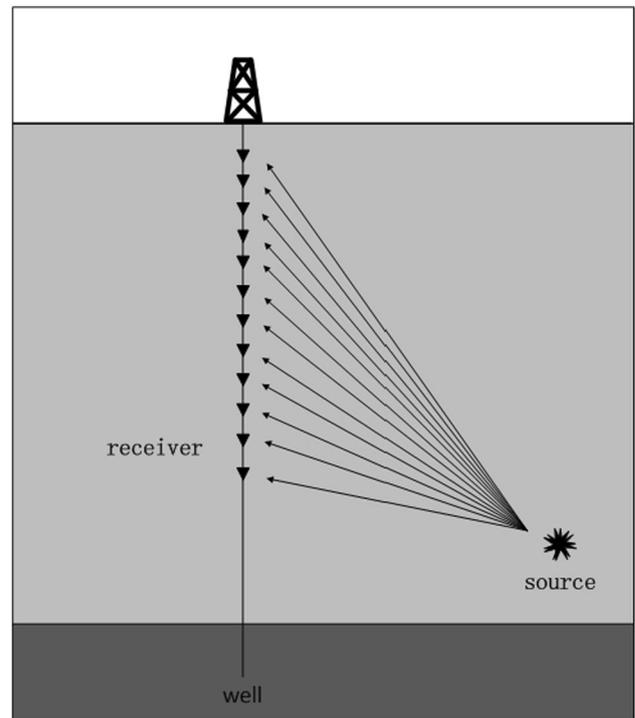


Fig. 1 Schematic diagram of the microseismic monitoring via borehole geophones

detected. Radon transform is suggested to detect the curve and get the parameters of the curve.

Radon transform has a Linear Radon transform and a Parabolic Radon transform. The Parabolic Radon transform should be applied on the microseismic datasets because of Eq. 2. The Parabolic Radon transform and its inverse transformation formula are given in Eqs. 3 and 4, respectively:

$$m(\tau, q) = \int_{-\infty}^{\infty} d(t = \tau + qz^2, z) dz \quad (3)$$

$$d(t, z) = \int_{-\infty}^{\infty} m(\tau = t - qz^2, q) dq \quad (4)$$

To implement the Radon transform on a computer, discrete calculation can be given. Assuming the microseismic data can be described using a given matrix \mathbf{X}_0 of size $N \times M$, where M is the number of receivers and N is the number of sample points of each receiver. The parameter of time (t_p or t_s) can be discrete by the sampling interval T_s , and expressed by $t_i = t_h + (i - 1)T_s$, $i = 1, 2, \dots, N$, where t_h means the start time of the given datasets. The parameter of receiver’s position at z -axis can be discrete by the space interval of geophones d_z , and expressed by $z_j = z_0 + (j - 1)d_z$, $j = 1, 2, \dots, M$, where z_0 means the start position of the first geophone. The data in the matrix \mathbf{X}_0 can be expressed as $x_0(i, j)$. Applying discrete Radon transform on the data matrix, we can detect the

P- and S-waves of microseismic events can get the estimation of the parameters.

Filter processing

After fitting the curve of arrival time vs the receiver's position, a filter window can be designed along the curve and VMF can be applied to filter the microseismic data caused by the same source with same wave type. The VMF has drawn a lot of attention in image processing area. Astola et al. (1990) proved that the VMF can reduce noise while preserving sharp boundaries in color images, and compared the performance with scalar median filter (SMF). Liu et al. (2009) and Liu (2013) introduced the VMF into seismic data denoising processing, and compared the performance with SMF, f-x deconvolution, and mean filters. The result verified that the VMF can reduce noise while preserving discontinuities very effectively.

Application of VMF

We assume that $\mathbf{X}_0 = \{x_0(i, j); i = 1, 2, \dots, N_1; j = 1, 2, \dots, M\}$ is the matrix consisting of original data vectors without noise, where N_1 is the length of the signal, M is the number of receivers, $\mathbf{X} = \{x(i, j); i = 1, 2, \dots, N_1; j = 1, 2, \dots, M\}$ is the matrix consisting of noised data vectors, and $\hat{\mathbf{X}} = \{\hat{x}(i, j); i = 1, 2, \dots, N_1; j = 1, 2, \dots, M\}$ is the matrix consisting of the filtered data vectors processed by VMF. The VFM is applied for the noised data matrix $\mathbf{X} = \{x(i, j); i = 1, 2, \dots, N_1; j = 1, 2, \dots, M\}$. Given the filtering window $n_1 \times n_1$, the processing scheme using VMF in this window can be realized as:

1. Calculating the sum of the distances between each vector and the other vectors in the filtering window as the following equation:

$$d_{ij} = \sum_{l=1}^{n_1} \sum_{k=1}^{n_1} \|x_{ij} - x_{kl}\| \quad i = 1, 2, \dots, n_1; j = 1, 2, \dots, n_1. \quad (5)$$

2. Comparing with all the d_{ij} to find the smallest distance d_{\min} ;
3. Corresponding to this d_{\min} , the associated median vector is determined, then it is used to replace the current center vector in the filtering window;

Moving the filtering window and repeating step 1, 2, and 3 until all the data is filtered, then the filtered data matrix $\hat{\mathbf{X}} = \{\hat{x}(i, j); i = 1, 2, \dots, N_1; j = 1, 2, \dots, M\}$ is obtained.

When we use the VMF, we need to determine the parameter of the VMF (i.e., the size of the window) at first. In the VMF, the filter window design is of paramount

importance for the noise-suppression and edge-preserving. To determine the best size for VMF, the S-transform is applied. It should be noted that if the durations of the microseismic signals can be obtained or estimated at first, the use of S-transform can be ignored.

Application of S-transform

The S-transform produces a time frequency representation of a time series. It can be derived from short time Fourier or wavelet transform, and has some unique advantages: its time–frequency resolution is related to frequency resolution and can be treated as a special case of the Fourier Transform with the characteristic of multiresolution. According to Stockwell (2007), the S-transform of a function $h(t)$ can be defined as a wavelet transform with a specific mother wavelet multiplied by a phase factor:

$$S(\tau, f) = \int_{-\infty}^{\infty} h(t) w(\tau - t) e^{-j2\pi ft} dt. \quad (6)$$

where $w(t)$ is the Gaussian function defined as:

$$w(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}}. \quad (7)$$

The parameter σ represents the window width in time domain. However, it is also a frequency-dependent quantity as follows:

$$\sigma = \frac{1}{|f|} \quad (8)$$

We should note that the basic wavelet of S-transform does not have to meet the admissibility condition.

Using S-transform, the time–frequency spectrum of the data collected by each geophone can be obtained. As mentioned by Zheng et al. (2013), the main frequency, delay and duration information of the seismic wavelets can be estimated by S-transform effectively. Assuming there is one microseismic event in the data matrix, and P- and S-waves are contained. Therefore, the data matrix can be separated into three parts: the sub-part of P-wave, sub-part of S-wave and the residual. Combining with the analysis using Radon, the positions of P- and S-waves in delay axis in S-transformed domain can be confirmed and can be treated as centers of signals to be processed. Setting a threshold, we can find out the relatively strong value of power in the S-transformed plane around the centers. The detection results are shown in Figs. 5 and 11. The marked edge in the delay axis can be treated as the effective duration of the signals to be processed and noted as L . It should be noted that the duration of different waves is not the same; therefore, L is different if the duration of the signals is not the same.

The partition size of the P- and S-waves to be processed in the time axis can be chosen according to the duration L obtained above. The size can be chosen between L and $2L$. And the length of the window in the VMF can be chosen between $L/8$ and $L/4$, and the shape of the window can be chosen as square. If the duration of P-wave is 20-sample points, the partition size of the P-wave to be processed can be chosen as 40-sample points. If the number of receivers is 12, then the size of the dataset separated from the total dataset which contains P-wave and need to be filtered by VMF is 40×12 . The size of the window of VMF can be set as 5×5 . It means that we apply a 5×5 VMF to process a dataset whose size is 40×12 .

Proposed method

This section includes two parts: first is the introduction of the proposed scheme, second is the experiments and analysis. In the second part, the experiments will be achieved using not only synthetic data but also field data to verify the performance of the proposed scheme.

Proposed scheme of AD-VMF

The scheme proposed for microseismic denoising is described in Fig. 2. Major steps are given as follows: first apply Radon transform on the microseismic data to detect the microseismic events and obtain the parameters of the P- and S-waves; then perform S-transform to determine the parameters for filter design; finally, adaptive VMF is applied to denoise the data. For a certain time series $x(i, j)$, $i = 1, 2, \dots, N$; $j = 1, 2, \dots, M$, the processed steps of AD-VMF can be shortly described as:

1. Apply Radon transform which is described in “Travel-time curves for microseismic events” on the series $x(i, j)$ to detect the P- and S-waves for every microseismic event and obtain the rough arrival times of the

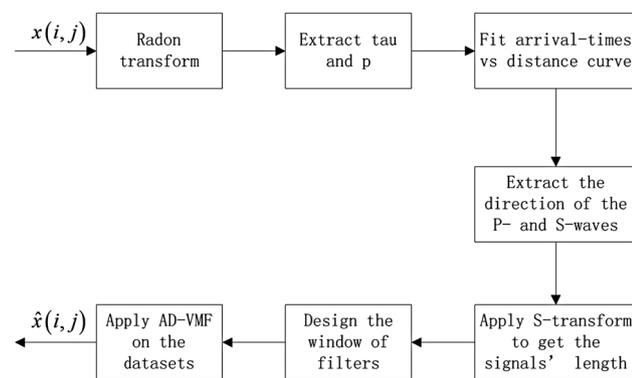


Fig. 2 Proposed scheme for AD-VMF

2. Get the time–frequency spectrum $S_j(k, m)$ by transposing the S-transformed results of the microseismic data $x(i, j)$; the S-transform can be implemented as described in “Application of VMF”, where j is the channel number, m is the delay information and k is the spectral information; it should be noted that $k = 0, 1, \dots, N - 1$, $m = 0, 1, \dots, N - 1$, where N is the length of data to be analyzed for each channel;
3. Search and estimate the length L for each type of waves for every microseismic event in the S-transformed plane with the information of arrival times of each geophone;
4. Design the filter window for VMF; the direction and center of the filter are determined by the fitted curve, the scope of one curve can be chosen between L and $2L$, the length of the window can be chosen between $L/8$ and $L/4$;
5. Apply the filter on the data series $x(i, j)$ as described in “Application of VMF” and get the denoised data $\hat{x}(i, j)$.

The scheme will be analyzed and verified by synthetic signals and field datasets.

Tests

In this part, the effect of the proposed method is tested by synthetic data. Figure 3 is a synthetic microseismic record used in this example. Figure 3a is the original data without noise, and Fig. 3b is the noised data at SNR 0 dB. The event is delayed by the same wavelet (i.e., Ricker wavelet of 50 Hz) in accordance with the different travel-times. In addition, the weighted coefficients are different for P-wave and S-waves of the event, and these coefficients are

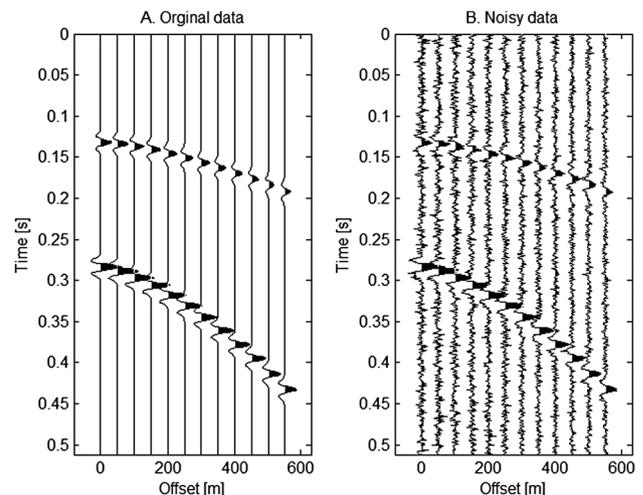


Fig. 3 Synthetic data with original and noisy data

attenuated by exponential law among the channels. The purpose is to simulate the attenuation effect of different propagation paths on seismic waves.

Now we focus on the processing of the noisy signals shown in Fig. 3b. The Radon transformed results are shown in Fig. 4a. With the parameters detected in Fig. 4a given by Radon transformed results, we can fit the curve of arrival times. The fitting results are shown in Fig. 4b. At the same time, we apply S-transform on the noised data and achieve the parameter of the scope containing the P- or S-wave of an event. The detection result of the first geophone is shown in Fig. 5. Finally, VMF is applied, and the filtered result is shown in Fig. 6a. To further demonstrate the performance of this method, we added the results obtained by normal VMF and showed it in Fig. 6b. From the denoised figures, we note that the arrival times can be clearly shown in the corresponding figure via the proposed method and the random noise can be effectively suppressed.

In the simulation case, we can evaluate the performance by the mean square error (MSE). MSE is defined as the ratio of the square of the difference between the original data and the filtered data to the number of elements. The expression is

$$\text{MSE} = \frac{\sum_{n=1}^N (\hat{s}(n) - s(n))^2}{N}, \quad (9)$$

where $\hat{s}(n)$ refers to the data after filtering, $s(n)$ to the original data, N is the sampling number for microseismic signals of each channel. It can not only evaluate the enhancement of SNR after denoising, but also measure the distortion between the signals processed by the proposed method and the original signals without noise. The evaluation result is shown in Fig. 7.

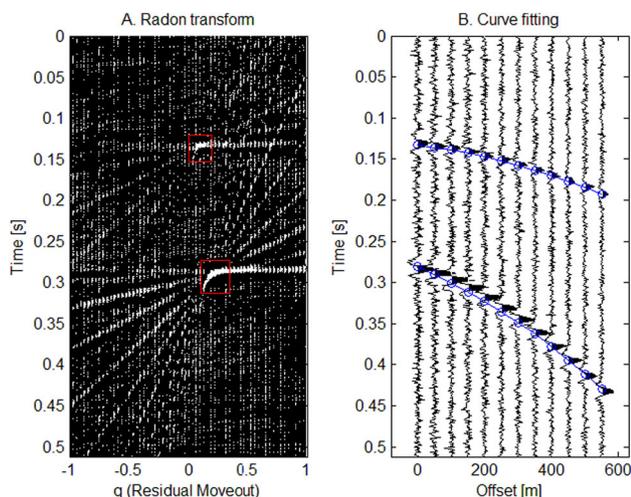


Fig. 4 Radon transformed and curve fitting results

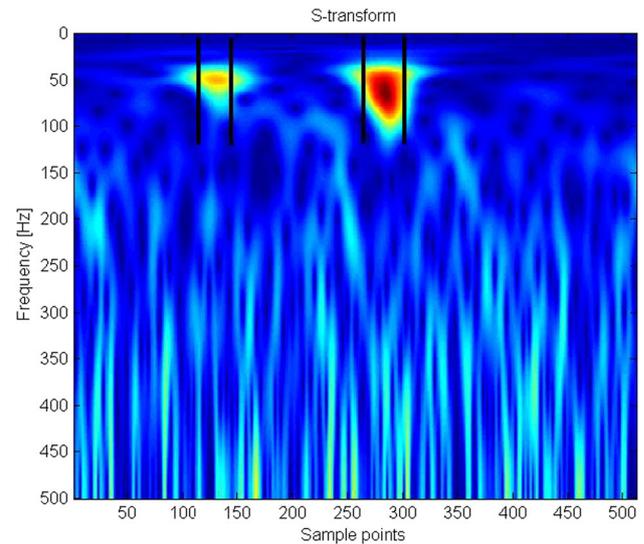


Fig. 5 S-transformed result of the data recorded by the first geophone

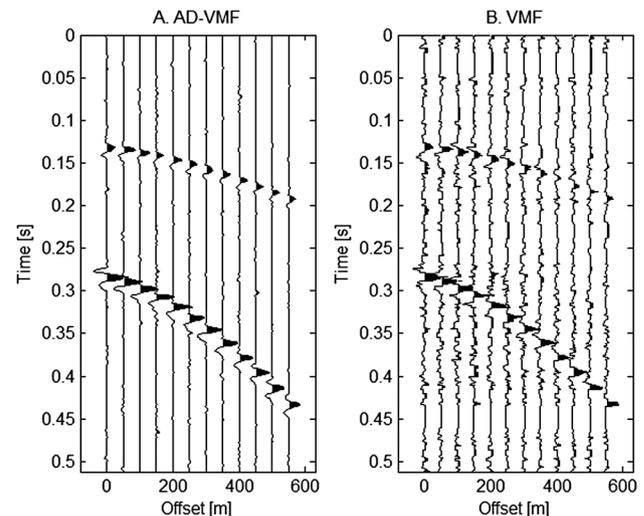


Fig. 6 Denoised results via AD-VMF

To verify the performance of the proposed method, we added an example in which the SNR is -10 dB. The noisy dataset is shown in Fig. 8a. It can be found that the noise level is high and the picking of P- and S- arrivals is difficult. Figure 8b showed the processed dataset. We can find that although the noise cannot be completely suppressed in Fig. 8b, the picking of P- and S- arrivals can be done more easily in Fig. 8b. It means that the method can process data at low SNR. We should note that when the noise level is higher, the filtering performance of the method becomes worse.

To evaluate the performance further, we process and analyze a group of field data via the proposed scheme. The actual data contains the microseismic event in a coal mine. The observation system is a downhole string of 12 geophones. The space distance is 10 m, and the sampling

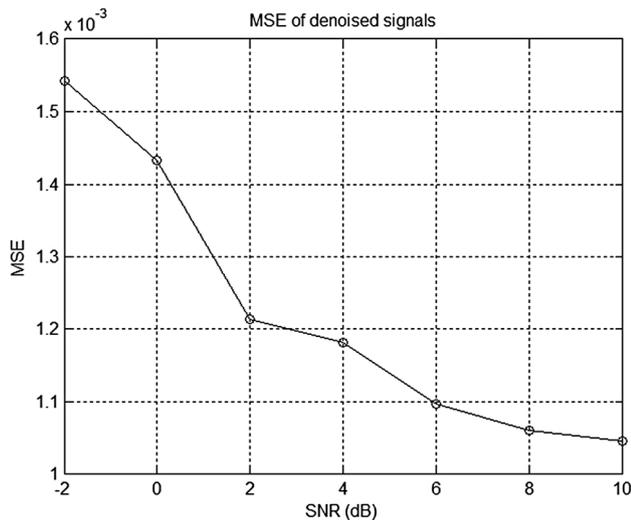


Fig. 7 MSE of the signals filtered by AD-VMF

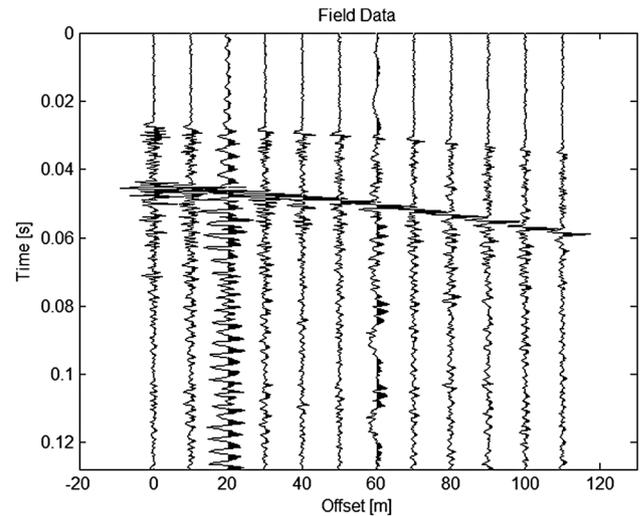


Fig. 9 Field data example with processed data

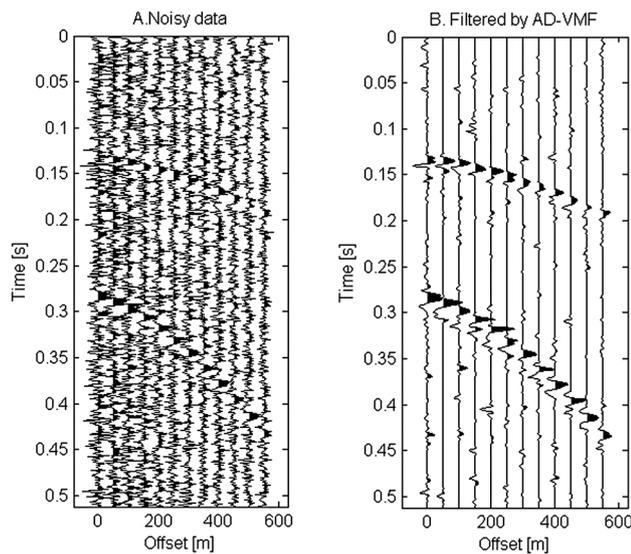


Fig. 8 Example with high level noise at SNR -10 dB

frequency is 4 kHz. Both P- and S-waves are generated by an effective microseismic event. The original data is shown in Fig. 9. From the original record, it was difficult for us to ascertain the arrival time for all the channels, especially of the P-wave.

Taking the same steps as those mentioned in “Proposed scheme of AD-VMF”, the P- and S-waves are detected in Radon domain (Fig. 10a). After fitting the curve, the arrival time can be achieved as shown in Fig. 10b. The window length can be estimated by S-transform. Figure 11 is an example for the data received by the first geophone. Finally, VMF is applied; the denoised result is shown in Fig. 12. We can find that after using the proposed method, the valid signal can be enhanced while the noise can be removed. However, the noised effect depends on the results

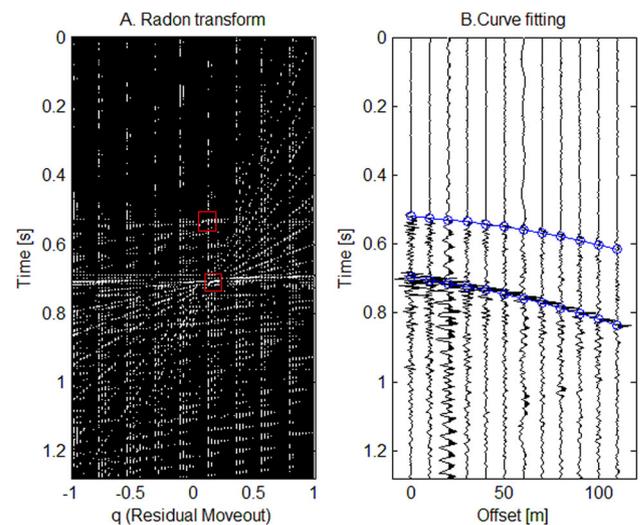


Fig. 10 Radon transformed and curve fitting results

of the event detection. If the events cannot be detected, the denoised signals will distort and miss some waves. For further research, a more effective method to detect the events automatically will be studied.

Conclusions

In this paper, an algorithm to denoise the microseismic datasets based on modified VMF method is proposed. The novel method contains three major steps for microseismic downhole data processing: (1) applying Radon transform on the microseismic data to detect the travel-times of microseismic events and obtain the parameters of the P- and S-waves, (2) performing S-transform to determine the

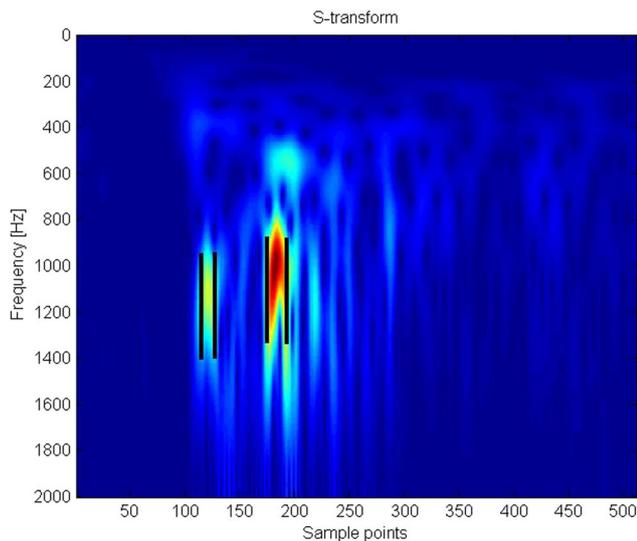


Fig. 11 S-transformed result of the data recorded by the first geophone

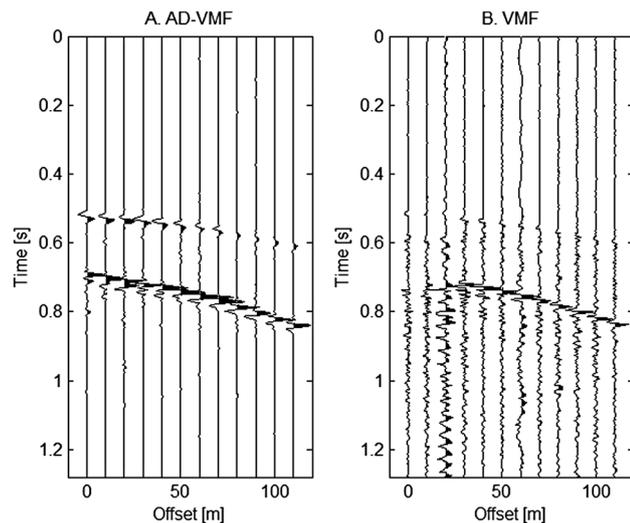


Fig. 12 Denoised result of field dataset

parameters for filter design and (3) applying the AD-VMF to denoise the data.

Not only the waveforms with higher energy (the S-wave in our examples) can be detected clearly in the denoised figures, but also the waveforms with lower energy (the P-wave in our examples) can be detected clearly in the denoised results. The test over a noisy synthetic dataset demonstrates that the algorithm significantly enhanced the SNR of the input data for both the P- and the S-wave arrivals. The MSE performance in the simulation case shows that the proposed method can reserve the useful waveforms better than the normal VMF while the noise is suppressed. Regarding the field data examples, the method succeeded in detecting and denoising the event for the poor

SNR. The P- and S-wave arrivals can be clearly recognized in the denoised version.

Acknowledgements The work was supported by the National Natural Science Foundation of China (Grant No. 41504041), State Key Laboratory of Coal Resources and Safe Mining, China University of Mining and Technology (No. SKLCSRSM14KFA05). The authors appreciate the editors Prof. Zbigniew Wisniewski, Prof. Filippos Vallianatos and anonymous reviewers for the guidance and help. The authors also thank Suping Peng for his support.

References

- Astola J, Haavisto P, Neuvo Y (1990) Vector median filter. *Proc IEEE* 78:678–689. doi:[10.1109/5.54807](https://doi.org/10.1109/5.54807)
- Deans SR (1984) Book-review—the radon transform and some of its applications. *Science* 223(1):3–4
- Forghani-Arani F, Willis M, Haines S, Batzle M, Behura J, Davidson M (2013) An effective noise-suppression technique for surface microseismic data. *Geophysics* 78(6):KS85–KS95. doi:[10.1190/GEO2012-0502.1](https://doi.org/10.1190/GEO2012-0502.1)
- Han J, van der Baan M (2015) Microseismic and seismic denoising via ensemble empirical mode decomposition and adaptive thresholding. *Geophysics* 80(6):69–80. doi:[10.1190/geo2014-0423.1](https://doi.org/10.1190/geo2014-0423.1)
- Jones GA, Kendall J-M, Bastow I, Raymer DG, Wuestefeld A (2014) Characterization of fractures and faults: a multi-component passive microseismic study from the Ekofisk reservoir. *Geophys Prospect* 62:779–796. doi:[10.1111/1365-2478.12139](https://doi.org/10.1111/1365-2478.12139)
- Kushnir A, Varypaev A, Dricker I, Rozhkov M, Rozhkov N (2014) Passive surface microseismic monitoring as a statistical problem: location of weak microseismic signals in the presence of strongly correlated noise. *Geophys Prospect* 62:819–833. doi:[10.1111/1365-2478.12124](https://doi.org/10.1111/1365-2478.12124)
- Kwiatk G, Bohnhoff M, Dresen G, Schulze A, Schulte T, Zimmermann G, Huenges E (2010) Microseismicity induced during fluid-injection: a case study from the geothermal site at Groß Schönebeck, north German Basin. *Acta Geophys* 58(6):995–1020. doi:[10.2478/s11600-010-0032-7](https://doi.org/10.2478/s11600-010-0032-7)
- Kwiatniak A (2015) Detection of the long period long duration (LPLD) events in time- and frequency-domain. *Acta Geophys* 63(1):201–213. doi:[10.2478/s11600-014-0262-1](https://doi.org/10.2478/s11600-014-0262-1)
- Liu Y, Luo Y, and Wang Y (2009) Vector median filter and its applications in geophysics. In: 79th Annual International SEG Meeting, Expanded Abstracts, pp 3342–3347, doi: [10.1190/1.3255554](https://doi.org/10.1190/1.3255554)
- Liu Y (2013) Noise reduction by vector median filtering. *Geophysics* 78(3):V79–V86. doi:[10.1190/geo2012-0232.1](https://doi.org/10.1190/geo2012-0232.1)
- Sabbione JI, Sacchi MD, Velis DR (2015) Radon transform-based microseismic event detection and signal-to-noise ratio enhancement. *J Appl Geophys* 113:51–63. doi:[10.1016/j.jappgeo.2014.12.008](https://doi.org/10.1016/j.jappgeo.2014.12.008)
- Stockwell RG (2007) A basis for efficient representation of the S-transform. *Digit Signal Process* 17:371–393. doi:[10.1016/j.dsp.2006.04.006](https://doi.org/10.1016/j.dsp.2006.04.006)
- Usher PJ, Angus DA, Verdon JP (2013) Influence of a velocity model and source frequency on microseismic waveforms: some implications for microseismic locations. *Geophys Prospect* 61(s1):334–345. doi:[10.1111/j.1365-2478.2012.01120.x](https://doi.org/10.1111/j.1365-2478.2012.01120.x)
- Velis D, Sabbione JI, Sacchi MD (2015) Fast and automatic microseismic phase-arrival detection and denoising by pattern recognition and reduced-rank filtering. *Geophysics* 80(6):25–38. doi:[10.1190/geo2014-0561.1](https://doi.org/10.1190/geo2014-0561.1)

- Vera Rodriguez I, Bonar D, Sacchi M (2012) Microseismic data denoising using a 3C group sparsity constrained time-frequency transform. *Geophysics* 77(2):V21–V29. doi:[10.1190/geo2011-0260.1](https://doi.org/10.1190/geo2011-0260.1)
- Zheng J, Peng SP, Liu M, Liang Z (2013) A novel seismic wavelet estimation method. *J Appl Geophys* 90:92–95. doi:[10.1016/j.jappgeo.2013.01.007](https://doi.org/10.1016/j.jappgeo.2013.01.007)

A theoretical derivation of the dilatancy equation for brittle rocks based on Maxwell model

Jie Li^{1,2} · Houxu Huang¹ · Mingyang Wang^{1,2}

Received: 10 January 2017 / Accepted: 16 January 2017 / Published online: 28 January 2017
© Institute of Geophysics, Polish Academy of Sciences & Polish Academy of Sciences 2017

Abstract In this paper, the micro-cracks in the brittle rocks are assumed to be penny shaped and evenly distributed; the damage and dilatancy of the brittle rocks is attributed to the growth and expansion of numerous micro-cracks under the local tensile stress. A single crack's behaviour under the local tensile stress is generalized to all cracks based on the distributed damage mechanics. The relationship between the local tensile stress and the external loading is derived based on the Maxwell model. The damage factor corresponding to the external loading is represented using the p -alpha (p - α) model. A dilatancy equation that can build up a link between the external loading and the rock dilatancy is established. A test of dilatancy of a brittle rock under triaxial compression is conducted; the comparison between experimental results and our theoretical results shows good consistency.

Keywords Subcritical crack growth · Dilatancy · Distributed damage mechanics · Maxwell model · p - α Model

Introduction

In rock engineering, the dilatancy before macro-failure is a manifest characteristic of rocks (Cook 1970; Cristescu 2002). Dilatancy plays a crucial role in the stability of

underground structures (Tan 1982; Bosman et al. 2000) and also has a close relation with dynamic disasters such as rockburst and seismicity. Direct evidence of crustal dilatancy has been observed prior to earthquakes. Scientists, such as Scholz (Scholz et al. 1973) and Tan (Tan and Kang 1983), took the dilatancy-expansion mode as a physical basis for seismic prediction.

There are many experimental studies of the dilatancy of rocks (Bridgman 1949; Brace et al. 1966; Scholz 1968; Cook 1970; Cogan 1976; Tapponnier et al. 1976; Tan et al. 1989; Haimson and Chang 2000; Yang et al. 2012; Zhang et al. 2015), which have contributed greatly to our understanding of rock dilatancy. These studies provide us with perfect stress-strain curves and clear thresholds corresponding to different stages of compression. But for the practical engineering, it would be convenient and helpful if we knew the rock dilatancy corresponding to any given external loading conditions of the rock masses. However, the theoretical studies of dilatancy, especially the ones that can directly reflect the correlation between the dilatancy and the external loading, are rare. Besides, the dilatancy of brittle rocks is mainly the result of numerous micro-cracks' growth and expansion; to build up the link between the dilatancy and the development of cracks, we need to describe the behaviour of numerous cracks, but the difficulty is that the growth and expansion of cracks in the process of rock deformation, which involve many factors and span several scales, are so complicated that there is no effective method to analyse them at present. Moreover, although the condition for the occurrence and expansion for a single crack can be successfully evaluated based on the fracture mechanics, there would be an unsolvable problem if numbers of cracks are given a separated study. What is more, although the effect of a single crack's growth and expansion on the rock can be well evaluated, it

✉ Jie Li
lijierf@163.com

¹ State Key Laboratory of Disaster Prevention and Mitigation of Explosion and Impact, PLA University of Science and Technology, Nanjing, China

² School of Mechanical Engineering, Nanjing University of Science and Technology, Nanjing, China

still seems impossible to evaluate the effects of numerous cracks' behaviour on the rock.

According to the study of Kachanov (1986), during the deformation process, three factors, i.e. the dilatancy caused by the growth and expansion of cracks, the damage degree, and the stress distribution, are directly related to each other. Therefore, the most important and difficult problem is to establish the relationship among the aforementioned three factors. The study of Kachanov (1986) laid the foundation for the continuum damage mechanics and provided us with the basic direction for our further study. In this paper, the cracks in the cylindrical rock specimen are assumed to be penny shaped, in the same size, evenly distributed and their directions are paralleled to the maximum principal stress. The volume of rock specimen is divided into two parts: one is the volume of the elastic matrix, which is regarded to be invariant during the rock deformation; the other is the volume of cracks, which is affected by the micro-cracking process in the specimen. The damage degree is represented by parameter α , i.e. the ratio of the specimen's total volume to the matrix volume. The expression of the local tensile stresses that cause the growth and expansion of cracks is derived based on the Maxwell model. Besides, the derived expression also successfully establishes the correlation between the external loading which is at the macroscopic scale, and the internal local stress which is at the microscopic scale. A single crack's effect on the damage degree and the rock dilatancy is carefully analysed; based on the distributed damage mechanics, a single crack's behaviour is generalized to all cracks. Based on these, we finally establish the correlation between the external loading and the dilatancy equation for brittle rocks.

Damage factor

According to Seaman et al. (1976), two problems have to be addressed to describe the deformation and failure of continuum in the framework of the distributed damage mechanics: one is to choose a proper physical parameter as a damage variable; the other is to establish a corresponding dynamic evolution equation to describe the quantitative variation of the specified damage degree in the process of a body's deformation and failure.

p - α Model and the corresponding equations

The essences of the porous media model (p - α model) are:

- it takes the relative volume of micro-cracks and pores as a measure of materials' damage (Carroll and Holt 1972; Bhatt et al. 1975);

$$\alpha = V/V_m \quad (1)$$

- it builds up the link between micro-cracks development and macro-dilatancy through writing the medium unit volume V as the sum of volume of the fissures (pores) and matrix, i.e. V_p and V_m ;

$$V = V_p + V_m \quad (2)$$

- it assumes that under the state of compression, the external loading's work on the continuum is only accumulated in the matrix material;

$$pdV = p_m dV_m, \quad (3)$$

therefore, from Eqs. (1)–(3) we can obtain the relation between pressure in porous material p and that in the matrix, p_m .

$$p = p_m/\alpha. \quad (4)$$

Provided the equation of state (EOS) of the matrix is given as

$$p_m = p_m(\rho_m), \quad (5)$$

then the porous material's EOS can be written as

$$p = p_m(\alpha\rho)/\alpha. \quad (6)$$

For rock media, the EOS of the matrix material is assumed to obey the Hooke's law

$$p_m = K_m \varepsilon_{vm} = K_m \left(\frac{\rho_{m0}}{\rho_m} - 1 \right), \quad (7)$$

where ρ_0 and ρ are the initial density and density of the material at any moment; $\rho_{m0} = \alpha_0 \rho_0$ and $\rho_m = \alpha \rho$ are the densities of the matrix corresponding to parameters ρ_0 and ρ ; α_0 and α are the initial damage degree and damage degree at any moment of the material; K_m is the bulk modulus of the matrix.

By substituting Eq. (7) into Eq. (6), and considering the relations $p_m = \alpha p = K_m \left(\frac{\alpha_0 \rho_0}{\alpha \rho} - 1 \right)$, i.e. $\frac{\rho_0}{\rho} = \frac{\alpha}{\alpha_0} \left(\frac{2p}{K_m} + 1 \right)$, and $\frac{\rho_0}{\rho} = 1 + \varepsilon_v$, the rock's volumetric dilatancy equation is obtained as follows

$$\varepsilon_v = \frac{p\alpha^2}{K_m\alpha_0} + \frac{\alpha}{\alpha_0} - 1. \quad (8)$$

However, due to the fact that the α^2 term in Eq. (8) is small, a simpler form of Eq. (8) can be obtained by performing Taylor extension of the function $f(\alpha) = p\alpha^2/K_m\alpha_0$ and retaining the first two terms

$$\varepsilon_v = \frac{p\alpha_0^2}{K_m\alpha_0} + \frac{2p}{K_m}(\alpha - \alpha_0) + \frac{\alpha}{\alpha_0} - 1. \quad (9)$$

Because $|p| \ll K_m$, and for dense rock media, $\alpha_0 \approx 1$, we have

$$\frac{2|p|}{K_m}(\alpha - \alpha_0) = \frac{2\alpha_0|p|}{K_m} \left(\frac{\alpha}{\alpha_0} - 1 \right) << \frac{\alpha}{\alpha_0} - 1. \quad (10)$$

So, Eq. (8) can be written in a simpler form as

$$\varepsilon_v = \frac{p\alpha_0}{K_m} + \frac{\alpha}{\alpha_0} - 1. \quad (11)$$

According to Eq. (11), if we know the evolution of the $p(\alpha)$ equation corresponding to different stages of the rock compression process, then we can obtain the rock dilatancy. In the next section, we will attempt to establish the $p(\alpha)$ equation corresponding to different deformation and failure processes of the rock.

Establishment of dilatancy equations

Based on the existing studies, the stress–strain behaviour of brittle rocks in the compression process can be divided into four stages:

1. Closure of pre-existing cracks. This stage can be neglected in hard rocks at high confining pressures;
2. Elasticity stage. There is fairly good linearity in the stress–strain curves in this stage;
3. Development of dilatancy associated with subcritical crack growth. This stage starts when $\sqrt{J_2}$ exceeds the strength for shear dilatancy, f_3 (Tan et al. 1989; Martin and Read 1992; Szczepanik et al. 2003), where $J_2 = s_{ij}s_{ij}/2$ is the second invariant of the deviatoric stress tensor. Acoustic emissions and dilatancy are found in this stage;
4. Unstable microcracking leads to the growth of macroscopic fractures. This stage begins when $\sqrt{J_2}$ reaches the breaking strength σ_F (Tan et al. 1989), and in this stage the material is broken into fragments and can no longer be regarded as a continuum. The fracturing and relative movement of these fragments result in a large irreversible deformation.

According to the aforementioned statements, to simplify the following analyses, we will neglect the first stage and try to establish the $p(\alpha)$ equation corresponding to the remaining three stages of rock compression.

Elasticity stage

In the elasticity stage, the stress accumulation is not strong enough for crack initiation and propagation, so the fracture porosity remains constant; by substituting $\alpha = \alpha_0$ into Eq. (11), the following dilatancy equation is obtained:

$$\varepsilon_v = \varepsilon_{ve} = \frac{p\alpha_0}{K_m}. \quad (12)$$

Subcritical crack growth stage

The rock deformation goes into a subcritical crack growth stage when $K_I \geq K_{I0}$ (Das and Scholz 1981; Atkinson 1987). Because of the growth and expansion of the cracks in the rocks, dilatancy phenomena occur. The distribution of initial cracks in the rock is so complicated that there is no effective method of observation and predication. To avoid complex derivation, the micro-cracks are assumed to be penny shaped, evenly distributed, and in the same size with their normal perpendicular to the direction of the maximum principal stress. Besides, it is also assumed that there is no interaction between the neighbouring cracks (as shown in Fig. 1).

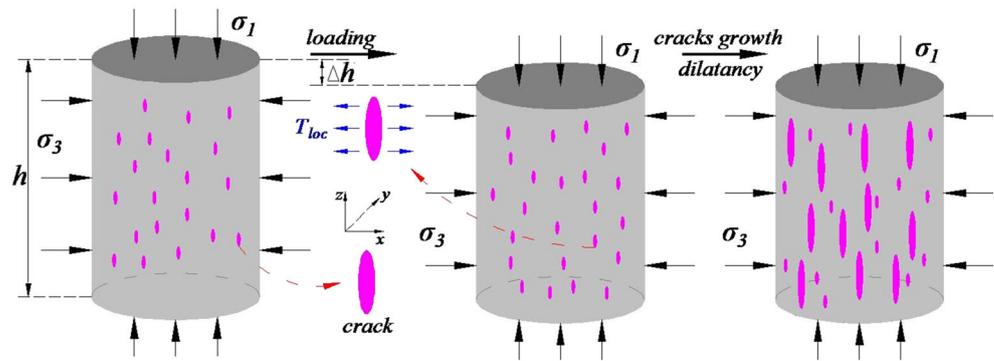
Local tensile stress

The brittle rock has complicated internal structure with a wide range of scales and often shows visible characteristics of tensile failure when subjected to compressive loading (Diederichs et al. 2004; Read 2004; Han et al. 2016). The deformation and fracture processes of brittle rock are mainly governed by elastic deformations and cracking, whereas the plastic deformation in brittle rock is insignificant and can be neglected (Qi et al. 2014). Direct observation (Tapponnier et al. 1976; Atkinson 1987) has shown that, for the most part, micro-fractures initiate from local stress concentrations resulting from mismatches in elastic properties along grain boundaries or natural flaws, driven by localized tensile stresses in the region near the stress concentration zone (Vásárhelyi and Bobet 2000; Paterson and Wong 2005). Cracks nucleate from pre-existing flaws and propagate predominantly as mode I (tensile opening) fractures (Ashby and Sammis 1990; Qian and Zhou 2011; Han et al. 2016) in a direction parallel to the maximum principal stress. Besides, in situ investigations as well as experimental and theoretical studies have all shown that the deformation and fracture of rock-like materials are all governed by the laws of Maxwell bodies and can be described using Maxwell model (Qi et al. 2014, 2016; Huang et al. 2016).

During the loading process, there are stress concentration, successive damage and fracture in the brittle rock. The occurrence of fracture will lead to stress relaxation; therefore, the stresses in the brittle rock can be divided into two components, one of which is the elastic stress caused by the reversible deformation, and the other is the local inelastic stress caused by the cracks which also exists around the cracks.

The relationship between the elastic stress and the reversible deformation can be described by the Hooke's law; the evolution of the additional stress $\Delta\sigma_{ij}$, i.e. the

Fig. 1 Local tensile stresses and rock dilatancy



inelastic stress which exists around the cracks and belongs to the stress at the microscopic scale, may be described using the following Maxwell model

$$\frac{d\Delta\sigma_{ij}}{dt} = k\rho c_s^2 \dot{\epsilon}_{ij} - v \frac{\Delta\sigma_{ij}}{l}, \quad (13)$$

where $\Delta\sigma_{ij}$ are the additional stress components caused by the crack with characteristic scale l ; $\dot{\epsilon}_{ij}$ is the corresponding strain-rate component; ρ is the density of the medium; v is the relaxation velocity, which can be interpreted as the propagation velocity of single or multiple cracks, depending on the loading conditions; c_s is the propagation velocity of the elastic shear wave, k is the stress concentration coefficient around the crack. To simplify the analysis, all the additional stresses are assumed to relax with the same relaxation time $\tau = l/v$. The first term on the right-hand side of Eq. (13) describes the elastic loading, while the second term on the right-hand side of Eq. (13) represents the stress relaxation.

For the constant strain rate, the solution of Eq. (13) is

$$\Delta\sigma_{ij} = k\rho c_s^2 \dot{\epsilon}_{ij} \frac{l}{v} \left(1 - e^{-vt/l}\right). \quad (14)$$

In this paper, during the loading process of brittle rock, the occurrence of dilatancy is attributed to be the occurrence of fracture or even macroscopic fracture, whereas for the occurrence of fracture, especially the macroscopic fracture, it is necessary that the loading time t be greater than the relaxation time, i.e. $t > \tau$ (Qi et al. 2014, 2016; Huang et al. 2016). This is consistent with the loading condition of the triaxial compression test, which always takes long time before the final failure of the specimen. Therefore, Eq. (14) can be rewritten into the following form

$$\Delta\sigma_{ij} = k\rho c_s^2 \dot{\epsilon}_{ij} \frac{l}{v}. \quad (15)$$

Equation (15) is appropriate for the study of macroscopic fracture of rock sample (Qi et al. 2014, 2016; Huang et al. 2016). The additional stress in the direction perpendicular to the axial direction of the cylindrical rock sample can be expressed as

$$\Delta\sigma_3 = k\rho c_s^2 \dot{\epsilon}_3 \frac{l}{v}, \quad (16)$$

where $\Delta\sigma_3$ represents the additional stress in the same or opposite direction to σ_3 ; considering that $\epsilon_v = 2\epsilon_3 + \epsilon_1$, and during the subcritical crack growth stage, the strain rate in the radial and circumferential directions can be approximately expressed as $\dot{\epsilon}_3 \approx \epsilon_3/t_{B0}$, $\dot{\epsilon}_1 \approx \epsilon_1/t_{B0}$ (Yang et al. 2012) and, therefore, Eq. (16) can be rewritten as

$$\Delta\sigma_3 \approx \frac{1}{3} k\rho c_s^2 \frac{l}{v} \frac{(\epsilon_3 - \epsilon_1)}{t_{B0}} = \frac{1}{3} k\rho c_s^2 (\epsilon_3 - \epsilon_1) \frac{l}{l_c - l_0}, \quad (17)$$

where t_{B0} is the time taken by the subcritical crack growth stage, l is the length of crack at any time in the subcritical growth stage, l_c is the length of crack at the end of subcritical crack growth stage, and l_0 is the length of crack at the beginning of subcritical crack growth stage. Besides, during the subcritical crack growth stage, the rock specimen is still a continuum and the stress–strain relationship has the following form

$$\begin{aligned} \epsilon_3 - \epsilon_1 &= \frac{1 + \nu}{E_s} (\sigma_3 - \sigma_1) = \frac{3(1 + \nu)(\sigma_3 - \sigma_0)}{E_s} \\ &= \frac{3(1 + \nu)}{E_s} s_3; \end{aligned} \quad (18)$$

therefore, the additional stress $\Delta\sigma_3$ can be expressed as

$$\Delta\sigma = k\rho c_s^2 \frac{l}{l_c - l_0} \frac{(1 + \nu)}{E_s} s_3, \quad (19)$$

where s_3 is the applied deviatoric stress, $\sigma_0 = (\sigma_1 + 2\sigma_3)/3$ is the hydrostatic component of the applied stress; ν and E_s are, respectively, Poisson's ratio and Young's modulus of the rock in the subcritical crack growth stage.

During the dilatancy process of the rock sample, $\epsilon_3 - \epsilon_1 > 0$ (here we adopt the convention that compression strains are negative), according to Eq. (17), $\Delta\sigma_3 > 0$, which means the additional stress $\Delta\sigma_3$, is tensile. Except for the applied deviatoric stress, s_3 , the other parameters, such as k , ρ , c_s , l , l_0 , l_c , ν , and E_s in Eq. (19), are all determined by the rock properties and loading conditions and it is difficult for us to know their exact values.

The growth and expansion of the cracks are simultaneously determined by both the additional tensile stress $\Delta\sigma_3$ at the microscopic scale and the applied external loading at the macroscopic scale. Therefore, let us introduce the concept of effective tensile stress T_{loc} , which means the tensile stress around the crack at the microscopic scale will directly cause the growth and expansion of the crack. Considering the expression of Eq. (19), and the fact that the general stress state can be divided into the hydrostatic stress component, σ_0 , and the deviatoric stress component, $s_i = \sigma_i - \sigma_0$, ($i = 1, 2, 3$), then, we have the following expression

$$T_{loc} = \eta s_3 + \sigma_0, \quad (20)$$

where T_{loc} is the effective local tensile stress at the microscopic scale that will directly cause the growth and expansion of the cracks, $\eta = k\rho c_s^2 l(1+\nu)/(l_c - l_0)E_s$ (the expression of η shows that its value is determined by both the rock properties and the loading conditions).

Equation (20) means that T_{loc} is proportional to the value of s_3 , i.e. the effective local tensile stress is proportional to the component of the applied deviatoric stress which is normal to the crack surface; this is consistent with the concept proposed by Costin (1983).

Based on Eq. (20), the model I stress intensity factor, K_I , for the penny-shaped micro-cracks is obtained as follows

$$K_I = \frac{2}{\pi} T_{loc} \sqrt{\pi l}. \quad (21)$$

Considering that $s_3 = \Theta\sqrt{J_2}$, $\Theta = \frac{2}{\sqrt{3}}\sin(\theta_\sigma - \frac{2}{3}\pi)$, $J_2 = \frac{1}{2}s_{ij}s_{ij}$ is the second invariant of the deviatoric stress tensor, and θ_σ is the Lode angle, then Eq. (20) can also be written as

$$T_{loc} = \kappa\sqrt{J_2} + \sigma_0, \quad (22)$$

where $\kappa = \eta\Theta$. When $T_{loc} > 0$, the micro-cracks start to open and dilatancy occurs. According to the study of Tan (Tan et al. 1989), the rock dilatancy occurs when $\sqrt{J_2} > f_3$, which means that when $T_{loc} = 0$, there is $\sqrt{J_2} = f_3$; this relation can be substituted into Eq. (22) to obtain

$$T_{loc} = \kappa(\sqrt{J_2} - f_3). \quad (23)$$

Crack opening

As mentioned above, the rock dilatancy is attributed to the growth and expansion of cracks; therefore, we need to know the increased volume that was caused by cracks. To obtain the void volume caused by the opening of cracks in the medium, in addition to the aforementioned idealized assumptions, the following assumptions are also required:

1. No new cracks emerge in the loading process, i.e. the number of cracks N_0 in a unit volume is invariant;

2. The volume of the medium matrix is unchanged in the process of crack opening;
3. The behaviour of a set of cracks resembles a single crack in an overall average sense.

As shown in Fig. 2, micro-crack openings are caused by effective local tensile stresses. According to Seaman et al. (1976), when disregarding the effects of the interaction between neighbouring cracks, the penny-shaped cracks are presumed to be opened elastically as

$$\delta_0 = \frac{2}{\pi} \frac{1-\nu}{\mu_m} R T_{loc} = \frac{1-\nu}{\pi} \frac{1-\nu}{\mu_m} l T_{loc}, \quad (24)$$

where δ_0 is half the maximum separation gap between the crack faces, μ_m is the shear modulus of the matrix material, and $R = l/2$ is the radius of the crack. It is assumed that an ellipsoidal surface will generate as the crack opening. Then, the void volume V'_t caused by the opening of a single crack is

$$V'_t = \frac{4}{3} \pi R^2 \delta_0 = \frac{8(1-\nu)}{3\mu_m} R^3 T_{loc} = \frac{(1-\nu)}{3\mu_m} l^3 T_{loc}. \quad (25)$$

The materials' deformation caused by the opening of cracks with a characteristic length l is

$$V_t = N_0 V_0 V'_t = N_0 V_0 \frac{(1-\nu)}{3\mu_m} l^3 T_{loc}. \quad (26)$$

According to Eqs. (2) and (26), the damage factor caused by the opening of numerous cracks is

$$\alpha = \frac{V_0 + V_t}{V_0/\alpha_0} = \alpha_0 \left(1 + N_0 \kappa \frac{(1-\nu)}{3\mu_m} l^3 (\sqrt{J_2} - f_3) \right) \quad (27)$$

where N_0 is the number of initial cracks with characteristic length l in a unit volume, V_0 is initial volume of the medium element, and $N_0 V_0$ is the number of initial cracks in the medium element.

Substituting Eq. (27) into Eq. (11), the dilatancy during the subcritical crack growth stage is

$$\varepsilon_v = \frac{V - V_0}{V_0} = \frac{p\alpha_0}{K_m} + N_0 \kappa \frac{(1-\nu)f_3}{3\mu_m} l^3 \left(\frac{\sqrt{J_2} - f_3}{f_3} \right), \quad (28)$$

where

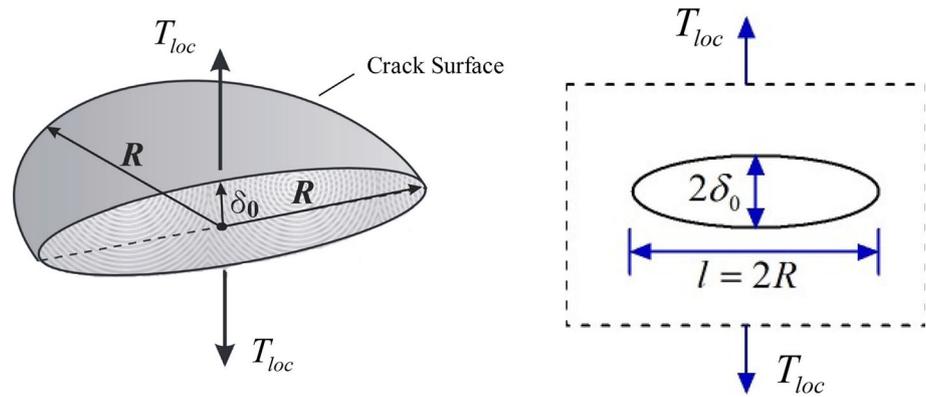
$$\varepsilon_{vd} = \varepsilon_v - \varepsilon_{ve} = N_0 \kappa \frac{(1-\nu)f_3}{3\mu_m} l^3 (\sqrt{J_2} - f_3/f_3)$$

is the volumetric increment due to the opening of cracks, and $\varepsilon_{ve} = p\alpha_0/K_m$ is the elastic strain of the specimen.

Crack length

Theoretical considerations have pointed out that the subcritical growth of cracks emerges when $K_I \geq K_{I0}$ (K_{I0} is the subcritical lower limit, according to Atkinson (1987); K_{I0}

Fig. 2 Crack opening under local tensile stresses



generally lies between $0.2 K_{IC}$ and $0.5 K_{IC}$; besides, the crack growth can be described by the following exponential expression (Wiederhorn et al. 1982)

$$v = v_0 \cdot \exp [\beta(K_I - K_{I0})], \quad (29)$$

where $v_0 \sim 10^{-9}$ m/s and $\beta \sim 10/(\text{MPa m}^{1/2})$ are constant material parameters determined experimentally (Atkinson 1987). Substituting the relation $K_{I0} = aK_{IC}$ ($a = 0.2\text{--}0.5$) and Eq. (21) into Eq. (29), we have

$$v = v_0 \cdot \exp \left[2\kappa(1-a)\beta \sqrt{\frac{l}{\pi}} (\sqrt{J_2} - f_3) \right], \quad (30)$$

where $l_0 \leq l \leq l_c$. The statistical results show that the initial and critical lengths of cracks are influenced by factors such as the loading conditions and rock properties, but, in general, $l_0 < l_c < 3l_0$ (Klein and Reuschle 2003, 2004). Therefore, according to Eq. (30), in the subcritical crack growth stage, the obvious effect on the crack growth velocity is coming from the loading conditions, i.e. mainly determined by expression $\sqrt{J_2} - f_3$, whereas the effect of the crack length's change on the velocity is not obvious. Without loss of generality, Eq. (30) can be approximated by

$$v = v_0 \cdot \exp \left[2\kappa(1-\alpha)\beta \sqrt{\frac{(l_0 + l_c)}{2\pi}} (\sqrt{J_2} - f_3) \right], \quad (31)$$

and the crack length at the end of the subcritical crack growth stage is

$$\begin{aligned} l_c &= \int_0^{t_{B0}} v dt \\ &= v_0 t_{B0} \exp \left[2\kappa(1-\alpha)\beta f_3 \sqrt{\frac{(l_0 + l_c)}{2\pi}} \left(\frac{\sqrt{J_2} - f_3}{f_3} \right) \right] \\ &= A \exp \left[B \left(\frac{\sqrt{J_2} - f_3}{f_3} \right) \right], \end{aligned} \quad (32)$$

where $A = v_0 t_{B0}$ and $B = 2\kappa(1-\alpha)\beta f_3 \sqrt{(l_0 + l_c)}/2\pi A$ are also regarded as the parameters related to the loading

conditions and the properties of the given material, t_{B0} is the time taken by the subcritical crack growth stage, as mentioned above. It is difficult to know the exact values of A , B , and η , but that does not matter, and our following analysis will show that they are just auxiliary parameters and would not be included in the finally dilatancy equation.

By substituting Eq. (32) into Eq. (28), the rock dilatancy at the end of the subcritical crack growth stage is

$$\varepsilon_v = \frac{p\alpha_0}{K_m} + \frac{D^* f_3}{\mu_m} \left[\exp \left(B \cdot \left(\frac{\sqrt{J_2} - f_3}{f_3} \right) \right) \right]^3 \left(\frac{\sqrt{J_2} - f_3}{f_3} \right), \quad (33)$$

where $D^* = \kappa \frac{1-\nu}{3} N_0 A^3$ is the material's dilatancy coefficient (Tan et al. 1989) that can be determined through linear fit, as shown in the test part of this paper. Based on the parameters given in the above sections, we can find that the value of B is very small. Besides, as mentioned above, in the subcritical crack growth stage, the crack growth velocity is mainly affected by the loading conditions; therefore, Eq. (33) is also appropriate to describe the rock dilatancy during whole the subcritical crack growth stage corresponding to the given loading conditions. To make it more convenient for use, we will simplify Eq. (33) into the following form

$$\begin{aligned} \varepsilon_v &= \frac{p\alpha_0}{K_m} + \frac{D^* f_3}{\mu_m} \left[\exp \left(B \cdot \left(\frac{\sqrt{J_2} - f_3}{f_3} \right) \right) \right]^3 \left(\frac{\sqrt{J_2} - f_3}{f_3} \right) \\ &\approx \frac{p\alpha_0}{K_m} + \frac{D^* f_3}{\mu_m} \left[1 + B \cdot \left(\frac{\sqrt{J_2} - f_3}{f_3} \right) \right]^3 \left(\frac{\sqrt{J_2} - f_3}{f_3} \right) \\ &\approx \frac{p\alpha_0}{K_m} + D^* \frac{f_3}{\mu_m} \left(\frac{\sqrt{J_2}}{f_3} - 1 \right)^m \approx \frac{p\alpha_0}{K_m} + D^* \frac{f_3}{\mu_m} \left(\frac{\sqrt{J_2}}{f_3} \right)^{m'}, \end{aligned} \quad (34)$$

where $B(\sqrt{J_2}/f_3 - 1) \ll 1$ stands for a weak-nonlinear term (for that, B is very small and $B \ll 1$, as mentioned above), and m' is a parameter that can be determined using conventional triaxial dilatancy experimental data.

Unstable crack growth stage

When $K_I \geq K_{IC}$, the rock mass transfers to the post-peak stage with unstable extension of cracks. The emergence of macro-cracks slices the rock mass into a discontinuous medium; then, under the effect of the external loading, granular fragments start to flow and swell due to the slip and rotation between blocks. From that instant, the calculation theory for continuum fails, and granular medium theory is needed. Hertel and Kerley (1998) suggested the following interpolation equation with experimental data to calculate rock damage factor

$$\alpha = \alpha_R - (\alpha_R - \alpha_F) \left(\frac{|\sigma_1 - \sigma_3| - \sigma_R^*}{\sigma_F^* - \sigma_R^*} \right)^\lambda, \tag{35}$$

where σ_F^* and σ_R^* are the peak strength and residual strength of the rock; α_F and α_R are the damage degrees corresponding to the rock reaches its peak and residual strengths, respectively. According to Eq. (11), the values of α_F and α_R can be obtained using $\alpha = \alpha_0(1 + \varepsilon_v - p\alpha_0/K_m)$, where $\alpha_0 \approx 1$, ε_v is the volume strain obtained from the triaxial compression test curve, and λ is a coefficient derived by fitting. The rock dilatancy in the unstable crack growth stage can be obtained by substituting Eq. (35) into Eq. (11)

$$\varepsilon_v = \frac{p\alpha_0}{K_m} + \frac{\alpha_R - \alpha_0}{\alpha_0} - \left(\frac{\alpha_R - \alpha_F}{\alpha_0} \right) \left(\frac{|\sigma_1 - \sigma_3| - \sigma_R^*}{\sigma_F^* - \sigma_R^*} \right)^\lambda. \tag{36}$$

Triaxial compression tests

The testing specimens with no visible flaws, $d = 50$ mm in diameter and $h = 100$ mm in height, were collected from Linyi, Shandong Province, and belong to a type of hard brittle red sandstone. The physical parameters for the specimen were $E = 18.24$ Gpa, $\rho_0 = 2450$ kg/m³, $\nu = 0.20$, and $\alpha_0 \approx 1$. The constant loading rate was 5×10^{-5} s⁻¹, which is consistent with the recommendation of ISRM. The lateral confining pressures are listed in Table 1 and the corresponding stress–strain curves of the tests are depicted in Fig. 4.

Table 1 Characteristic strength of the red sandstone under different lateral confining pressures

| Confining pressure [MPa] | 0 | 4 | 10 | 15 | 20 |
|--------------------------------|--------|--------|--------|--------|--------|
| Long term strength f_3^* | 53.77 | 67.76 | 74.16 | 89.25 | 106.58 |
| Peak stress σ_F^* | 105.42 | 126.43 | 143.37 | 183.47 | 201.58 |
| Residual strength σ_R^* | – | 31.05 | 45.82 | 80.03 | 97.07 |

Characteristic strength

We adopt the convention that compression strains are negative. About the dilatancy process, Tan et al. (1989) have given the following description: (1) at the beginning, due the closure of the pre-existing cracks, the volumetric strain is negative; (2) after $|\sigma_1 - \sigma_3|$ surpasses f_3^* ($f_3^* = \sqrt{3}f_3$), the positive volumetric dilatancy emerges; (3) after $|\sigma_1 - \sigma_3|$ reaches its maximum σ_F^* ($\sigma_F^* = \sqrt{3}\sigma_F$), the rock undergoes strain softening until it reaches the residual strength of σ_R^* ($\sigma_R^* = \sqrt{3}\sigma_R$). In our tests, due to the rigidity of the test machine and the brittleness of the rock, the stress–strain curve of the rock’s post-peak stage is hard to directly obtain; therefore, the value of σ_R^* can only be roughly evaluated by experience (for the brittleness of the red sandstone, it is observed that, after peak, the stress–strain curve drops rapidly from the peak value to the residual value; this characteristic is the basis of the residual value evaluation). The testing characteristic strengths of the rock samples are listed in Table 1.

Analysis of dilatancy

By performing a natural log operation on both sides of Eq. (34), we obtain:

$$\begin{aligned} \ln \varepsilon_{vd} &= \ln \left(\varepsilon_v - \frac{p\alpha_0}{K_m} \right) = \ln \left(\varepsilon_v - \frac{(\sigma_1 + 2\sigma_3)\alpha_0}{3K_m} \right) \\ &= \ln D^* \frac{f_3}{\mu_m} + m' \ln \left(\frac{\sqrt{J_2}}{f_3} \right), \end{aligned} \tag{37}$$

besides, according to Eq. (11), we also have

$$\ln \varepsilon_{vd} = \ln \left(\frac{\alpha}{\alpha_0} - 1 \right). \tag{38}$$

In Eq. (37), ε_v can be directly obtained from the test curve, σ_1 and σ_3 are, respectively, the maximum and minimum principal stress components of the specimen, $K_m = E/3(1-2\nu)$ is the bulk modulus, and the initial damage degree α_0 is assumed to be 1.

As shown in Fig. 3, the dots represent the testing data and the solid lines are the fitting curves. The curves of $\ln(\sqrt{J_2}/f_3)$ and $\ln \varepsilon_{vd}$ are nearly parallel, with slopes that are the values of parameter m' and intercepts showing the

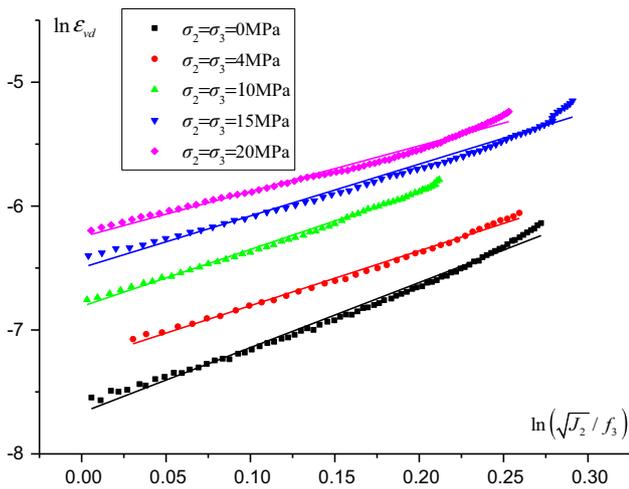


Fig. 3 Plots of $\ln(\sqrt{J_2}/f_3)$ against $\ln(\epsilon_{vd})$

influences of f_3 , which coincide well with the law proposed by Tan et al. (1989). According to the linear regression, the average value of m' is $\bar{m}' = 4.426$; by substituting $\bar{m}' = 4.426$ into Eq. (37), we have $\bar{D}^* = 0.573$.

The dilatancy equation in the post-peak softening stage can take the form of Eq. (36), in which α_F can be gained through Eqs. (37) and (38) by taking $\sqrt{J_2} = \sigma_F$. Besides, through experimental data, $\alpha_R = 1.017$ and $\lambda = 0.5$. The testing and calculated dilatancy curves of the triaxial compression tests are plotted in Fig. 4.

As shown in Fig. 4, considering that the differences between the specimens' real physical parameters and the testing parameters are inevitable, and to make the analysis simpler, the assumptions and the model for the cracks are idealized. Therefore, when ignoring these factors, the curves plotted based on the dilatancy equation presented in this paper are acceptable and can well describe the deformation trend of a brittle rock under triaxial compression tests.

Conclusions

In this paper, the idealized assumptions are made to the cracks in the rock samples, the p - α model is introduced to describe the damage degree, and the distributed fracture mechanics is employed as the basis of attributing a single crack's behaviour to all cracks. Based on our analyses, the following conclusions are drawn:

- During the compression process, the growth and expansion of micro-cracks are caused by the effective local tensile stresses around them; the effective local tensile stresses are directly related to the materials' physical parameters, the cracks' length and the external

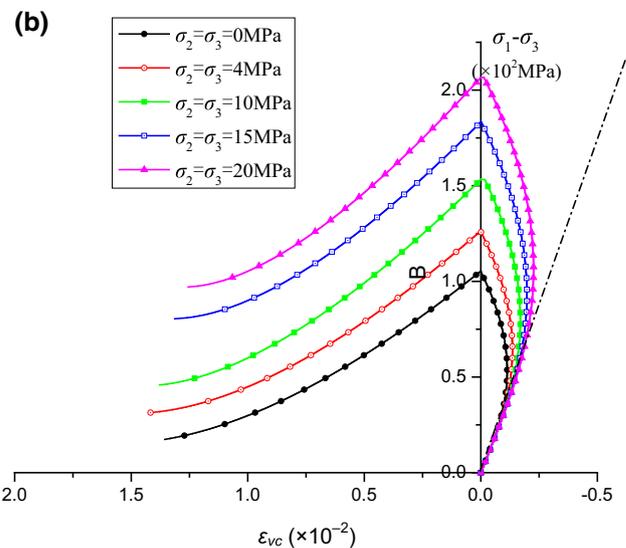
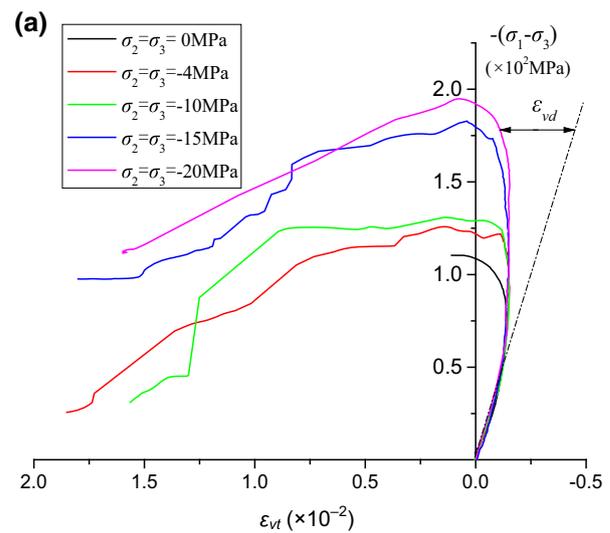


Fig. 4 Experimental (ϵ_{vt}) and calculated (ϵ_{vc}) dilatancy curves of triaxial compression tests with red sandstone

loadings; the complicated relationship among these factors can be described using the Maxwell model.

- According to the newly established dilatancy equation, when considering that most of the physical parameters in the equation are determined by the material itself, then we can find that the most important external factor that can affect the dilatancy is the strength of the shear stress $\sqrt{J_2}$, or, in other words, the most important relation that can affect the dilatancy is the one between $\sqrt{J_2}$ and the material's shear dilatancy strength f_3 ; dilatancy occurs when $\sqrt{J_2} > f_3$.
- Considering that the differences between the rock specimens' real physical parameters and the testing parameters are inevitable and that the crack model adopted in this paper is idealized, the triaxial compression tests which are

carried out with the red sandstone demonstrate that the dilatancy equation presented in this paper describes the whole deformation process of brittle rocks under triaxial compression test well.

- The method of simplifying the complicated expressions which consist of parameters determined by the rock properties and loading conditions as parameters that can be obtained through test may be useful in the study of rock-like materials and rock engineering.
- It is necessary to note that to know the exact values of the parameters k , ρ , c_s , l , l_0 , l_c , v , and E_s , especially the value of the crack length, is very difficult and almost impossible. But, in fact, it is not necessary to know their exact values, for they are only auxiliary parameters which are helpful to finally establish the dilatancy Eq. (34).
- Linear Eq. (37) is transformed from dilatancy Eq. (34); we can obtain D^* and m' in Eq. (37) by fitting without knowing the exact values of k , ρ , c_s , l , l_0 , l_c , v , and E_s , as shown in Fig. 3.

Acknowledgements The authors would like to acknowledge the financial support from the National Basic Research Program of China (973 Program, No. 2013CB036005) and the National Natural Science Foundation of China (Grant No. 51527810, No. 51679249, No. 51309233), which is greatly appreciated by the authors.

References

- Ashby MF, Sammis CG (1990) The damage mechanics of brittle solids in compression. *Pure appl Geophys* 133(3):489–521. doi:10.1007/BF00878002
- Atkinson BK (1987) *Fracture mechanics of rock*. Academic Press, London
- Bhatt JJ, Carroll MM, Schatz JF (1975) A spherical model calculation for volumetric response of porous rocks. *J Appl Mech* 42(2):363–368. doi:10.1115/1.3423582
- Bosman JD, Malan DF, Drescher K (2000) Time-dependent tunnel deformation at Hartbeestfontein Mine. *J S Afr Inst Min Metall* 100(6):333–340
- Brace WF, Paulding BW, Scholz C (1966) Dilatancy in the fracture of crystalline rocks. *J Geophys Res* 71:3939–3953. doi:10.1029/JZ071i016p03939
- Bridgman PW (1949) Volume changes in the plastic stages of simple compression. *J Appl Phys* 20:1241–1251. doi:10.1063/1.1698316
- Carroll MM, Holt AC (1972) Static and dynamic pore collapse relations for ductile porous materials. *J Appl Phys* 43(4):1626–1635. doi:10.1063/1.1661372
- Cogan J (1976) Triaxial creep tests of Ophongong limestone and Ophir shale. *Int J Rock Mech Min Sci Geomech Abstr* 13(1):1–10. doi:10.1016/0148-9062(76)90221-7
- Cook NGW (1970) An experiment proving that dilatancy is a pervasive volumetric property of brittle rock loaded to failure. *Rock Mech* 2(4):181–188. doi:10.1007/BF01245573
- Costin LS (1983) A microcrack model for the deformation and failure of brittle rock. *J Geophys Res* 88(B11):9485–9492. doi:10.1029/JB088iB11p09485
- Cristescu ND (2002) New trends in rock mechanics. *Int Appl Mech* 38(1):1–22. doi:10.1023/A:1015364607665
- Das S, Scholz CH (1981) Theory of time-dependent rupture in the Earth. *J Geophys Res* 86(B7):6039–6051. doi:10.1029/JB086iB07p06039
- Diederichs MS, Kaiser PK, Eberhardt E (2004) Damage initiation and propagation in hard rock during tunnelling and the influence of near-face stress rotation. *Int J Rock Mech Min Sci* 41(5):785–812. doi:10.1016/j.ijrmms.2004.02.003
- Haimson B, Chang C (2000) A new true triaxial cell for testing mechanical properties of rock, and its use to determine rock strength and deformability of westerly granite. *Int J Rock Mech Min Sci* 37(99):285–296. doi:10.1016/S1365-1609(99)00106-9
- Han LJ, He YN, Zhang HQ (2016) Study of rock splitting failure based on Griffith strength theory. *Int J Rock Mech Min Sci* 83:116–121. doi:10.1016/j.ijrmms.2015.12.011
- Hertel ES, Kerley GI (1998) CTH reference manual: The equation of state package, Sandia National Laboratories, SAND98-0947
- Huang H-X, Fan P-X, Li J, Wang M-Y, Rong X-L (2016) A theoretical explanation for rock core disk in triaxial unloading test by considering local tensile stress. *Acta Geophys* 64(5):1430–1445. doi:10.1515/acgeo-2016-0068
- Kachanov LM (1986) *Introduction to the continuum damage mechanics*. Springer-Science+Business Media, Dordrecht
- Klein E, Reuschle T (2003) A model for the mechanical behaviour of Bentheim sandstone in the brittle regime. *Pure Appl Geophys* 160:833–849. doi:10.1007/978-3-0348-8083-1_3
- Klein E, Reuschle T (2004) A pore crack model for the mechanical behaviour of porous granular rocks in the brittle deformation regime. *Int J Rock Mech Min Sci* 41(6):975–986. doi:10.1016/j.ijrmms.2004.03.003
- Martin CD, Read RS (1992) Strength of massive granite around underground excavation. In: *Proc. 16th Canadian Rock Mechanics Symposium*, Sudbury, pp 1–11
- Paterson MS, Wong TF (2005) *Experimental rock deformation—the brittle field*. Springer, Berlin
- Qi C-Z, Wang M-Y, Bai J-P, Li K-R (2014) Mechanism underlying dynamic size effect on rock mass strength. *Int J Impact Eng* 68:1–7. doi:10.1016/j.ijimpeng.2014.01.005
- Qi C-Z, Wang M-Y, Bai J-P, Wei X-K, Wang H-S (2016) Investigation into size and strain rate effects on the strength of rock-like materials. *Int J Rock Mech Min Sci* 86:132–140. doi:10.1016/j.ijrmms.2016.04.008
- Qian Q-H, Zhou X-P (2011) Non-euclidean continuum model of the zonal disintegration of surrounding rocks around a deep circular tunnel in a non-hydrostatic pressure state. *J Min Sci* 47(1):37–46. doi:10.1134/S1062739147010059
- Read RS (2004) 20 years of excavation response studies at AECL's underground research laboratory. *Int J Rock Mech Min Sci* 41(8):1251–1275. doi:10.1016/j.ijrmms.2004.09.012
- Scholz CH (1968) Microfracturing and the inelastic deformation of rock in compression. *J Geophys Res* 73(4):1417–1432. doi:10.1029/JB073i004p01417
- Scholz CH, Sykes LR, Aggarwal YP (1973) Earthquake prediction: a physical basis. *Science* 181(4102):803–810. doi:10.1126/science.181.4102.803
- Seaman L, Curran DR, Shockey DA (1976) Computational models for ductile and brittle fracture. *J Appl Phys* 47(11):4814–4826. doi:10.1063/1.322523
- Szczepanik Z, Milne D, Kostakis K, Eberhardt E (2003) Long term laboratory strength tests in hard rock. 10th ISRM Congress “Technology Roadmap for Rock Mechanics”. Sandton, South Africa, pp 1179–1184
- Tan TK (1982) The mechanical problems for the long-term stability of underground galleries. *Chin J Rock Mech Eng* 1(1):1–20

- Tan TK, Kang WF (1983) Time dependent dilatancy prior to rock failure and earthquakes. In: Proc. 5th ISRM Congress, Melbourne
- Tan T-K, Shi Z-Q, Yu Z-H, Wu X-Y (1989) Dilatancy, creep and relaxation of brittle rocks measured with the 8000 kN multipurpose triaxial apparatus. *Phys Earth Planet Inter* 55(3–4):335–352. doi:[10.1016/0031-9201\(89\)90081-2](https://doi.org/10.1016/0031-9201(89)90081-2)
- Tapponnier P, Brace WF, Tapponnier P, Brace WF (1976) Development of stress-induced microcracks in westerly granite. *Int J Rock Mech Min Sci Geomech Abstr* 13:103–112. doi:[10.1016/0148-9062\(76\)91937-9](https://doi.org/10.1016/0148-9062(76)91937-9)
- Vásárhelyi B, Bobet A (2000) Modeling of crack initiation, propagation and coalescence in uniaxial compression. *Rock Mech Rock Eng* 33(2):119–139. doi:[10.1007/s006030050038](https://doi.org/10.1007/s006030050038)
- Wiederhorn SM, Freiman SW, Fuller Jr ER, Simmons CJ (1982) Effects of water and other dielectrics on crack growth. *J Mater Sci* 17:3460–3478. doi:[10.1007/BF00752191](https://doi.org/10.1007/BF00752191)
- Yang S-Q, Jing H-W, Wang S-Y (2012) Experimental investigation on the strength, deformability, failure behavior and acoustic emission locations of red sandstone under triaxial compression. *Rock Mech Rock Eng* 45:583–606. doi:[10.1007/s00603-011-0208-8](https://doi.org/10.1007/s00603-011-0208-8)
- Zhang Y, Shao J-F, Xu W-Y, Zhao H-B, Wang W (2015) Experimental and numerical investigations on strength and deformation behavior of cataclastic sandstone. *Rock Mech Rock Eng* 48(3):1083–1096. doi:[10.1007/s00603-014-0623-8](https://doi.org/10.1007/s00603-014-0623-8)

Changes of snow cover in Poland

Małgorzata Szwed¹ · Iwona Pińskwar¹ · Zbigniew W. Kundzewicz^{1,2} ·
Dariusz Graczyk¹ · Abdelkader Mezghani³

Received: 29 December 2016 / Accepted: 16 January 2017 / Published online: 1 February 2017
© The Author(s) 2017. This article is published with open access at Springerlink.com

Abstract The present paper examines variability of characteristics of snow cover (snow cover depth, number of days with snow cover and dates of beginning and end of snow cover) in Poland. The study makes use of a set of 43 long time series of observation records from the stations in Poland, from 1952 to 2013. To describe temporal changes in snow cover characteristics, the intervals of 1952–1990 and of 1991–2013 are compared and trends in analysed data are sought (e.g., using the Mann–Kendall test). Observed behaviour of time series of snow-related variables is complex and not easy to interpret, for instance because of the location of the research area in the zone of transitional moderate climate, where strong variability of climate events is one of the main attributes. A statistical link between the North Atlantic Oscillation (NAO) index and the snow cover depth, as well as the number of snow cover days is found.

Keywords Snow cover depth · Snowy season · Variability · NAO index · Poland

Introduction

Unequivocal warming of the global climate system is now evident from observations of increases in air temperatures at various spatial scales, from local to national, regional,

continental, hemispheric, and global (Kundzewicz and Huang 2010; Stanisławska et al. 2013; IPCC 2013). Almost the entire globe has experienced surface warming. Many of the observed changes have been unprecedented over the time scales of decades to millennia. The globally averaged combined land and ocean surface temperature data show a warming of 0.85 (0.65–1.06) °C, over the period 1880–2012, when multiple independently produced data-sets exist (IPCC 2013). Each of the sixteen years of the 21st century, 2001–2016, belongs to a set of seventeen globally warmest years on record. Global temperature changes are accompanied, expectedly, by changes in other environmental elements, e.g., the sea level or the state of the cryosphere.

The air temperature is only one of several climate elements and factors influencing snow cover. Other factors include: atmospheric circulation, precipitation phase and amount, wind speed, solar radiation, sunshine, etc. Snow cover plays a key role in environmental and socio-economic systems. It is regarded as a crucial component of the global climatic system in relation to the role in modifying the transfer of energy and water vapour between the surface of the Earth and the atmosphere (Brown and Goodison 1996). From the climatological point of view, snow cover functions as a thermal insulation between the atmosphere and the ground, thus lowering cold extremes of temperature of the ground surface. Its occurrence is of essential importance for water resources (via snow water contents). By acting as a seasonal reservoir of water, snow likewise strongly alters the hydrological cycle. Earlier snow melting in winter or spring causes ground to dry earlier, increasing the likelihood of drought in summer (Jaagus 1997). Snowmelt plays an important role in hydrological (and river) regime. The hydrological regime is closely related to seasonal changes in climate. In Poland, rainfall amount and

✉ Małgorzata Szwed
mszwed@man.poznan.pl

¹ Institute for Agricultural and Forest Environment, Polish Academy of Sciences, Poznań, Poland

² Potsdam Institute for Climate Impact Research, Potsdam, Germany

³ Met Norway, Oslo, Norway

evaporation are controlling factors of the hydrological regime in a warmer part of the year, while precipitation (snowfall and rainfall) and air temperature are the essential factors in a colder part of the year. Hence, the dynamics of the cryosphere in the snowmelt-fed river catchments are of crucial importance for water resources management. Snow cover may also influence many human activities, e.g., agriculture, energy production or transport. Thus, snow cover decides on hibernating conditions of winter crops, protecting them from frost damage during the winter months. Hydro-power and energy supply are largely dependent on, and sensitive to, changes in snow amount and duration. Heavy snowfall (and, as a result, deep snow cover) tends to cause traffic hazards and communication problems. Finally, snow cover is important in sports and recreation (skiing, sledging, walking), aesthetics, and tradition.

It seems to be of much interest to examine long time series of national observation records, wherever available, to look for changes. Snow cover in Poland, similarly to other countries located in central European lowlands, is dominated by strong inter-winter and intra-winter variability. Year-to-year variability of snow cover reflects well the overall winter climatic variability. The variability throughout the season may be represented by cold periods alternating with the mild ones and by the short- or long-lasting snow cover (Bednorz 2013; Falarz 2004). Other research studies also report large variability between individual seasons, and the lack of distinct, statistically significant trends in snow cover depth (Brown and Petkova 2007; Czarnecka 2012, etc.). Several authors, such as Falarz (2004), Nowosad and Bartoszek (2007) and Czarnecka (2012) spotted a slight decreasing trend of snow cover characteristics (number of days with snow cover, depth of snow cover, occurrence frequency) in Poland during the second half of the 20th century, but no change was distinguished for a longer period. Some authors noted a tendency of gathering of snow-poor winters (e.g., Jaagus 1997; Niedźwiecki 1998) and reported about winters with particularly abundant snow cover (e.g., 1970, 1979, 2010) and outstandingly snowless (e.g., 1949, 1930, 2008), cf. Hejduk and Hejduk (2014). In general, the Northern Hemisphere spring snow cover has decreased in extent (IPCC 2013).

There have been several recent publications indicating a North Atlantic Oscillation (NAO) track in snow cover records. However, these studies analyze the connections of the NAO with different specific characteristics of the snow cover, e.g., the number of days with snow cover (e.g., Bednorz 2002), the depth of snow cover or the date of disappearance of the snow (Hori and Yasunari 2003). The study areas of these research works are different, like a part of Poland (e.g., Bednorz 2002), the whole of Poland (Czarnecka 2011) or a macro-region (e.g., western Eurasian continent in Hori and Yasunari 2003).

In general, the strong coupling of the NAO and snow extent during winter (especially in January and February) was confined to the area of Western Europe (Hori and Yasunari 2003). This connection becomes less significant towards the East (Gutzler and Rosen 1992; Clark et al. 1999; Bednorz 2004). The strong negative correlation for the number of days with snow cover in Poland was found (Bednorz 2002). The strongest connection of the NAO index vs depth of snow cover in Poland exists in January and in February (Czarnecka 2012). The NAO index can also explain the variability of dates of disappearance of the snow cover (Czarnecka 2011). During warmer winters, in the positive NAO phase, a large variability of the melt supply and earlier culminations of flows were noted by Wrzeński (2010). This indicates decreasing number of days with snow cover.

The present paper examines the set of long time-series of snow cover characteristics in Poland based on daily values of snow cover depth at 43 meteorological stations in 1952–2013. The snow cover characteristics includes: snow cover depth, number (derived) of days with snow cover during winter (also in adjacent seasons: autumn and spring), and the dates of beginning and end of snow cover.

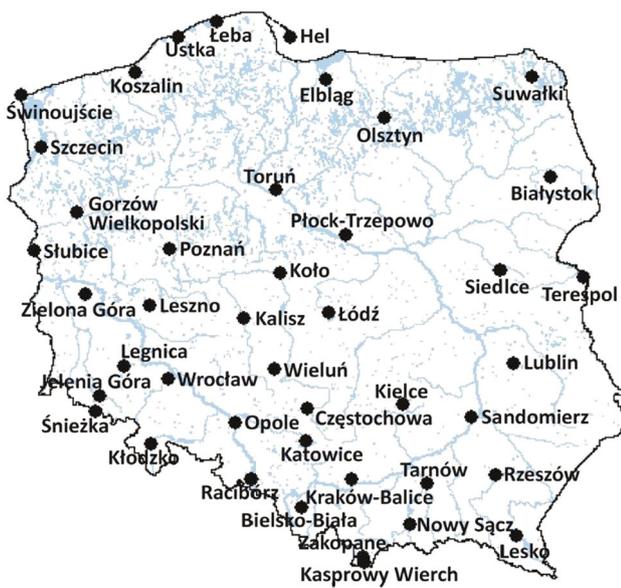
Data and methodology

Although instrumental temperature measurements in Polish lands, like in other European countries, started around the end of the 18th century, due to historical reasons (armed conflicts, long lasting lack of independent state, and major changes of national boundaries before 1945), records of good-quality daily meteorological data in Poland are shorter than in many European countries with more stable and less convoluted history. Polish Meteorological and Hydrological Service (the Institute of Meteorology and Water Management-National Research Institute, Polish acronym: IMGW-PIB) provided a time series of depth of snow cover records observed at meteorological stations in Poland. The basic criteria for selection of stations were: (1) the length of the available time series, and (2) the spatial distribution of stations with the goal of covering the whole territory of Poland and its climate regions, in a possibly uniform way.

In this study, daily depth of snow cover from 43 stations in Poland was analysed. The studied data started on 1 October 1951 (38 meteorological stations) or later (5 stations) and continued until 31 December 2013. The shorter series of observation data concern stations in: Rzeszów (data start in 1952), Nowy Sącz and Lesko (from 1954), Terespol (from 1955) and Leszno (the shortest time series, starting in 1958). The meteorological stations used in the present study are listed in Table 1 and mapped in Fig. 1.

Table 1 List of meteorological stations used in this work

| Station | Latitude | Longitude | Elevation [m. a.s.l.] | Station | Latitude | Longitude | Elevation [m. a.s.l.] |
|------------------------------|----------|-----------|-----------------------|---------------------|----------|-----------|-----------------------|
| Białystok | 53°06′ | 23°10′ | 148 | Olsztyn | 53°46′ | 20°25′ | 133 |
| Bielsko-Biała-Aleksandrowice | 49°48′ | 19°00′ | 398 | Opole | 50°40′ | 17°58′ | 176 |
| Częstochowa | 50°49′ | 19°06′ | 295 | Płock-Trzepowo | 52°35′ | 19°44′ | 106 |
| Elbląg-Milejewo | 54°10′ | 19°26′ | 38 | Poznań | 52°25′ | 16°50′ | 86 |
| Gorzów Wielkopolski | 52°45′ | 15°17′ | 72 | Racibórz-Studzienna | 50°05′ | 18°13′ | 190 |
| Hel | 54°36′ | 18°49′ | 1 | Rzeszów-Jasionka | 50°06′ | 22°03′ | 200 |
| Jelenia Góra | 50°54′ | 15°48′ | 342 | Sandomierz | 50°42′ | 21°43′ | 217 |
| Kalisz | 51°44′ | 18°05′ | 140 | Siedlce | 52°11′ | 22°16′ | 146 |
| Kasprowy Wierch | 49°14′ | 19°59′ | 1991 | Słubice | 52°21′ | 14°36′ | 21 |
| Katowice | 50°29′ | 19°05′ | 317 | Suwałki | 54°08′ | 22°57′ | 184 |
| Kielce-Suków | 50°51′ | 20°37′ | 268 | Szczecin-Dąbie | 53°24′ | 14°37′ | 1 |
| Kłodzko | 50°26′ | 16°39′ | 316 | Śnieżka | 50°44′ | 15°44′ | 1603 |
| Koło | 52°12′ | 18°40′ | 116 | Świnoujście | 53°55′ | 14°14′ | 6 |
| Koszalin | 54°12′ | 16°09′ | 33 | Tarnów | 50°02′ | 20°59′ | 209 |
| Kraków-Balice | 50°05′ | 19°48′ | 237 | Terespol | 52°04′ | 23°37′ | 133 |
| Legnica | 51°13′ | 16°10′ | 122 | Toruń | 53°03′ | 18°35′ | 69 |
| Lesko | 49°28′ | 22°20′ | 386 | Ustka | 54°35′ | 16°52′ | 6 |
| Leszno-Strzyżewice | 51°50′ | 16°32′ | 91 | Wieluń | 51°13′ | 18°35′ | 195 |
| Lublin-Radawiec | 51°13′ | 22°24′ | 238 | Wrocław | 51°06′ | 16°53′ | 120 |
| Łeba | 54°45′ | 17°32′ | 2 | Zakopane | 49°18′ | 19°57′ | 857 |
| Łódź-Lublinek | 51°44′ | 18°24′ | 187 | Zielona Góra | 51°56′ | 15°30′ | 180 |
| Nowy Sącz | 49°37′ | 20°42′ | 292 | | | | |

**Fig. 1** Map of locations of meteorological stations used in this work

Based on time series of daily depth of snow cover records the following characteristics of snow cover were estimated:

- mean value of snow cover depth for every month of the cold half-year (from October to April) in cm;
- mean value of snow cover depth for winter (December–January–February) in cm;
- maximum value of snow cover depth for winter (December–January–February) in cm;
- number of days with snow cover during every winter (DJF) and mean number of days with snow cover during DJF;
- number of days with snow cover during snow season (from the beginning of October to the end of May);
- first day with snow cover during every winter and mean date of appearance of snow cover in winter season;
- last day with snow cover during winter and mean date of disappearance of snow cover in winter season;
- the Pearson correlation coefficient between the NAO index and analysed snow cover characteristics (snow cover depth and the number of days with snow cover).

In this paper, the link of the NAO index value with snow cover depth, as well as with the number of days with snow cover is analysed based on mean value of the NAO index for DJF and corresponding values of snow cover depth and the number of days with snow cover in DJF. This research

uses the North Atlantic Oscillation Index defined first by Barnston and Livezey (1987), worked up by Hurrell (1995). The NAO consists of a north–south dipole of anomalies, with one center located over Greenland and the other center of opposite sign spanning the central latitudes of the North Atlantic between 35° and 40°N. Data of the monthly NAO index for the 65-year period of 1950–2015 originate from the Climate Prediction Center (available at <http://www.cpc.noaa.gov>).

To describe temporal changes in snow cover characteristics, for every station:

- the intervals of 1952–1990 and of 1991–2013 were compared;
- the Mann–Kendall statistic was used;
- the rate of change per year for the whole analysed period was calculated.

To detect existing changes/trends in time series of snow cover data, the Hydrospect 2.0 software was used (Radziejewski and Kundzewicz 2000). Statistical significance of trend for every station at various levels was determined. The Hydrospect 2.0 software package allows changes in long time-series of hydrological records to be detected. Among the statistical tests available in the package the Mann–Kendall test was used in the research reported in this paper. Additionally, simple linear regression was used, that allows the user to find the slope of the linear trend and the significance level, if the records are independent and normally distributed (else, normalization is offered under the normal-scores regression test).

Apart from trend detection in the complete data records covering 1952–2013, differences between values of indices in two compared intervals, 1952–1990 and 1991–2013, were illustrated. There is a disproportion in the length between analysed periods (38 years vs 23 years). Moreover, after preliminary analysis, it turned out that characteristic of observed data changed considerably after 1990. Therefore, for the purposes of the present paper, the authors decided on such a division of time series.

To describe spatial variability of snow cover characteristics, IDRISI GIS platform was used. For spatial interpolation of the snow-related characteristics, the Kriging method was used. Authors are aware that ordinary Kriging may underestimate high values and overestimate low values in case when the purpose of interpolation is to delineate areas that exceed a threshold level of the predicted surface values, but it seemed to be an optimal predictor in this paper.

Strength of the relationship between the value of the NAO index and the depth of snow cover, and the number of days with snow cover was expressed by the R^2 value (the square of the Pearson correlation coefficient) of polynomial and linear regression, respectively.

Results

Snow cover depth

Mean snow cover during DJF

The mean depth of snow cover in winter (DJF) in 1952–2013 in Poland varies from 2.2 to 11.8 cm (except for mountain areas, where the depth of snow cover is considerably higher). The lowest mean depth of snow cover is characteristic for the western part of lowland Poland, while the highest—for the north-eastern part (Suwałki District)—Fig. 2.

Comparing the intervals 1952–1990 and 1991–2013, one can draw a conclusion that the former one was more snowy (in terms of mean depth of snow cover) in 38 out of 43 analysed stations (for spatial variability see Fig. 3a, b).

Figure 4 illustrates the time series of observed changes in mean depth of snow cover [cm] for the Toruń station for DJF, as an example. No significant trend on the 95% level was detected for most of the country, using statistic of Mann–Kendall. In general, after 1980, a cluster of years with small cover occurred (including a whole 20-year interval, 1988–2007, with mean winter snow cover depth below 5 cm). However, more snowy winters (with higher depth of mean snow cover) have also appeared more recently.

Spatial variability of temporal changes in mean depth of snow cover in winter (DJF) in 1952–2013 is presented in Fig. 5. It is presented as a slope of linear regression (rate of change per unit time–year), measured in mm/year. There is no distinct spatial pattern in this parameter; the value of change per year is positive in some locations, while negative in others. There are many more locations with

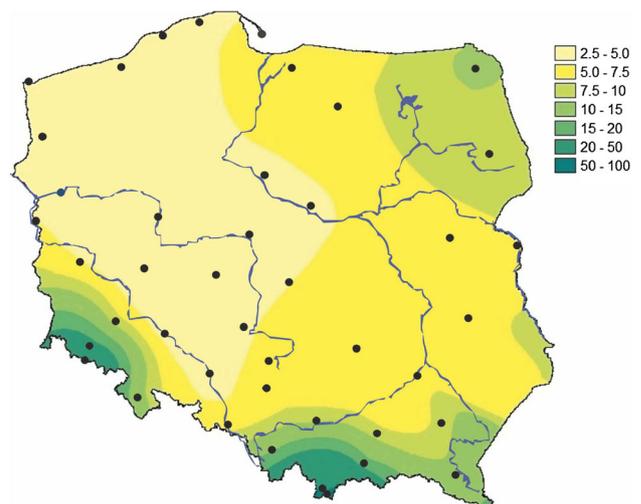


Fig. 2 The mean depth of snow cover [cm] in winter (DJF) in the period of 1952–2013 in Poland

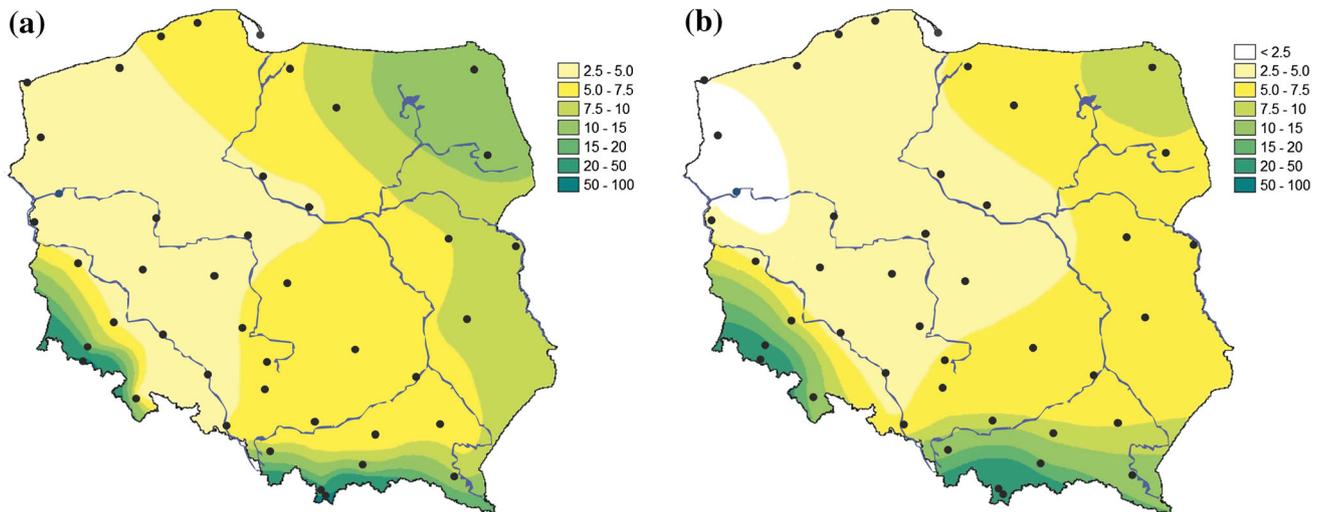


Fig. 3 The mean depth of snow cover [cm] in winter (DJF) in Poland in the intervals of (a) 1952–1990 and (b) 1991–2013

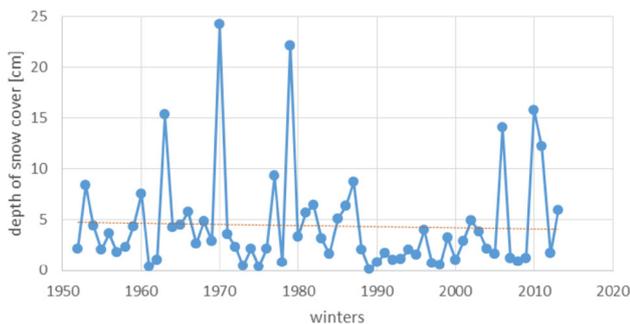


Fig. 4 Time series of mean depth of snow cover [cm] in DJF in Toruń (1952–2013). Red line—linear regression of time series

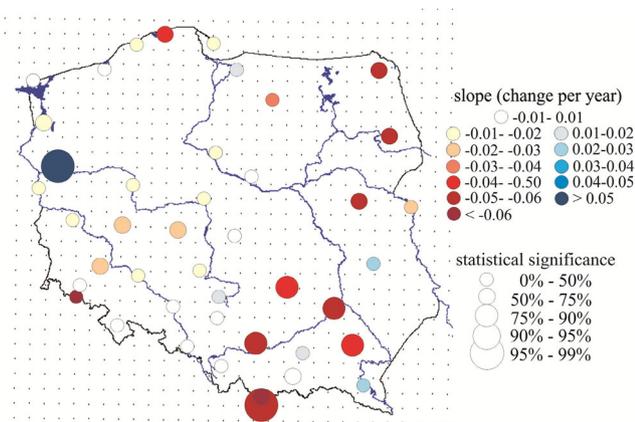


Fig. 5 Temporal changes in mean depth of snow cover in winter (DJF) in 1952–2013 expressed by the rate of change per year and the statistical significance. The values of change per year and statistical significance have been established with the help of the Hydrospect package, based on linear regression

negative values of slope than with positive values of slope. However, it seems that in the East, there are larger decreases of the mean depth of snow cover in the analysed

period. Nevertheless, great majority of these changes are not statistically significant at the level of 95% (Fig. 5).

NAO index vs mean depth of snow cover

The NAO index appears to be negatively correlated with the mean snow cover depth. So, for a positive value of the NAO index, there is a fairly high probability that the mean depth of snow cover would be below the multi-year mean. In contrast, for a negative value of the NAO index, there is a fairly high probability of snow cover depth exceeding the mean. This relationship gets weaker to the East of Poland. Figure 6 presents examples of different strength of relationship between the value of the NAO index and the depth of snow cover (Szczecin—an example of a stronger relation with $R^2 = 0.43$ and Rzeszów—an example of a weaker relation, $R^2 = 0.25$). At the end of this section, Fig. 7 illustrates strength of the relationship between the value of the NAO index and the mean depth of snow cover expressed by the R^2 value of non-linear (quadratic) regression for the entire area of Poland. In this figure, one can clearly note that the connection of the mean depth of snow cover with the NAO index decreases towards the East. Significance level of correlation coefficient for all the stations (excluding mountain ones) is on the level of 99%. For the mountain stations (Kasprowy Wierch, Zakopane, Śnieżka) a relation between the NAO index and mean depth of snow cover does not exist.

Maximum snow cover depth

The maximum depth of snow cover in winter (DJF) in 1952–2013 in Poland varies from 34 cm in Poznań and Słubice to 85 cm in Kraków (while in mountain areas

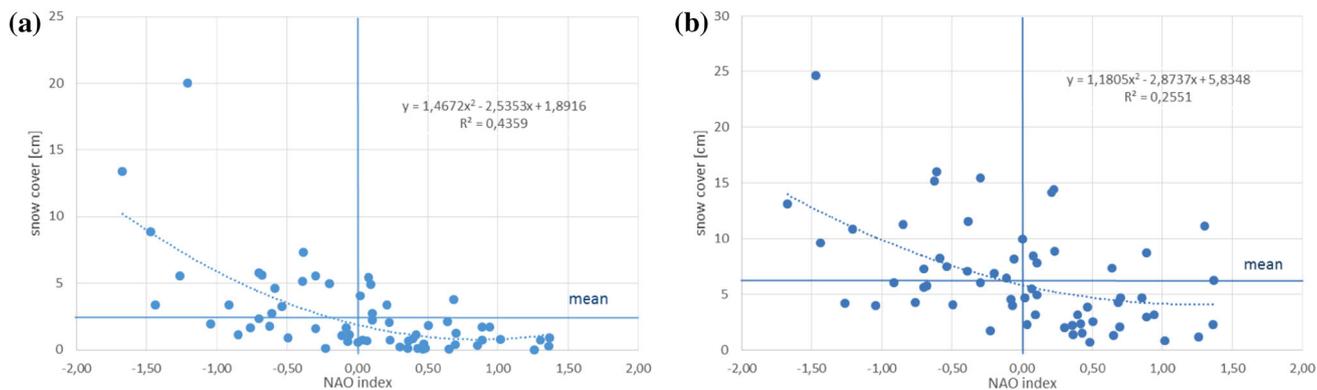


Fig. 6 NAO index vs depth of snow cover in DJF in (a) Szczecin (North–West Poland) and (b) Rzeszów (South–East Poland)

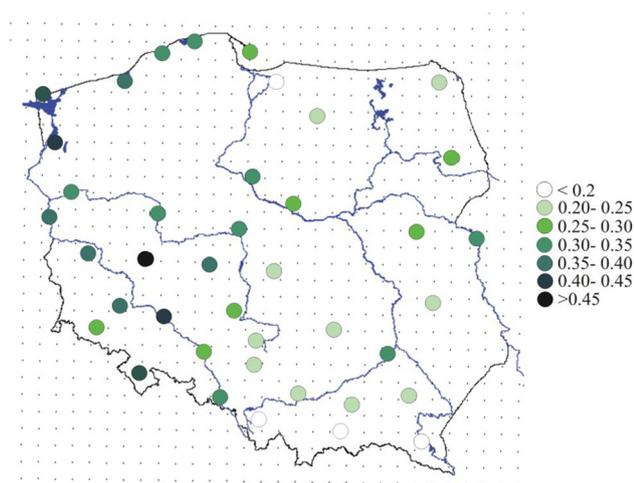


Fig. 7 Strength of the relationship between the value of the NAO index and the mean depth of snow cover in Poland in the period of 1952–2013 expressed by the R^2 value of non-linear regression

maximum depth of snow cover exceeds 300 cm). In 1952–2013, the pattern of spatial variability of the maximum depth of snow cover in DJF in Poland is similar to the one of the mean depth, i.e., the maximum depth of snow cover increases from the West to the East (North–East) cf. Fig. 8.

The maximum depth of snow cover in winter (DJF) recorded in the period of 1952–2013 is in most cases equal to the maximum value recorded in the period of 1952–1990. The values recorded later, in the period of 1990–2013, are definitely lower (Fig. 9). In the period of 1952–1990, the maximum value of snow cover depth was recorded for 39 stations, while for the period of 1991–2013 only for four stations. The years 1963 and 1979 are absolute record holders. In these years, maximum of snow cover depth occurred for 13 and 12 stations, respectively.

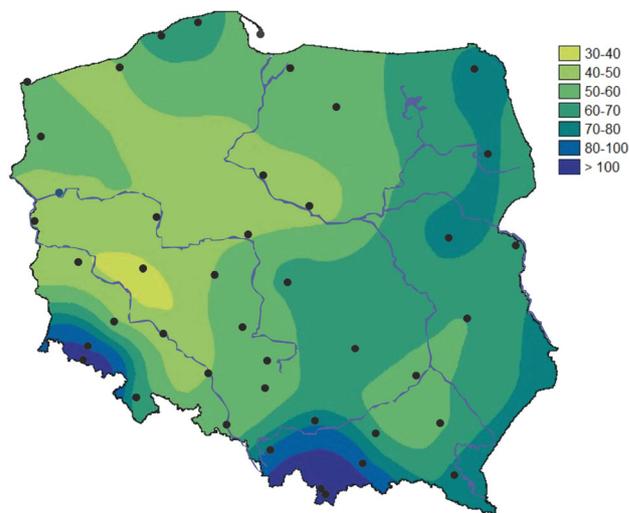


Fig. 8 Spatial distribution of the maximum depth of snow cover [cm] in winter (DJF) in 1952–2013 in Poland

Number of days with snow cover

The mean (over all winters in the interval 1952–2013) number of days with snow cover increases from the West to the East (North–East) and varies from 29.6 days in Szczecin to 65.8 in Suwałki (Fig. 10) with the mean value of 47.2 days. Even higher values are recorded in the mountains where snow cover remains over the whole winter, as a rule.

If the entire snowy period including the non-winter months (OCT, NOV, MAR, APR) is taken into consideration, the average number of days with the snow cover in Poland amounts to 60.6 days (mountain stations not included). It means that before and after winter, the snow cover persists for two additional weeks, and lower numbers of extra days are observed in the western and north-western parts of Poland while higher numbers are observed in the areas of upland in southern Poland and in North–

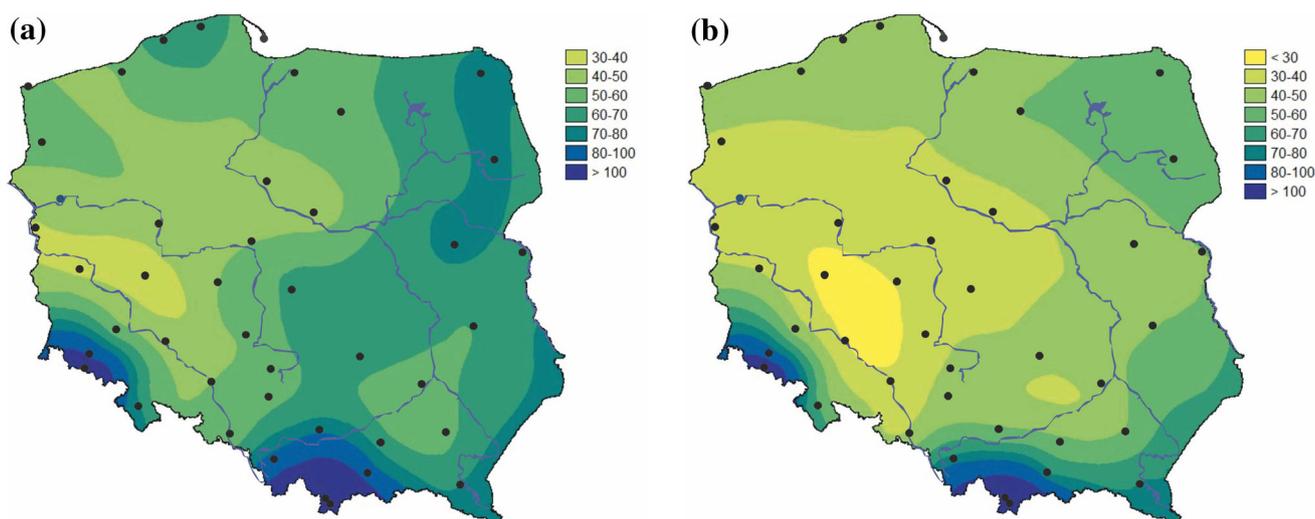


Fig. 9 The maximum depth of snow cover [cm] in winter (DJF) in Poland in the intervals of (a) 1952–1990 and (b) 1991–2013

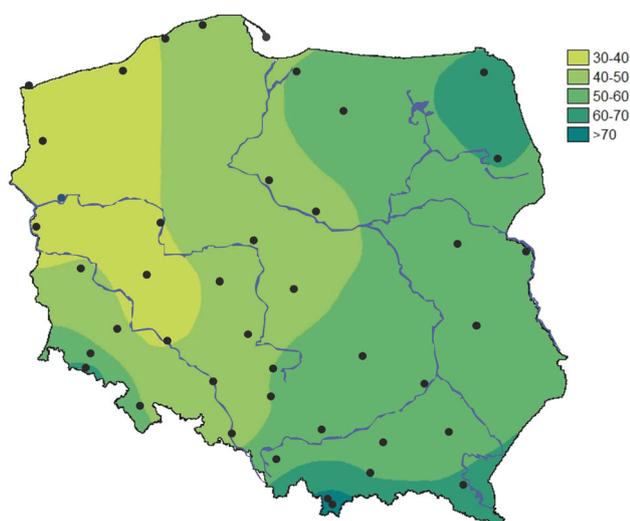


Fig. 10 Mean number of days with snow cover in DJF for 1952–2013

East. In the mountains, the number of days with snow cover is much higher, almost doubled as compared to DJF.

The mean number of days with snow cover in winter (DJF) decreases in time (Fig. 11), but no statistically significant trend has been observed in the number of days with snow cover in DJF using Mann–Kendall statistic. However, in most cases, linear regression goes down (cf. an example of Zielona Góra–Fig. 12). A quasi-periodicity (approx. 10 years) of maxima can be spotted, with maximum in DJF of 1970, but with strong area-dependence. The winter (DJF) of 1970 was the most snowy one in terms of the number of days with snow cover for 32 out of 43 stations. Even if weak decreases in the number of days with snow cover in DJF for 1952–2013 were noted for entire Poland, they are not significant on the 95% level for most of the

country. The lower changes are noted in the southern part of Poland (Fig. 13).

NAO vs mean number of days with snow cover

Observational evidence suggests that the NAO index is negatively correlated with the number of days with snow cover in Poland. The link is even stronger than for snow depth.

Thus, for a positive value of the NAO index, there is quite a high probability that the number of days with snow cover would be below the multi-year mean. Consequently, for a negative value of the NAO index, there is a fairly high probability that the number of days with snow cover would be higher than the mean. Figure 14 presents examples of different strength of relationship between the NAO and the number of days with snow cover (Szczecin—an example of a strong relation with $R^2 = 0.47$ and Rzeszów—an example of a weak relation, with $R^2 = 0.25$).

The spatial variability of the relationship between the NAO index and the number of days with snow cover in 1952–2013 for the entire area of Poland, expressed by the R^2 value of linear regression, is presented in Fig. 15. In general, this relationship is weaker in the southern upland part of Poland, while for the lowland areas it is stronger. Still, significance level of correlation coefficient for all the stations (excluding mountain ones) is on the level of 99%.

Start of the snowy season

In this section, strictly first day with appearance of snow cover during snow season is analysed.

Throughout Poland, the mean day of snow cover appearance in the winter season in the analysed period of

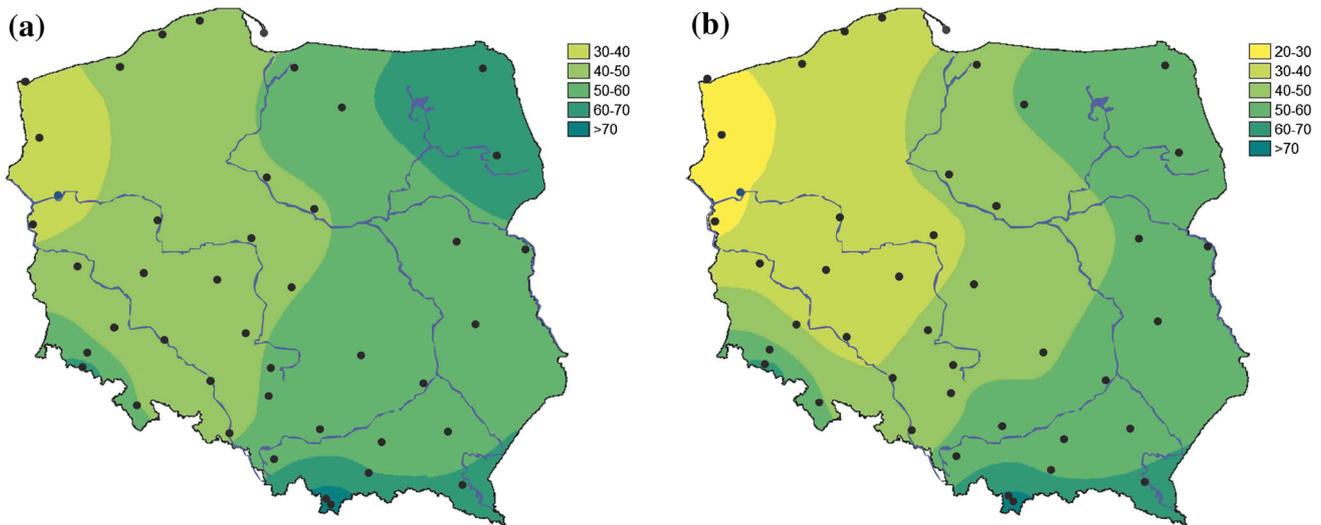


Fig. 11 Mean number of days with snow cover in winter (DJF) in Poland in the intervals of (a) 1952–1990 and (b) 1991–2013

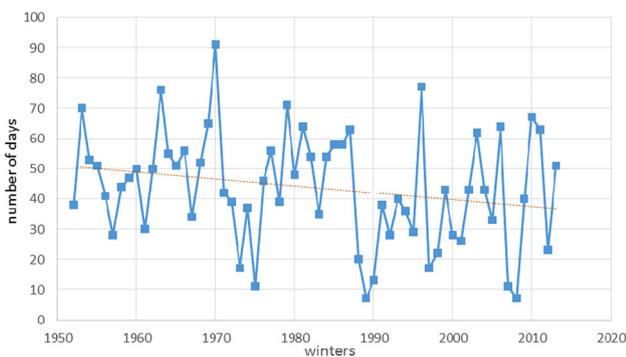


Fig. 12 Number of days with snow cover in DJF in Zielona Góra (1952–2013). *Red line*—linear regression of time series

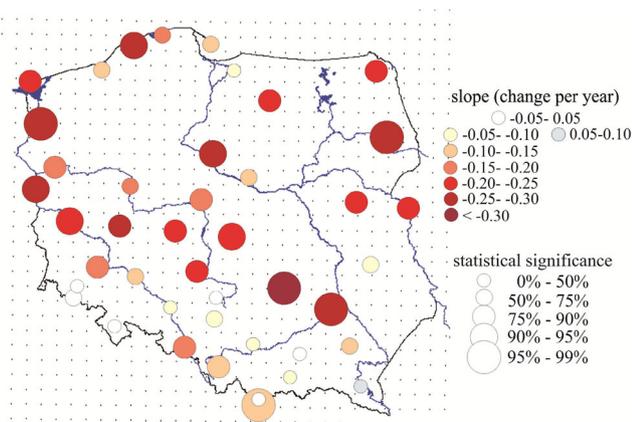


Fig. 13 Temporal changes in number of days with snow cover in winter (DJF) in 1952–2013 expressed by the rate of change per year and the statistical significance. The values of change per year and statistical significance have been established with the help of the Hydrospect package, based on linear regression

1952–2013 is 23 November. However, if the mountain stations, where the snow cover appears considerably earlier (in October), were not taken into consideration, the beginning of the snowy season would be 26 November. Excluding mountain areas, the snowy season, on average, starts between 14 November in Lesko and 10 December in Hel. In general, earlier appearance of snow cover is characteristic for the eastern part of Poland (in mid-November), in the western and north-western (coastal) parts snow cover appears later (first decade of December)—Fig. 16. When comparing the date of first appearance of snow cover for the intervals of 1952–1990 and 1991–2013, it seems that more recently snowy season started later (Fig. 17). However, no distinct tendency has been observed in the date of the first day of snow cover across Poland.

In mountain areas, snow cover has been recorded several times on 1 October, which is the first day of data of snow cover analysis in this research. Except for mountain areas, the earliest start of the snowy season took place on 3 October 1973 in Nowy Sącz, while the latest start of the snowy season was on 24 February 1988 in Szczecin.

Figure 18 illustrates time series of recorded date of the first day of snow cover on the example of Elbląg. No distinct tendency has been observed in the date of the first day of snow cover across Poland, except for Lesko, where the start of the snowy season takes place gradually earlier (on the level of confidence 95% using Mann–Kendall test). In general, the beginning of the snowy season depends on the region of the country. The appearance of snow cover is strongly place- and time-dependent and inter-annual variability is very strong.

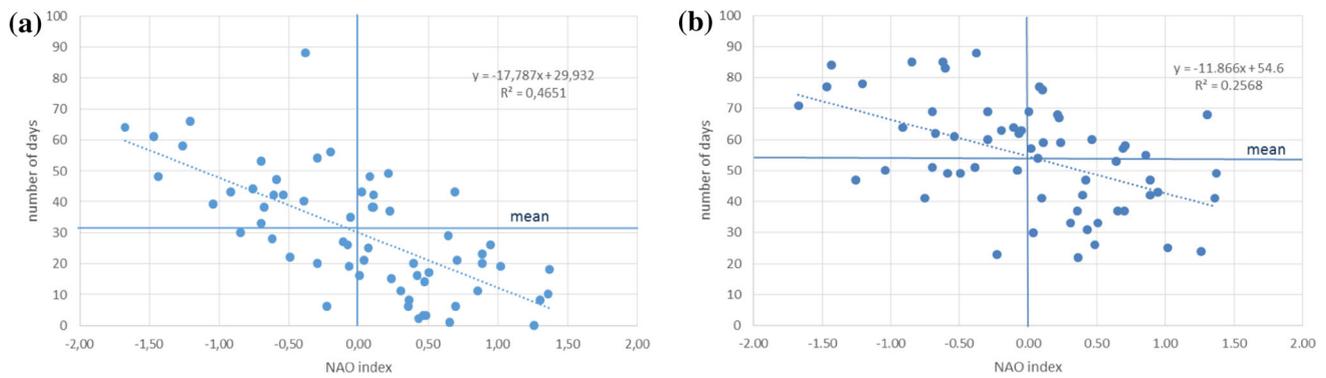


Fig. 14 NAO index vs number of days with snow cover in DJF in (a) Szczecin (North–West Poland) and (b) Rzeszów (South–East Poland)

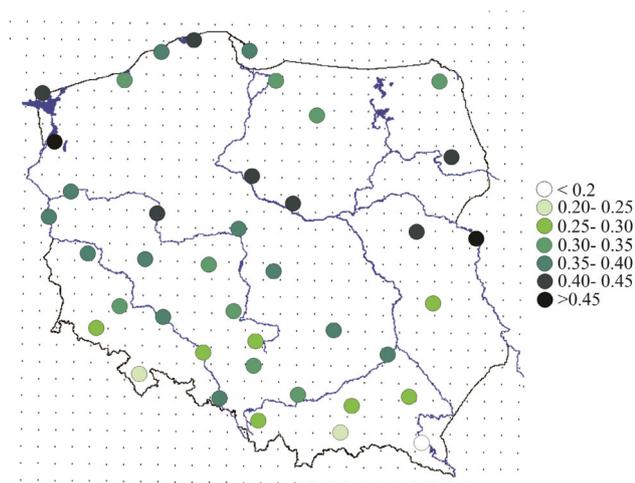


Fig. 15 Strength of the relationship between the NAO index and the number of days with snow cover in Poland in the period of 1953–2013, expressed by the R^2 value of linear regression

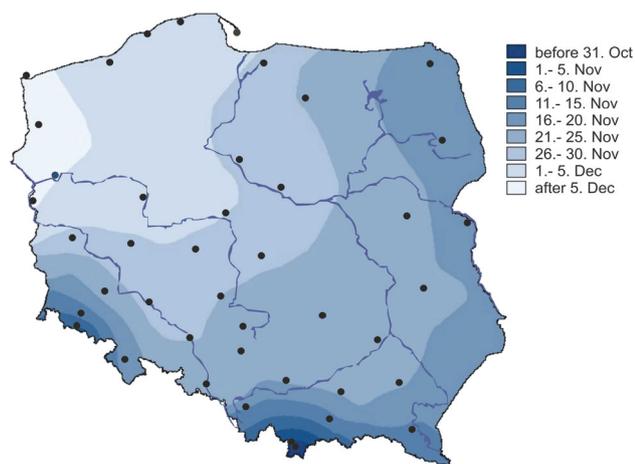


Fig. 16 The first day of snow cover in Poland; mean for all years in 1952–2013

End of the snowy season

In this section the absolutely last day with snow cover in every winter is analysed.

In the analyzed period of 1952–2013, except mountain areas, the snowy season ends, on average, between 11 March in Szczecin and Słubice, and 8 April in Bielsko Biała. In general, snow cover persists longest in the eastern and the southern part of Poland, while in the north–western parts it disappears early—see Fig. 19. The mean day of snow cover disappearance in Poland is 28 March, or 26 March, if mountain areas are not included.

The comparison of the dates of disappearance of snow cover for the periods of 1952–1990 and 1991–2013 is presented in Fig. 20. For the period of 1991–2013, snow cover stays shorter. The last day of snow cover is recorded several days earlier in some regions (especially in the North–East part of the country), however, no distinct tendency has been noted across Poland.

In mountain and foothill areas, snow cover may not disappear until 30 April. In the lowland part of Poland, the snow did not disappear until the end of April in Zielona Góra in 1985 and in Łódź in 1962. The earliest disappearance of snow cover took place on 22 November 1989 in Szczecin. This means that in some years there was no snow cover in Szczecin during the whole winter. In 1989, snow cover disappeared on 22 November and was absent throughout the winter. In contrast, in 1987/1988, there was no snow cover over nearly whole winter. The first snow cover day occurred towards the end of climatic winter—on 24 February 1988.

Figure 21 illustrates the time series of recorded last day of snow cover on the example of Białystok. No distinct tendency has been observed in the date of last days of snow cover across Poland. Similarly, to the first day of the snow cover in the winter season, the last day is strongly place- and time-dependent and inter-annual variability is very strong.

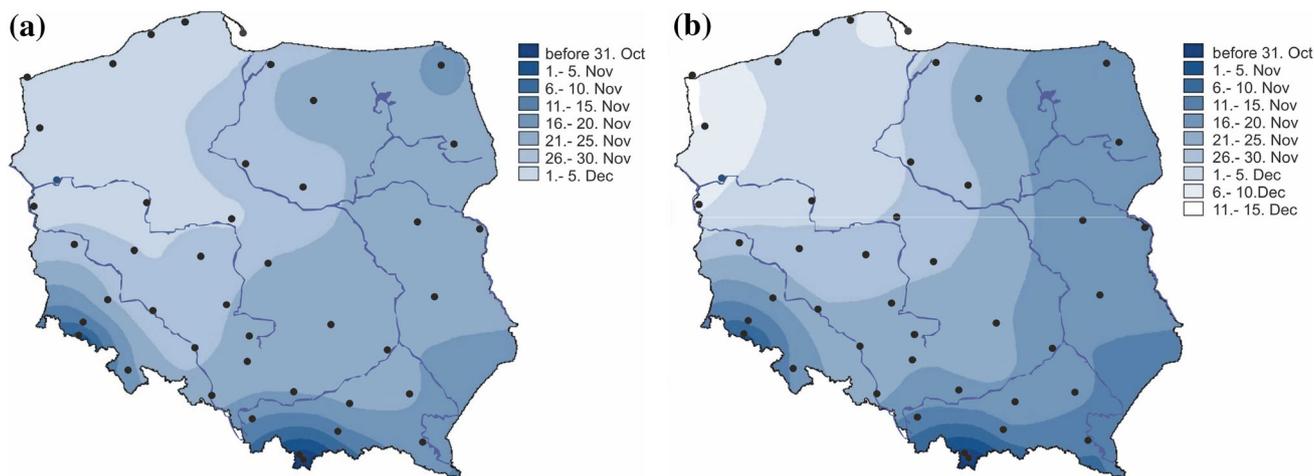


Fig. 17 The first day of snow cover in Poland in the intervals of (a) 1952–1990 and (b) 1991–2013

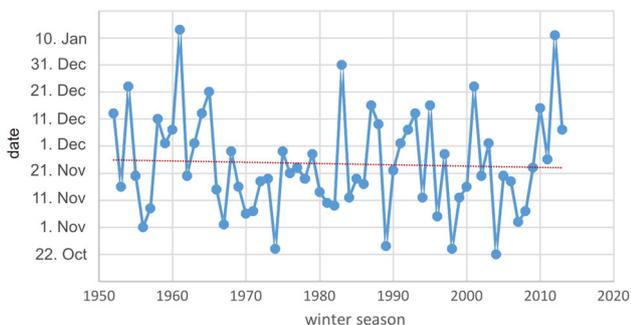


Fig. 18 Time series of recorded first day of snow cover in Elbląg. Red line—linear regression of time series

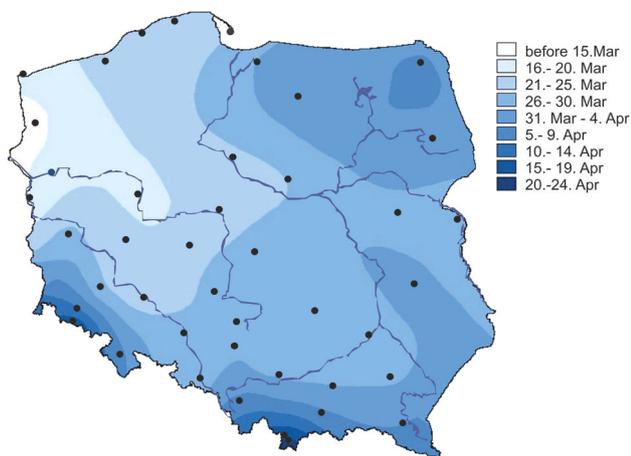


Fig. 19 The last day of snow cover in Poland; mean for all years in 1952–2013

Concluding remarks

The present paper examines the variability of characteristics of snow cover (snow cover depth, number of days with snow cover and dates of beginning and end of snow cover)

in Poland. Observed behaviour of time series of snow is complex and not easy to interpret. Even if shrinking snow cover is typically expected in the warming climate of the moderate climate zone, typically the change in Poland is dominated by strong inter-winter variability, rendering trend detection difficult. Our findings coincide with other research studies, which also report large variability between individual seasons e.g., Falarz (2004), Nowosad and Bartoszek (2007), Czarnecka (2012) and Hejduk and Hejduk (2014).

During typical, mean winter (DJF) in Poland (1952–2013) depth of snow cover varies from 2.2 to 11.8 cm (except for mountain areas, where it is considerably larger) and is rising to the North–East. The mean number of days with snow cover increases from the West to the East (North–East) and varies from 29.6 to 65.8 days per winter. Higher values are recorded in mountain stations. On average, the snow cover lasts from 26 November to 26 March.

It seems that the increase in temperature should delay first snow cover appearance and accelerate and increase snowmelt. However, for Poland, dates of appearance and disappearance of snow cover show slight changes—some delay in appearance and advance in disappearance. Still, no statistically significant trend has been detected. Czarnecka (2011) reached similar conclusions. Similarly, no distinct, statistically significant, trend has been observed in the mean snow cover depth in DJF. However, comparing the intervals 1952–1990 and 1991–2013, the former one was more snowy (in terms of mean depth of snow cover). Other research studies also report the lack of distinct, statistically significant trends in snow cover depth in Poland, e.g., Falarz (2004), Czarnecka (2012), etc. No distinct statistically significant trend has been observed in the number of days with snow cover in DJF as well. However, in most

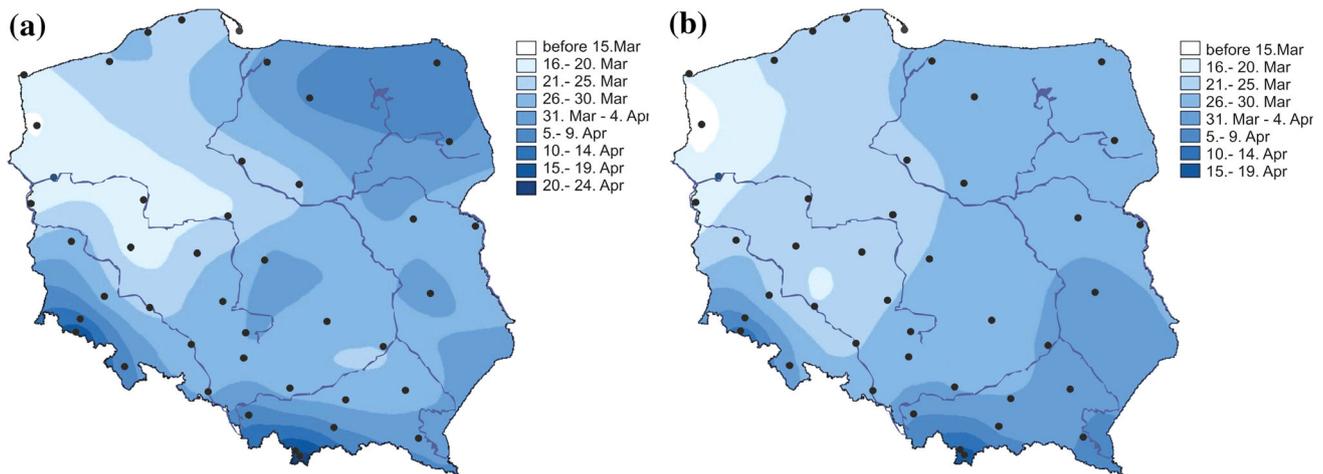


Fig. 20 The last day of snow cover in Poland in the intervals of (a) 1952–1990 and (b) 1991–2013

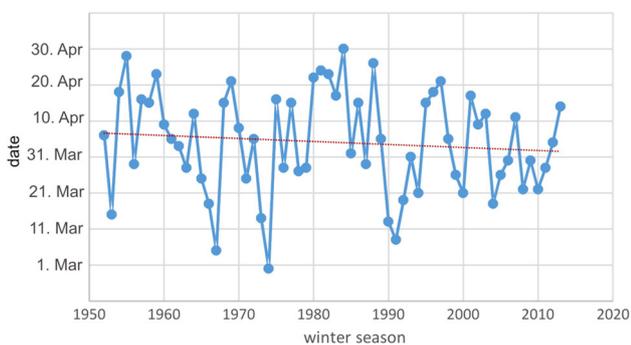


Fig. 21 Time series of recorded last day of snow cover in Białystok. Red line—linear regression of time series

cases, linear regression for time series of number of days with snow cover has a negative slope. The obtained results corroborate findings in recent papers of other researchers. Several authors, such as Falarz (2004), Nowosad and Bartoszek (2007) and Czarnecka (2012) spotted a slight decreasing trend (statistically insignificant) in the number of days with snow cover in Poland during the second half of the 20th century.

The pattern of air circulation in the North Atlantic region (described by the NAO index) has influence on characteristics of snow cover in Poland in winter. The positive phase of the NAO can be associated with the occurrence of less snowy winters, whereas the negative phase increases the probability of the occurrence of more snowy winters. A statistical link between the value of NAO index and the snow cover depth, as well as the number of snow cover days was found. The connection of snow cover depth and the NAO index decreases to the East, while connection of the number of days with the snow cover and the NAO index is almost equal for the lowland areas of Poland and much weaker in the upland areas on the South. For comparison, a strong negative correlation between the

NAO index and the number of days with snow cover in Poland was found by Bednorz (2002). Next, the strongest connection of the NAO index with the depth of snow cover in Poland was reported in Bednorz (2004), while Czarnecka (2012) reports it for January and February only.

Due to substantial importance of snow cover for thermal insulation, water availability and river regime, a continued analysis of snow cover characteristics seems to be still necessary for solving several practical problems. There are benefits and threats in many fields of the national economy, such as agriculture/farming, water management, transport and tourism related to snow cover.

Acknowledgements The reported research has been carried out within the CHASE-PL (Climate change impact assessment for selected sectors in Poland) project of the Polish–Norwegian Research Programme in the frame of Project Contract No. Pol-Nor/200799/90/2014. Observation data used in this work were provided by the Institute of Meteorology and Water Management-State Research Institute (IMGW-PIB).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Barnston AG, Livezey RE (1987) Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Mon Weather Rev* 115:1083–1126
- Bednorz E (2002) Snow cover in Western Poland and macro-scale circulation conditions. *Int J Climatol* 22:533–541
- Bednorz E (2004) Snow cover in Eastern Europe in relation to temperature, precipitation and circulation. *Int J Climatol* 24:591–601

- Bednorz E (2013) Heavy snow in Polish–German lowlands–Large-scale synoptic reasons and economic impacts. *Weather Clim Extremes* 2:1–6
- Brown RD, Goodison BE (1996) Interannual variability in reconstructed Canadian snow cover, 1915–1992. *J Clim* 9:1299–1318
- Brown RD, Petkova N (2007) Snow cover variability in Bulgarian mountainous regions, 1931–2000. *Int J Climatol* 27:1215–1229. doi:[10.1002/joc.1468](https://doi.org/10.1002/joc.1468)
- Clark MP, Serreze MC, Robinson DA (1999) Atmospheric controls on Eurasian snow extent. *Int J Climatol* 19:27–40
- Czarnecka M (2011) Zmienność terminów początku i końca pokrywy śnieżnej o różnym czasie zalegania i ich uwarunkowania cyrkulacyjne. *Prace i Studia Geograficzne, ZUT Szczecin* 47:109–118
- Czarnecka M (2012) Częstość występowania i grubość pokrywy śnieżnej w Polsce. *Acta Agrophys* 19(3):501–514
- Falarz M (2004) Variability and trends in the duration and depth of snow cover in Poland in the 20th Century. *Int J Climatol* 24:1713–1727
- Gutzler DS, Rosen RD (1992) Interannual variability of wintertime snow cover across the Northern Hemisphere. *J Clim* 5(12):1441–1447
- Hejduk A, Hejduk L (2014) Thermal and snow conditions of winters and winter floods on example of Zagożdżonka River. *Ann Warsaw Univ Life Sci-SGGW Land Reclam* 46(1):3–15
- Hori ME, Yasunari T (2003) NAO impact towards the springtime snow disappearance in the western Eurasian continent. *Geophys Res Lett* 30(19). doi:[10.1029/2003GL018103](https://doi.org/10.1029/2003GL018103)
- Hurrell JW (1995) Decadal trends in the North Atlantic Oscillation: regional temperatures and precipitation. *Science* 269:676–679
- IPCC (2013) Summary for Policymakers. In: Stocker TF, Qin D, Plattner GK, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds), *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, New York
- Jaagus J (1997) The impact of climate change on the snow cover pattern in Estonia. *Clim Chang* 36:65–77
- Kundzewicz ZW, Huang S (2010) Seasonal temperature extremes in Potsdam. *Acta Geophys* 58(6):1115–1133. doi:[10.2478/s11600-010-0026-5](https://doi.org/10.2478/s11600-010-0026-5)
- Niedźwiecki M (1998) Charakterystyka pokrywy śnieżnej w Łodzi w latach 1950–1989. *Acta Univ Lodziensis Folia Geograph Phys* 3:265–277
- Nowosad M, Bartoszek K (2007) Wieloletnia zmienność grubości pokrywy śnieżnej w okolicy Lublina. In: Piotrowicz K, Twardosz R (red) *Wahania klimatu w różnych skalach przestrzennych i czasowych*. Instytut Geografii i Gospodarki Przestrzennej UJ Kraków. pp 411–421
- Radziejewski M, Kundzewicz ZW (2000) HydroSpect-Software for detecting changes in hydrological data. In: Robson A, Kundzewicz ZW (eds) *Detecting trend and other changes in hydrological data*, App. 2. World Climate Programme-Applications and Services, World Meteorological Organization, Geneva
- Stanisławska K, Kundzewicz ZW, Krawiec K (2013) Hindcasting global temperature by evolutionary computation. *Acta Geophys* 61(3):732–751. doi:[10.2478/s11600-012-0091-z](https://doi.org/10.2478/s11600-012-0091-z)
- Wrzesiński D (2010) *Przestrzenne zróżnicowanie stabilności reżimu odpływu rzek europejskich*. Bogucki Wydawnictwo Naukowe, Poznań, p 220

Spatio-temporal assessments of rockburst hazard combining b values and seismic tomography

Jing Li^{1,2} · Si-Yuan Gong^{1,2}  · Jiang He^{1,2} · Wu Cai^{1,2} · Guang-An Zhu^{1,2} · Chang-Bin Wang^{1,2} · Tian Chen^{1,2}

Received: 5 January 2017 / Accepted: 16 January 2017 / Published online: 31 January 2017
© Institute of Geophysics, Polish Academy of Sciences & Polish Academy of Sciences 2017

Abstract A better understanding of rockburst precursors and high stress distribution characteristics can allow for higher extraction efficiency with reduced safety concerns. Taking the rockburst that occurred on 30 January 2015 in the Sanhejian Coal Mine, Jiangsu Province, China, as an example, the mechanism of rockburst development in a roadway was analysed, and a combined method involving b values and seismic velocity tomography was used to assess the rockburst in both time and space, respectively. The results indicate that before the rockburst, b values dropped significantly from 0.829 to 0.373. Moreover, a good agreement between a significant decrease in b values and the increase of the number of strong tremors was found. Using seismic tomography, two rockburst risk areas were determined where the maximum velocity, maximum velocity anomaly and maximum velocity gradient anomaly were 6 km/s, 0.14 and 0.13, respectively. The high-velocity regions corresponded well with the rockburst zone and large seismic event distributions. The combination of b values and seismic tomography is proven to have been a promising tool for use in evaluating rockburst risk during underground coal mining.

Keywords Coal mining · Rockburst · b value · Seismic velocity tomography · Strong tremors

Introduction

Underground coal mining activities lead to disturbances in the natural stress field in the surrounding coal–rock mass and subsequent stress releases through seismic events. Rockbursts are particular cases of mining-induced seismic events, usually of a magnitude ranging from M_L 1.5–4.5, which result in roadway destruction due to elastic strain energy being emitted in a sudden, rapid, and violent way from a coal–rock mass, and in the worst case, they can cause casualties (Young et al. 1992; He and Qian 2010; Dou et al. 2012). Many countries suffer rockburst hazards. For example, a rockburst occurred in Springhill coal mine in Nova Scotia, Canada claiming, 75 lives (Ge 2005). A rockburst accident induced by a large thrust fault caused ten deaths and trapped 75 people at Yima Qianqiu coal mine, Henan Province, China (Cai et al. 2014). As the average mining depth and intensity increases continuously, rockbursts pose a persistent threat to mine safety.

The b value, a frequency-magnitude relationship defined by Gutenberg and Richter (1944), is one of the basic seismological parameters used to reflect the characteristics of seismogenic structures and stress distributions in specific regions (Wiemer et al. 1998). The decreases in b values with time normally imply the occurrence of high-energy events (Main et al. 1992; Amitrano 2003; Schorlemmer et al. 2005). A similar conclusion can also be reached during the risk evaluation of seismic events induced by underground coal mining, which shows that if the b value is large small seismic events are relatively common, whereas when the b value is small, there is a possibility of a strong

✉ Jing Li
lijingxchn@gmail.com

✉ Si-Yuan Gong
gsycumt@gmail.com

¹ Key Laboratory of Deep Coal Resource Mining, Ministry of Education of China, School of Mines, China University of Mining and Technology, Xuzhou, Jiangsu, China

² State Key Laboratory of Coal Resources and Safe Mining, China University of Mining and Technology, Xuzhou, Jiangsu, China

tremor or rockburst (Urbancic 1992; Holub 1996; Melnikov et al. 1996; Mutke et al. 2015). Thus, the temporal variations in b value can be used as a precursor to forecast rockburst hazards.

Recently, passive seismic velocity tomography, widely used for inferring stress redistribution in coal–rock masses, has been proven as a powerful tool for the evaluation of dynamic hazards in underground coal mining. Lurka (2008) and Cao et al. (2015) conducted passive tomographic imaging to locate zones of high seismic activity and to assess rockburst risk. The results indicate that the high velocity and velocity gradient regions correlated well with seismic activities. Luxbacher et al. (2008) evaluated three-dimensional time-lapse velocity tomograms using mining-induced microseismic events as sources, and some high-velocity regions were observed to redistribute as the longwall face retreated. Cai et al. (2014) and Wang et al. (2015) adopted seismic tomography to assess the rockburst risk in longwall panels, and revealed that seismic tomography can be applied to identifying stress redistribution and evaluating rockburst risk. Dou et al. (2014) established rockburst predication parameters based on the relationship between stress and wave velocity, and discovered that velocity tomography is an ideal tool for pre-warning of rockbursts.

As mentioned above, b values and passive seismic tomography are both powerful methods when used to infer rockburst hazards with respect to time and space. The spatial distribution of b values can also be used for inferring the stress state with a large number of long-term earthquake data points (Wiemer and Wyss 1997; Wiemer et al. 1998). However, in the short-term, the spatial imaging of b values in mine-scale regions cannot meet the accuracy requirements due to the paucity of seismic event data. The deficiency of b value imaging in mines can be overcome by seismic tomography. The incentive of the study stems from the ability to explore a mixed method of assessing rockburst risk, which uses temporal variations of b values as the precursor to a rockburst, and then delineates areas of rockburst risk in space using seismic tomography.

Theoretical considerations

The b value theory

The b value is the power law frequency-magnitude relationship defined by Gutenberg and Richter (1944) as:

$$\log N = a - bM, \quad (1)$$

where N is the cumulative number of earthquakes with magnitude larger than M , and parameter a denotes the level of seismic activity, while parameter b , or the b value, is the

slope of the log-linear relationship which indicates the seismotectonic level of activity in the region (Wiemer and Wyss 1997; Wiemer et al. 1998; Wyss et al. 2001); both a and b are constants.

Low-magnitude earthquakes in mines, namely mining-induced seismic events, make the temporal analysis of b value feasible. The relationship between the number of seismic events N and their energy size represented by the quantity of the released seismic energy E being given in joules, can be approximated in log–log scale by the following equation (Holub 1996):

$$\log N = a - b \log E. \quad (2)$$

Seismic events occurring in a defined region during a certain time interval can be used to set-up recurrence curves similar to those used in the Gutenberg–Richter concept. The parameter b in Eq. 2 generally characterises the ratio of weak and strong seismic events, alternatively, b values are often estimated using one of two methods: a least squares method (LSQ), or a maximum likelihood method (MLM). The MLM, which is based on theoretical considerations with fewer uncertainties, is claimed to be a better and more stable method (Aki 1965). In this method, the b value is defined as:

$$b = \frac{\log e}{M_{\text{mean}} - M_{\text{min}}}, \quad (3)$$

where M_{mean} is the average magnitude and M_{min} is the minimum magnitude in the data set.

Seismic velocity tomography

Mechanical properties of rocks, such as rock type, stress, angle of fracture plane, anisotropy, porosity, and modulus of elasticity are correlated with wave velocity, so it is possible to analyse stress distribution in a rock mass by measuring wave velocity variations therein. Research on the relationship between stress and wave velocity has been carried out to good effect, for instance, Nur and Simmons (1969) discovered that both P-wave and S-wave velocities increase with increasing stress, and their increased amplitude was related to the directions of stress and P-wave propagation. Yale (1985) found that the increase of wave velocity with stress was attributed to the closure of cracks and pore spaces. Gong et al. (2012) found that a power function governed the relationship between stress and P-wave velocity obtained from laboratory experiments. The above-mentioned research findings indicate that velocity tomography offers the possibility of assessing the state of stress in coal–rock masses during mining operations.

According to the types of wave sources, velocity tomography can be classified as “active” or “passive”

(Luxbacher et al. 2008; Lurka 2008). The sources of active tomography are normally produced by controlled explosives, hammer strikes, or cutting equipment at a known position; however, the practical limitations of active tomography are obvious, including its infeasibility for long time-lapse investigations and relatively small study area (less than $200 \text{ m} \times 200 \text{ m}$, in general). Those limitations can be overcome by passive tomography, which uses seismic events as sources, as this method can enlarge the investigation area to $2000 \times 2000 \text{ m}$ and shows the stress variations arising during mining.

In this study, passive tomographic imaging was used to infer high stress distribution in a coal–rock mass, and evaluate rockburst hazard zones. Figure 1 shows the principle of seismic velocity tomography in real underground mining operations. The travel time T from source position S (mine tremors) to receiver position R (geophones) represents a line integral of wave slowness P (inverse of velocity v), which can be expressed by Eq. 4 (Luxbacher et al. 2008; Hosseini et al. 2012):

$$T = \int_S^R \frac{dL}{v(x, y, z)} = \int_S^R p dL. \quad (4)$$

To solve Eq. 4, the inversion region is divided into M voxels and travel time T_i in the i th voxel can be described as:

$$T_i = \sum_{j=1}^M d_{ij} p_j \quad (i = 1, \dots, N), \quad (5)$$

Where d_{ij} is the distance of the i th ray travelling in the voxel j , and N is the total number of rays. For the entire set of rays, Eq. 5 can be written in matrix form as shown in Eq. 6 (Luxbacher et al. 2008):

$$T = DP, \quad (6)$$

where T are the column vectors of travel time per ray matrix, P is the slowness per grid cell matrix, and D is the rectangular matrix of distance per ray per voxel matrix ($N \times M$).

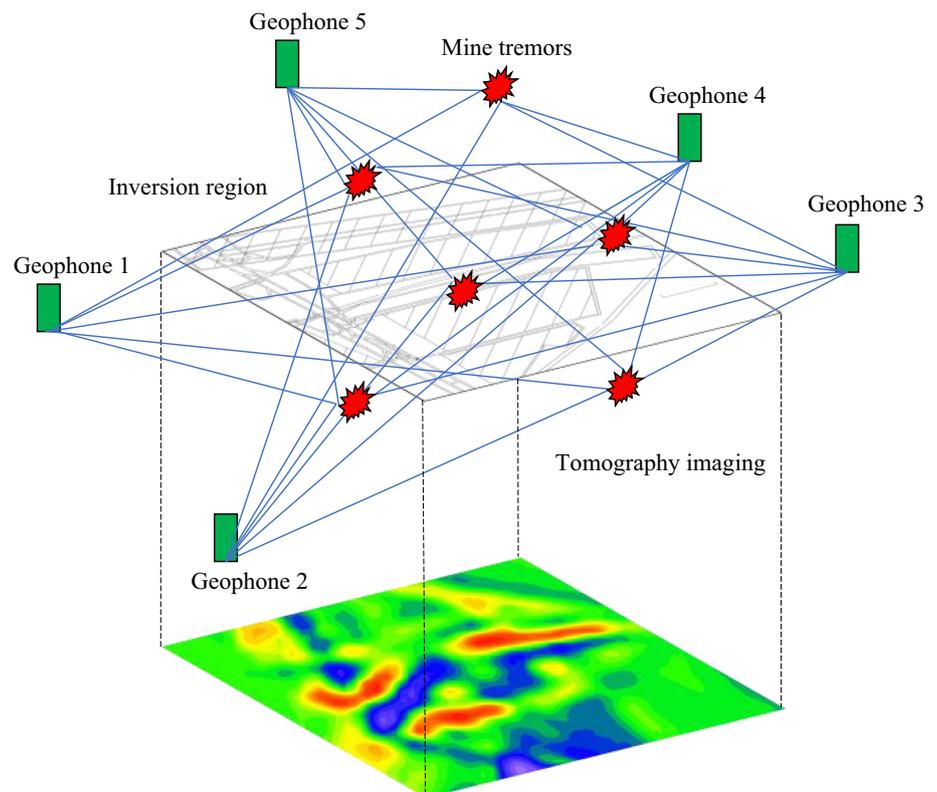
The essence of tomography is solving for the slowness vector P , and the damped least squares method is regarded as an appropriate method to solve for P :

$$P = (DTD + \lambda I)^{-1} D^T T, \quad (7)$$

where λ is the damping parameter and I is the identity matrix.

The method described above requires a substantial amount of time for the matrix inversion, and the most effective way to solve this problem is through an iterative process, such as an algebraic reconstructive technique (ART) or simultaneous iterative reconstructive technique (SIRT) (Gilbert 1972).

Fig. 1 Principle of seismic tomography as applied to underground mining



Rockburst risk assessment

The key to rockburst risk prediction is to determine the stress state and degree of stress concentration in a coal–rock mass. According to the experimental study on the coupling correlation between load and P-wave velocity for coal–rock samples conducted by Gong (2010), there is a positive anomaly in the P-wave velocity in high stress concentration regions, and a negative velocity anomaly in pressure-relief areas. The positive velocity anomaly is expressed as (Gong 2010; Cao et al. 2015):

$$A_n = \frac{V_p - V_p^a}{V_p^a}, \quad (8)$$

where V_p is the P-wave velocity at a certain point in the inversion region, and V_p^a is the average velocity of the model. Table 1 shows the relationship between the positive anomaly in the P-wave velocity and stress concentration, which can be used as the criterion determining potential rockburst risk.

In addition, seismic hazards are also likely to occur in the areas with high stress differences, where there are large velocity variation gradients. The VG value is used to express the degree of variation of velocity gradient, and the anomaly in VG can be shown by (Gong 2010):

$$A_n = \frac{VG - VG^a}{VG^a}, \quad (9)$$

where VG^a is the average value of VG in the model. Similarly, the relationship between VG anomalies and rockburst risk can be shown in Table 2.

Table 1 Relationship between positive anomaly in wave velocity and stress concentration (Gong 2010)

| Rockburst risk index | Stress concentration | Positive velocity anomaly, A_n (%) |
|----------------------|----------------------|--------------------------------------|
| 0 | None | <5 |
| 1 | Weak | 5–15 |
| 2 | Middle | 15–25 |
| 3 | Strong | >25 |

Table 2 Relationship between VG anomaly and rockburst risk (Gong 2010)

| Rockburst risk index | Stress difference | VG anomaly, A_n (%) |
|----------------------|-------------------|-----------------------|
| 0 | None | <5 |
| 1 | Weak | 5–15 |
| 2 | Middle | 15–25 |
| 3 | Strong | >25 |

Case study

Site description

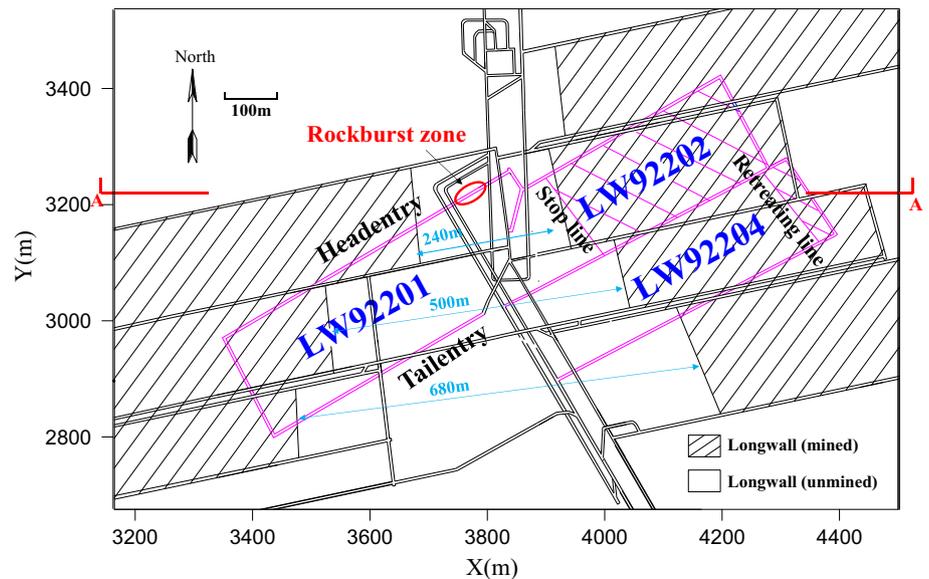
Sanhejian Coal Mine, owned by the Xuzhou Coal Group, is situated in the north of Jiangsu Province, China. The study area, South No. 2 District, includes 7# and 9# coal seams, and the vertical distance between the two coal seams normally ranges from 20 to 30 m. The accompanying surface subsidence was observed in the mining operations of panels in 7# coal seam, and irregular coal pillars with different widths (240, 500, and 680 m) between gobs were left to protect these main roadways, as shown in Fig. 2. Under the 7# coal seam is the 9# coal seam with a thickness ranging from 0.32 to 4.7 m (about 2.2 m on average). The average angle of inclination of the 9# seam is 12°. The roof of the 9# coal seam consists of 12.4 m of medium sandstone, 17.78 m of mudstone, 3.61 m of the 7# seam coal, and below the 9# coal seam there are 1.2 m of clay rock, 4.8 m of siltstone, and 5.2 m of fine sandstone (Fig. 3). There are three panels (LW92201, LW92202, and LW92204) with mining depths of about 1010 m underground in the 9# coal seam. Panel LW92204 is being mined adjacent to the LW92202 gob, and its position is 45 m from the beginning line (the rate of advance is about 1.5 m per day). Panel LW92201 is extracted after LW92204; now its headentry and tailentry are under preparation, as shown in Fig. 2.

At 3:47:04, on 30 January 2015, a strong tremor with an energy of 1.48×10^4 J occurred near the 240-m pillar and triggered a dynamic failure in the headentry of panel LW92201 (Fig. 2), which resulted in serious deformation of the roadway and damage to the communication system (Fig. 4). Five miners were seriously injured due to the strong shock wave released from rockburst. The event indicated that high stress concentration zones seen as the main threat to nearby panels existed in the South No. 2 District, so it was necessary to find the precursor to intensive seismic activities or rockbursts, and further implement destressing measures to prevent their recurrence.

Rockburst mechanism

In Fig. 5, a “T” structure in the overlying strata is presented due to the existence of coal pillars in the 7# coal seam when examining the cross-sectional geometry on line A–A (Fig. 2). According to research results (Dou and He 2014), the occurrence of a rockburst is related to such a “T” spatial structure. To study the stress distribution in the 9# coal seam under the influence of this “T” structure, FLAC^{3D} was adopted to establish a model for stress field analysis. The properties of the coal–rock in the model are consistent with measured borehole data, and model size is 2000 × 1000 × 200 m in length, width, and height,

Fig. 2 Layout of panels in South No. 2 District and the rockburst occurring on 30 January 2015 (panels in the 9# coal seam represented by pink lines lie above panels in the 7# coal seam represented by black lines)



| Lithology | Geology columnar | Thickness(m) | Depth(m) |
|------------------|------------------|--------------|----------|
| Grit sandstone | | 6.2 | 963.99 |
| Fine sandstone | | 4.7 | 968.69 |
| Siltstone | | 3.8 | 972.49 |
| Medium sandstone | | 2.5 | 974.99 |
| Siltstone | | 1.49 | 976.48 |
| 7# coal seam | | 3.61 | 980.08 |
| Mudstone | | 17.78 | 997.87 |
| Medium sandstone | | 12.4 | 1010.27 |
| 9# coal seam | | 2.2 | 1012.47 |
| Clay rock | | 1.2 | 1013.67 |
| Siltstone | | 4.8 | 1018.47 |
| Fine sandstone | | 5.2 | 1023.61 |
| Siltstone | | 3.09 | 1026.76 |

Fig. 3 Columnar illustration of coal and rock layers in South No. 2 District

respectively. Figure 6 shows that there are high-stress zones in the 9# coal seam under the “T” structure, and that the rockburst happened in that relatively highly stressed region. As the coal–rock mass, when subjected to high stresses, is a balanced system of energy input, storage and release, a rockburst is less likely to occur under high static load. However, mining disturbances, accompanied by roof fracture, pillar failure, and fault slip, often lead to tremors,



Fig. 4 Deformation of the roadway caused by a rockburst

among which strong tremors would be a trigger for a rockburst. According to the theory of dynamic and static load superposition (He et al. 2014; Dou et al. 2015), coal will be damaged when the superposition of dynamic load σ_d and static load σ_s exceeds the coal strength $[\sigma_c]$ (namely, $\sigma_d + \sigma_s > [\sigma_c]$).

Figure 7 describes the possible mechanism of rockburst evolution in a roadway: $[\sigma_c]$ is defined as the critical value, and as the excavation of the roadway progresses, the stress on the coal wall sides will be redistributed (Fig. 7a); as the peak stress is less than the critical value, the coal mass is in a stable state; however, due to disturbance by tremors, if dynamic and static load superposition in the coal causes the stress to exceed the critical value, a rockburst may occur (Fig. 7b). The coal where the stress exceeds the critical value is defined as region A, and the coals on both sides of A are defined as regions B and C, respectively (Fig. 7b).

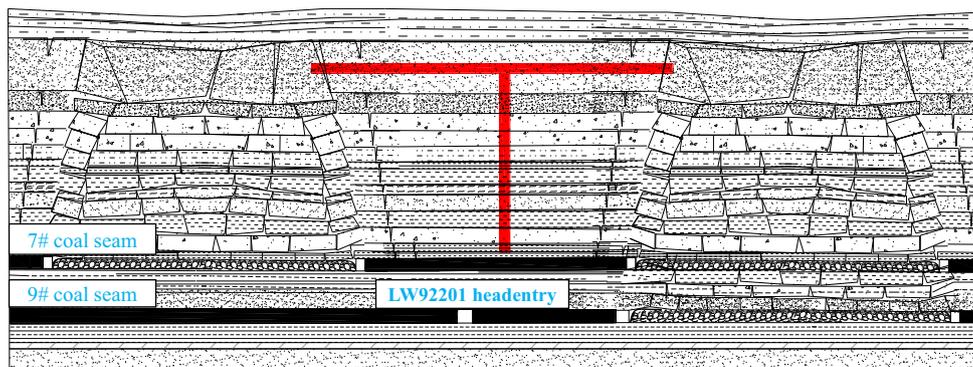


Fig. 5 “T” spatial structure of overlying strata

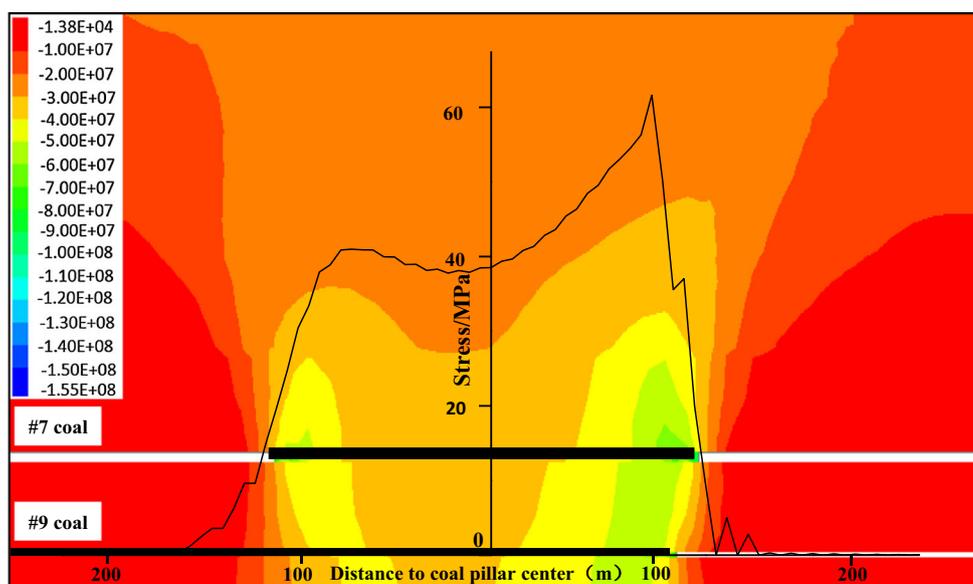


Fig. 6 The stress distribution in the 9# coal seam under the coal pillar

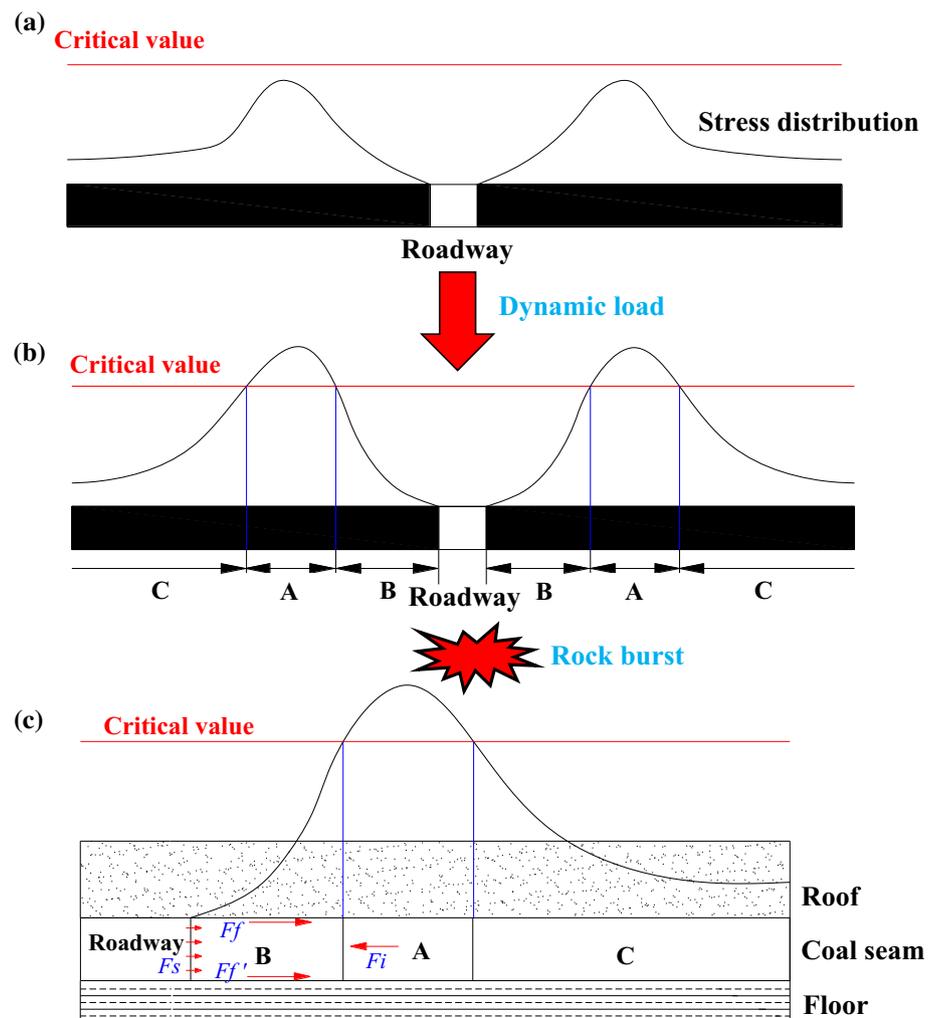
The coal in region *A* is in a dilatant stage and undergoes unstable rapid expansion, causing the impact force F_i acting on both regions, *B* and *C* (Fig. 7c). By comparison, the coal in region *B*, which is closer to the roadway, is more likely to be pushed out. The force acting across region *B* can be decomposed into three parts: a support force from the bolt and anchor acting on the coal wall (F_s), the friction force between the coal in region *B* and the roof (F_f), and the friction force between the coal in region *B* and the floor (F_f'). When the impact force is higher than the resistance (namely, $F_i > F_s + F_f + F_f'$), a rockburst will occur (Fig. 7c).

According to the above analysis, high stress in coal–rock mass and strong seismic events both need to be analysed when assessing a rockburst; thus, b values and seismic tomography were used for tremor analysis and the depiction of high-stress zones in this study.

The b value calculation

A microseismic monitoring system (“SOS”, developed by the Central Mining Institute (GIG) of Poland) was installed in the study area to record seismic events for b value calculation. From 1 to 31 January 2015, three strong tremors occurred with energies of over 10^4 J in the study area (Table 3). The frequency-energy distribution method was used to calculate the threshold energy M_e (Wiemer and Wyss 2000; Holub 1996), and the minimum energy for completeness of seismic events in the region of the investigation was found to be $M_e = 340$ J. Then 93 seismic events in January 2015 with an energy $M_e > 340$ J are used for b value calculations with a successive time-window length of five days and a time shift of one day. The b value in each window is calculated by MLM. The calculated b values are shown in Table 3.

Fig. 7 Analysis of rockburst mechanism in a roadway: **a** stress distribution on both sides of a roadway; **b** stress exceeding critical value due to dynamic load disturbance; **c** impact force larger than resistance leading to a rockburst



P-wave velocity inversion

In this study, velocity tomography was conducted using mining-induced seismic events in January 2015. To improve the inversion accuracy, the seismic events recorded by over four stations were adopted to avoid artificial anomalies affecting the tomograms. After data selection, 95 seismic events located in the study area were analysed using the MINESOSTOMO program developed by Gong (2010) to generate tomographic images. SIRT was chosen as the inversion algorithm because the solution tends to be slowly convergent or divergent. Here, $70 \times 80 \times 3$ voxel points were created and each voxel measured $45 \times 40 \times 620$ m in the X-, Y-, and Z-directions, respectively. To reduce indeterminacy therein, the maximum velocity was restricted to 6.0 km/s. The calculated velocity images were derived and sliced at the 9# coal seam level.

Results and discussion

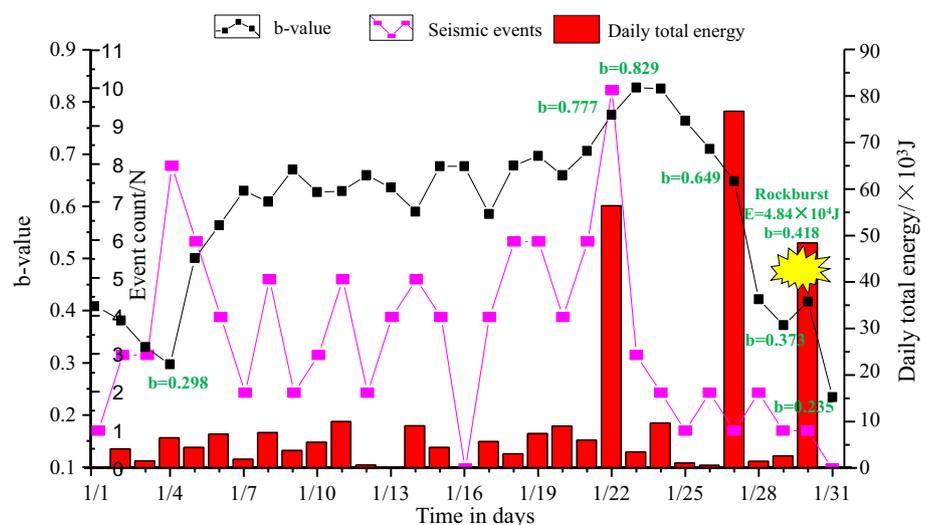
Temporal forecasting of rockbursts

Figure 8 shows the variations in b values, event counts, and daily total tremor energy from 1 to 31 January 2015 in the study area. It can be seen that the b values fluctuated from 0.298 to 0.777 between 1 and 22 January, but maintained a relatively high level as a whole, which indicated that seismic activities in this period were normal. However, after 22 January, b values dropped significantly, from 0.829 to 0.235. Before the rockburst (from 27 to 29 January), the b value suddenly decreased to an abnormally low level, and reached 0.373 on 29 January, which demonstrated that the danger level of rockburst at this stage was very high. Ultimately, the rockburst was triggered on 30 January, and therefore, it can be verified that the continuous decrease, and abnormally low, b values may be regarded as an effective precursor to a rockburst.

Table 3 Tremor details and *b* value calculation results in the study area in January 2015

| Date | Event counts | Daily total energy (J) | Maximum energy of tremors (J) | <i>b</i> value |
|-------------|--------------|--------------------------------------|--------------------------------------|----------------|
| 1 Jan 2015 | 1 | 2.13×10^2 | 2.13×10^2 | 0.409 |
| 2 Jan 2015 | 3 | 4.15×10^3 | 3.26×10^3 | 0.383 |
| 3 Jan 2015 | 3 | 1.56×10^3 | 8.44×10^2 | 0.332 |
| 4 Jan 2015 | 8 | 6.53×10^3 | 3.39×10^3 | 0.298 |
| 5 Jan 2015 | 6 | 4.45×10^3 | 2.56×10^3 | 0.502 |
| 6 Jan 2015 | 4 | 7.34×10^3 | 6.13×10^3 | 0.565 |
| 7 Jan 2015 | 2 | 1.92×10^3 | 1.71×10^3 | 0.631 |
| 8 Jan 2015 | 5 | 7.63×10^3 | 3.53×10^3 | 0.610 |
| 9 Jan 2015 | 2 | 3.79×10^3 | 3.26×10^3 | 0.672 |
| 10 Jan 2015 | 3 | 5.54×10^3 | 3.27×10^3 | 0.629 |
| 11 Jan 2015 | 5 | 1.00×10^4 | 6.94×10^3 | 0.630 |
| 12 Jan 2015 | 2 | 6.80×10^2 | 3.75×10^2 | 0.661 |
| 13 Jan 2015 | 4 | 1.96×10^2 | 8.15×10^2 | 0.637 |
| 14 Jan 2015 | 5 | 9.11×10^3 | 3.16×10^3 | 0.591 |
| 15 Jan 2015 | 4 | 4.47×10^3 | 2.88×10^3 | 0.678 |
| 16 Jan 2015 | 0 | 0 | 0 | 0.678 |
| 17 Jan 2015 | 4 | 5.70×10^3 | 2.46×10^3 | 0.586 |
| 18 Jan 2015 | 6 | 3.08×10^3 | 6.88×10^2 | 0.679 |
| 19 Jan 2015 | 6 | 7.43×10^3 | 1.60×10^3 | 0.698 |
| 20 Jan 2015 | 4 | 9.04×10^3 | 1.90×10^3 | 0.661 |
| 21 Jan 2015 | 6 | 5.99×10^3 | 2.32×10^3 | 0.707 |
| 22 Jan 2015 | 10 | 5.61×10^4 | 3.16×10^4 | 0.777 |
| 23 Jan 2015 | 3 | 3.50×10^3 | 2.54×10^3 | 0.829 |
| 24 Jan 2015 | 2 | 9.68×10^3 | 9.59×10^3 | 0.827 |
| 25 Jan 2015 | 1 | 1.11×10^3 | 1.11×10^3 | 0.765 |
| 26 Jan 2015 | 2 | 6.02×10^2 | 5.08×10^2 | 0.711 |
| 27 Jan 2015 | 1 | 7.67×10^4 | 7.67×10^4 | 0.649 |
| 28 Jan 2015 | 2 | 1.45×10^3 | 1.02×10^3 | 0.423 |
| 29 Jan 2015 | 1 | 2.65×10^3 | 2.65×10^3 | 0.373 |
| 30 Jan 2015 | 1 | 4.84×10^4 | 4.84×10^4 | 0.418 |
| 31 Jan 2015 | 0 | 0 | 0 | 0.235 |

Bold values represent high-energy tremors or special *b* values

Fig. 8 The *b* values, seismic event counts, and daily total energy released: January 2015

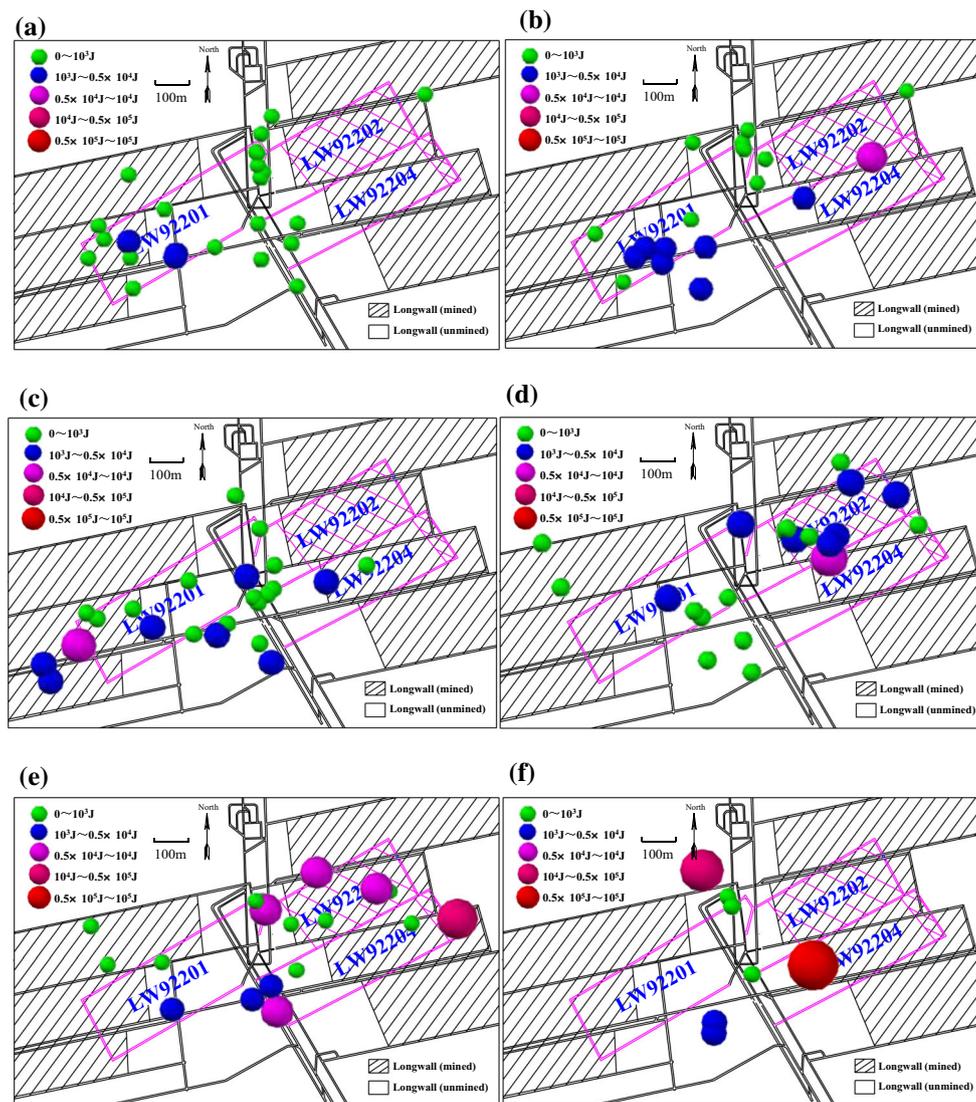


Fig. 9 Tremor distributions at five-day intervals, from 1 to 31 January (Panels in the 9# coal seam represented by pink lines lie above panels in the 7# coal seam and are represented by black lines. Panels in the 7# coal seam are already mined. In the 9# coal seam,

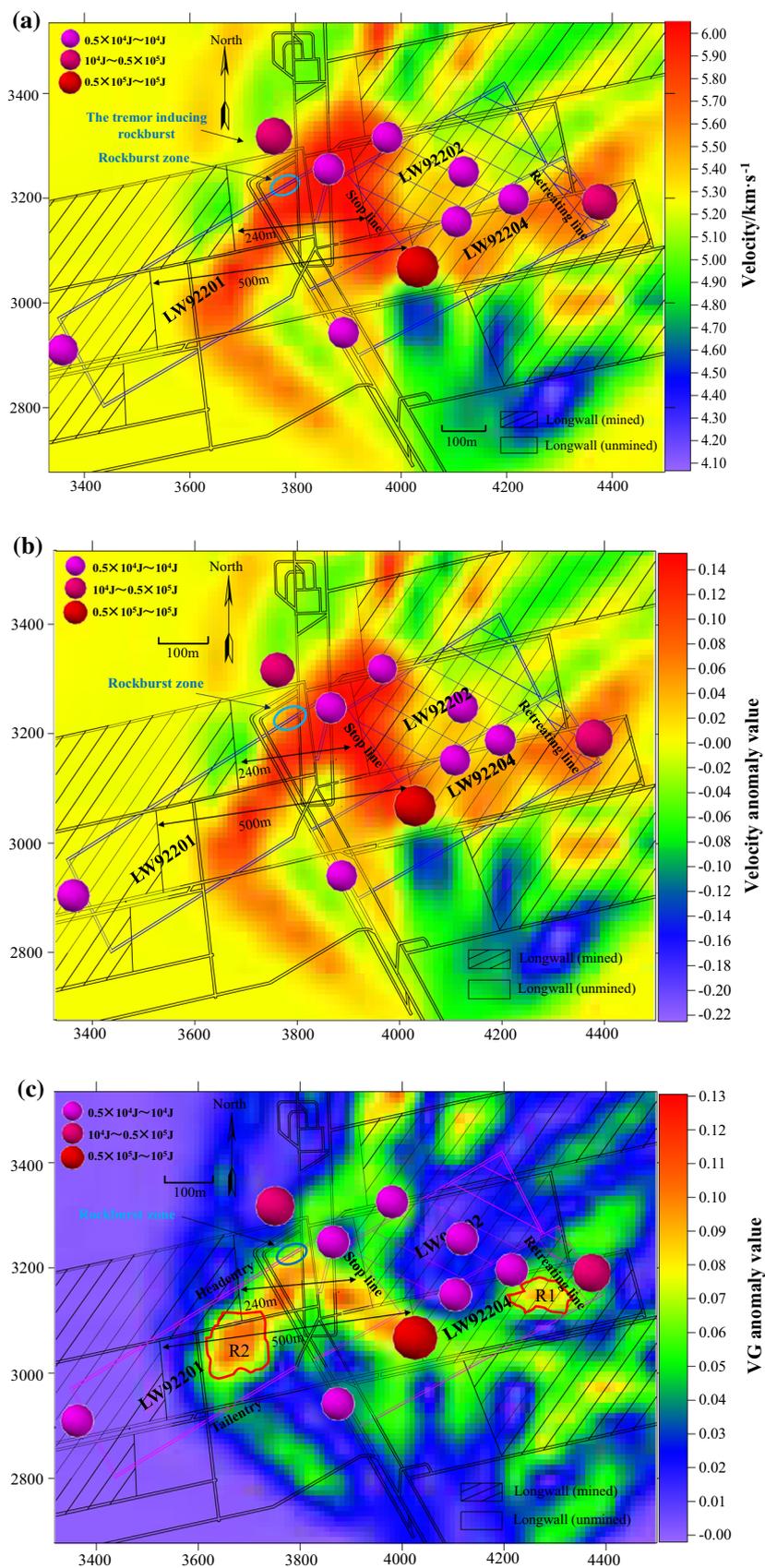
panel LW92202 has been mined, and panel LW92204 is being mined, but mining operations in panel LW92201 had not yet started). **a** 1–5 January, **b** 6–10 January, **c** 11–16 January, **d** 17–21 January, **e** 22–26 January, **f** 27–31 January

The variations in event counts and daily total tremor energy support the accuracy of the b value analysis. According to Fig. 8, before 21 January, the daily total energy was generally lower and basically remained stable, while the event count underwent fluctuations during mining operations, which indicated that coal–rock fracturing intensity, and the corresponding amount of energy released, gradually increased and a large number of micro-fissures formed. This corresponded to b value variations in this period. However, after 21 January, sharp increases in daily tremor total energy appeared on 22, 27, and 29 January, and simultaneously, event counts dropped significantly from 22 to 25 January, then remained at a lower level until the end of the period which demonstrated that

micro-fissures inside the coal–rock mass have propagated, converged, and interconnected to form macro-fracture. At this stage, a downward trend was seen in the b values.

Figure 9 shows the distribution of tremors every five days in January 2015. In that period, low-energy tremors initially played a dominant role in the region (Fig. 9a), then the number of larger-energy tremors increased gradually (Fig. 9b–d), and finally the proportion of strong tremors increased rapidly (Fig. 9e–f). A good agreement between a significant decrease in b values and an increased number of strong tremors was found, proving that the temporal variation in b values can be used as precursory information for rockburst prediction. (a) 1 to 5 January, (b) 6 to 10 January, (c) 11 to 16 January, (d) 17 to 21 January, (e) 22 to 26 January, (f) 27 to 31 January.

Fig. 10 Tomography results at the 9# coal seam obtained using seismic events in January 2015 (circles show the positions of large seismic tremors with energies $E > 0.5 \times 10^4$ J): **a** velocity inversion result; **b** velocity anomaly inversion result; **c** velocity gradient anomaly inversion result



The b value variation reflected the level of rockburst risk in the study area, which was basically consistent with the appearance of tremors with different energy levels. The deficiency of the variation in b values means that the engineer cannot obtain the risk distribution in the danger zones; thus, seismic tomography was further used for rockburst pre-warning in the study area.

Determining rockburst risk areas in spatial terms

Calculated tomographic images at the 9# coal seam level, $Z = -1010$ m, are shown in Fig. 10. High velocity and velocity gradient anomalies were observed in the 240 and 500-m coal pillar areas and in the vicinity of the panel LW92202 stop line, which corresponded to the “T” spatial structure and numerical simulation results (Figs. 5 and 6). The maximum P-wave velocity value reached 6 km/s, while the maximum values of positive velocity anomaly and velocity gradient anomaly were 0.14 and 0.13, respectively, which indicated that the aforementioned areas were in their rockburst risk potential stage (Tables 1, 2). Moreover, the rockburst which occurred on 30 January 2015 was also discovered to have been in a high velocity and velocity gradient anomaly zone, proving the accuracy of the seismic tomography with regard to its inferring the presence and location of such a high-stress zone, and indeed in forecasting rockburst risk.

During the study period, ten mine tremors with energies of over 0.5×10^4 J were recorded between 1 and 31 January in the area under investigation: most of them were located within the high velocity and velocity anomaly zones (Fig. 10a, b). In particular, a strong tremor with an energy release of 1.48×10^4 J occurred about 100 m from the panel LW92201 headentry in its north side, inducing a serious rockburst hazard (Figs. 4, 10a).

Tomography results also made it possible to forecast potential rockburst zones for the future mining. As shown in Fig. 10c, a high velocity gradient anomaly region named R1, which was circled by red irregular lines, was located about 20 m ahead of panel LW92204's retreating line, which indicated that this coal–rock mass was subjected to high stress differences in that area and that rockbursts were liable to occur under dynamic load disturbance (see Table 2). A similar region, named R2, with a high velocity gradient anomaly was observed inside panel LW92201 (Fig. 10c), which was also a potential rockburst risk area. Some pressure relief measures had to be taken in regions R1 and R2 before mining to prevent the occurrence of strong tremors or rockbursts.

Conclusions

A rockburst assessment method combining b values and velocity tomography was applied to longwall panels in Sanhejian Coal Mine, Jiangsu Province, China. The case study in panels LW92201, LW92202, and LW92204 indicated that a significant decrease of seismic b values from 0.829 to 0.373 was detected before a rockburst occurred on 30 January 2015. Simultaneously, seismic event counts declined markedly in this period while the daily tremor total energy increased significantly on 22 and 27 January. Besides, there was good agreement between the significant decrease in b values and the appearance of strong tremors.

Based on the seismic tomography, two rockburst risk areas with high velocity or velocity gradient anomalies were observed within irregular coal pillars and panel LW92204. The rockburst zone was located in the rockburst risk area mentioned above, where the maximum velocity, maximum velocity anomaly and maximum velocity gradient anomaly were 6 km/s, 0.14 and 0.13, respectively. There was also a good correlation between the areas of high velocity, or velocity anomalies, and areas of high seismic activity. The application of the proposed method gave the engineers the opportunity to temporally forecast strong tremors or rockbursts, and spatially depict high-stress regions, which was shown to be a promising rockburst assessment method.

Acknowledgements We acknowledge the financial support provided for this work by the National Natural Science Foundation of China (51404269), Fundamental Research Funds for the Central Universities (2013QNA46, 2014ZDPY09), the Jiangsu Natural Science Foundation (BK20130183) and the Research and Innovation Project for College Graduates of Jiangsu Province (KYZZ16_0223).

References

- Aki K (1965) Maximum likelihood estimate of b in the formula $\log N = a - bM$ and its confidence limits. Bull Earth Res Inst Univ Tokyo 43:237–239
- Amitrano D (2003) Brittle–ductile transition and associated seismicity: experimental and numerical studies and relationship with the b value. J Geophys Res 108(B1):2044. doi:10.1029/2001JB000680
- Cai W, Dou LM, Cao AY, Gong SY, Li ZL (2014) Application of seismic velocity tomography in underground coal mines: a case study of Yima mining area, Henan, China. J Appl Geophys 109:140–149. doi:10.1016/j.jappgeo.2014.07.021
- Cao AY, Dou LM, Cai W, Gong SY, Liu S, Jing GC (2015) Case study of seismic hazard assessment in underground coal mining using passive tomography. Int J Rock Mech Min Sci 78:1–9. doi:10.1016/j.ijrmms.2015.05.001
- Dou LM, He H (2014) Spatial structure evolution of overlying strata and inducing mechanism of rockburst in coal mine. Trans

- Nonferrous Met Soc China 24(4):1255–1261. doi:[10.1016/S1003-6326\(14\)63187-3](https://doi.org/10.1016/S1003-6326(14)63187-3)
- Dou LM, Chen TJ, Gong SY, He H, Zhang SB (2012) Rockburst hazard determination by using computed tomography technology in deep workplace. *Saf Sci* 5(4):736–740. doi:[10.1016/j.ssci.2011.08.043](https://doi.org/10.1016/j.ssci.2011.08.043)
- Dou LM, Cai W, Gong SY, Han RJ, Liu J (2014) Dynamic risk assessment of rock burst based on the technology of seismic computed tomography detection. *J China Coal Soc* 39(2):238–244. doi:[10.13225/j.cnki.jccs.2013.2016](https://doi.org/10.13225/j.cnki.jccs.2013.2016)
- Dou LM, He J, Cao AY, Gong SY, Cai W (2015) Rock burst prevention methods based on theory of dynamic and static combined load induced in coal mine. *J. China Coal Soc* 40(7):1469–1476. doi:[10.13225/j.cnki.jccs.2014.1815](https://doi.org/10.13225/j.cnki.jccs.2014.1815)
- Ge M (2005) Efficient mine microseismic monitoring. *Int J Coal Geol* 64:44–56. doi:[10.1016/j.coal.2005.03.004](https://doi.org/10.1016/j.coal.2005.03.004)
- Gilbert P (1972) Iterative methods for the three-dimensional reconstruction of an object from projections. *J Theor Biol* 36(1):105–117. doi:[10.1016/0022-5193\(72\)90180-4](https://doi.org/10.1016/0022-5193(72)90180-4)
- Gong SY (2010) Research and application of using mine tremor velocity tomography to forecast rockburst danger in coal mine. China University of Mining and Technology, Xuzhou
- Gong SY, Dou LM, He J, He H, Lu CP, Mu ZL (2012) Study of correlation between stress and longitudinal wave velocity for deep burst tendency coal and rock samples in uniaxial cyclic loading and unloading experiment. *Rock Soil Mech* 33(1):41–47. doi:[10.3969/j.issn.1000-7598.2012.01.007](https://doi.org/10.3969/j.issn.1000-7598.2012.01.007)
- Gutenberg B, Richter C (1944) Frequency of earthquakes in California. *Bull Seismol Soc Am* 34(4):185–188
- He MC, Qian QH (2010) *The Basis of Deep Rock Mechanics*. Science Press, Beijing (in Chinese)
- He J, Dou LM, Cai W, Li ZL, Ding YL (2014) Mechanism of dynamic and static combined load inducing rockburst in thin coal seam. *J China Coal Soc* 39(11):2177–2182. doi:[10.13225/j.cnki.jccs.2013.1603](https://doi.org/10.13225/j.cnki.jccs.2013.1603)
- Holub K (1996) Space-time variations of the frequency-energy relation for mining-induced seismicity in the Ostrava-Karvina mining district. *Pure Appl Geophys* 146(2):265–280. doi:[10.1007/BF00876493](https://doi.org/10.1007/BF00876493)
- Hosseini N, Oraee K, Shahriar K, Goshtasbi K (2012) Passive seismic velocity tomography and geostatistical simulation on longwall mining panel. *Arch Min Sci* 57:139–155. doi:[10.2478/v10267-012-0010-9](https://doi.org/10.2478/v10267-012-0010-9)
- Lurka A (2008) Location of high seismic activity zones and seismic hazard assessment in Zabrze Bielszowice coal mine using passive tomography. *J China Univ Min Technol* 18(2):177–181. doi:[10.1016/S1006-1266\(08\)60038-3](https://doi.org/10.1016/S1006-1266(08)60038-3)
- Luxbacher K, Westman E, Swanson P, Karfakis M (2008) Three-dimensional time-lapse velocity tomography of an underground longwall panel. *Int J Rock Mech Min Sci* 45(4):478–485. doi:[10.1016/j.ijrmms.2007.07.015](https://doi.org/10.1016/j.ijrmms.2007.07.015)
- Main IG, Meredith PG, Sammonds PR (1992) Temporal variations in seismic event rate and *b* values from stress corrosion constitutive laws. *Tectonophysics* 211(1–4):233–246. doi:[10.1016/0040-1951\(92\)90061-A](https://doi.org/10.1016/0040-1951(92)90061-A)
- Melnikov NN, Kozyrev AA, Panin VI (1996) Induced seismicity in large-scale mining in the Kola Peninsula and monitoring to reveal informative precursors. *Pure Appl Geophys* 147(2):263–276. doi:[10.1007/BF00877482](https://doi.org/10.1007/BF00877482)
- Mutke G, Józef D, Lurka A (2015) New criteria to assess seismic and rock burst hazard in coal mines. *Arch Min Sci* 60(3):743–760. doi:[10.1515/amsc-2015-0049](https://doi.org/10.1515/amsc-2015-0049)
- Nur A, Simmons G (1969) Stress-induced velocity anisotropy in rock: an experimental study. *J Geophys Res* 74(27):6667–6674. doi:[10.1029/JB074i027p06667](https://doi.org/10.1029/JB074i027p06667)
- Schorlemmer D, Weimer S, Wyss M (2005) Variations in earthquake-size distribution across different stress regimes. *Nature* 437:539–542. doi:[10.1038/nature04094](https://doi.org/10.1038/nature04094)
- Urbancic TI, Trifu C-I, Long JM, Young RP (1992) Space-time correlation of *b* values with stress release. *Pure Appl Geophys* 139(3/4):449–462. doi:[10.1007/BF00879946](https://doi.org/10.1007/BF00879946)
- Wang GF, Gong SY, Li ZL, Dou LM, Wu C, Mao Y (2015) Evolution of stress concentration and energy release before rock bursts: two case studies from Xingan coal mine, Hegang, China. *Rock Mech Rock Eng* 11:1–9. doi:[10.1007/s00603-015-0892-x](https://doi.org/10.1007/s00603-015-0892-x)
- Wiemer S, Wyss M (1997) Mapping the frequency-magnitude distribution in asperities: an improved technique to calculate recurrence times. *J Geophys Res* 102:15115–15128. doi:[10.1029/97JB00726](https://doi.org/10.1029/97JB00726)
- Wiemer S, Wyss M (2000) Minimum magnitude of completeness in earthquake catalogs: examples from Alaska, the western United States, and Japan. *Bull Seismol Soc Am* 90(4):859–869. doi:[10.1785/0119990114](https://doi.org/10.1785/0119990114)
- Wiemer S, McNutt SR, Wyss M (1998) Temporal and three-dimensional spatial analyses of the frequency-magnitude distribution near Long Valley Caldera, California. *Geophys J Int* 134(2):409–421. doi:[10.1046/j.1365-246x.1998.00561.x](https://doi.org/10.1046/j.1365-246x.1998.00561.x)
- Wyss M, Klein F, Nagamine K, Weimer S (2001) Anomalously high *b* values in the South Flank of Kilauea Hawaii: evidence for the distribution of magma below Kilauea's East Rift Zone. *J Volcanol Geotherm Res* 106(1–2):23–37. doi:[10.1016/S0377-0273\(00\)00263-8](https://doi.org/10.1016/S0377-0273(00)00263-8)
- Yale D (1985) Recent advances in rock physics. *Geophysics* 50(12):2480–2491
- Young RP, Maxwell SC, Urbancic TI, Feignier B (1992) Mining-induced microseismicity: monitoring and applications of imaging and source mechanism techniques. *Pure appl Geophys* 139(3):697–719. doi:[10.1007/BF00879959](https://doi.org/10.1007/BF00879959)

Insight into variability of spring and flash flood events in Lithuania

Diana Meilutytė-Lukauskienė¹ · Vytautas Akstinas¹ · Jūratė Kriauciūnienė¹ · Diana Šarauskiene¹ · Aldona Jurgelėnaitė¹

Received: 9 January 2017 / Accepted: 16 January 2017 / Published online: 31 January 2017
© Institute of Geophysics, Polish Academy of Sciences & Polish Academy of Sciences 2017

Abstract In this research, variability of spring (from 1 March to 30 May) and flash (from 1 June to 30 November) floods in rivers of different regions was analysed. The territory of Lithuania is divided into three regions according to hydrological regime of the rivers: Western, Central, and Southeastern. The maximum river discharge data of spring and flash floods [a total of 31 water gauging stations (WGS)] were analysed. Comparison of the data of four periods (1922–2013, 1941–2013, 1961–2013, and 1991–2013) with the data of the reference period (1961–1990) was performed. Analysis included the longest discharge data set of the Nemunas River at Smalininkai WGS (1812–2013) as well. Mixed patterns of flood changes in Lithuanian rivers were detected. The analysis of flood discharges of the Nemunas River indicated that both spring and flash floods in Lithuania were getting smaller.

Keywords Spring and flash floods · Lithuanian rivers · Mann–Kendall and Spearman–Rho tests

Introduction

Floods have been the most widely occurring natural disaster in human history. Recently, changing climate makes floods even more severe and threatening. A considerable amount of the literature has been published on this problem. Researchers endeavour to understand the nature of this hardly controllable process that can appear unexpectedly,

make huge damage and cause loss of human lives. According to European Environment Agency, Europe suffered over 213 major damaging floods in the period of 1998–2009. Floods in the EU have caused some 1126 deaths, the displacement of about half a million people and at least €52 billion in insured economic losses. Such statistics prompted the European Parliament to adopt Directive 2007/60/EC on the assessment and management of flood risks. This Directive requires Member States to make every effort to reduce the flood risk.

This natural phenomenon can have various origins (natural and/or anthropogenic) and may be grouped into different ways. For example, Han (2001) suggests classifying floods into flash, fluvial (river floods), pluvial (rain floods), urban (urban pluvial floods), groundwater, coastal, and floods by dam break. Brazdil et al. (2002) proposes four basic types: short but intense downpours or cloud-bursts (flash floods), long-lasting continuous rainfall, snowmelt, and ice damming. Barredo (2007) argues that floods are usually classified into three main types: river flood, flash flood, and storm surge.

The study of Barredo (2007) is one of those that deal with a large amount of information on flood disasters for long time periods in Europe. This researcher attempted to identify and classify the most destructive floods in 1950–2005. Italy, Spain, France, and Germany were the countries affected by the largest number of major floods. It was identified that all mentioned floods in Spain are classified as flash floods. This country was considered as the most affected by this type of flood; Italy and France are the next ones in this list. In the case of river floods, Italy is the first, followed by Germany, France, and Poland. The largest floods in Europe over the period of 1985–2009 are listed by Kundzewicz et al. (2013). The authors repeat the same affected countries as well as the Netherlands, the

✉ Diana Meilutytė-Lukauskienė
diana.meilutyte-lukauskiene@lei.lt

¹ Laboratory of Hydrology, Lithuanian Energy Institute, Kaunas, Lithuania

Table 1 Characteristics of WGS

| No. | Station | River | Basin | Basin area, km ² | Data series | Average of 1961–1990 | |
|------------------------|-----------|--------------|------------|-----------------------------|-------------|----------------------|--------------------|
| | | | | | | Q_{Spring} | Q_{Flash} |
| Western Lithuania | | | | | | | |
| 1 | Venta | Papilė | Venta | 1570 | 1948–2010 | 114 | 86 |
| 2 | Venta | Leckava | Venta | 4060 | 1949–2010 | 231 | 256 |
| 3 | Rešketa | Gudeliai | Venta | 84 | 1947–1999 | 9 | 6 |
| 4 | Bartuva | Skuodas | Baltic Sea | 612 | 1957–2010 | 71 | 112 |
| 5 | Jūra | Pajūris | Nemunas | 876 | 1946–1999 | 118 | 192 |
| 6 | Jūra | Tauragė | Nemunas | 1690 | 1925–2010 | 210 | 282 |
| 7 | Akmena | Paakmenis | Nemunas | 314 | 1955–2010 | 44 | 55 |
| 8 | Šešuvis | Skirgailiai | Nemunas | 1880 | 1941–2010 | 163 | 147 |
| 9 | Šyša | Jonaičiai | Nemunas | 174 | 1960–1999 | 16 | 27 |
| 10 | Minija | Kartena | Nemunas | 1230 | 1925–2010 | 117 | 221 |
| 11 | Veiviržas | Mikužiai | Nemunas | 336 | 1954–1999 | 44 | 93 |
| 12 | Nemunas | Smalininkai | Nemunas | 81,200 | 1812–2010 | 1881 | 1588 |
| Central Lithuania | | | | | | | |
| 13 | Agluona | Dirvonakiai | Lielupė | 66 | 1946–1999 | 9 | 5 |
| 14 | Mūša | Ustukiai | Lielupė | 2280 | 1958–2010 | 152 | 86 |
| 15 | Lėvuo | Kupiškis | Nemunas | 307 | 1955–1999 | 25 | 17 |
| 16 | Nevėžis | Dasiūnai | Nemunas | 5530 | 1961–2010 | 300 | 199 |
| 17 | Šušvė | Josvainiai | Nemunas | 1100 | 1941–1999 | 76 | 70 |
| 18 | Dubysa | Lyduvėnai | Nemunas | 1070 | 1941–2010 | 73 | 59 |
| 19 | Dubysa | Padubysys | Nemunas | 1840 | 1930–1999 | 113 | 104 |
| 20 | Alsa | Paalsys | Nemunas | 49 | 1957–1999 | 6 | 7 |
| 21 | Neris | Jonava | Nemunas | 24,600 | 1920–2010 | 660 | 561 |
| Southeastern Lithuania | | | | | | | |
| 22 | Merkys | Puvočiai | Nemunas | 4300 | 1946–2010 | 97 | 99 |
| 23 | Ūla | Zervynos | Nemunas | 679 | 1960–2010 | 23 | 22 |
| 24 | Verknė | Verbyliškės | Nemunas | 694 | 1952–2010 | 31 | 30 |
| 25 | Žeimena | Pabradė | Nemunas | 2580 | 1954–2010 | 49 | 56 |
| 26 | Šventoji | Anykščiai | Nemunas | 3600 | 1928–2010 | 119 | 101 |
| 27 | Šventoji | Ukmergė | Nemunas | 5440 | 1925–2010 | 190 | 155 |
| 28 | Šešupė | Kalvarija | Nemunas | 444 | 1954–2010 | 14 | 13 |
| 29 | Nemunas | Druskininkai | Nemunas | 37,100 | 1945–2010 | 718 | 610 |
| 30 | Nemunas | Nemajūnai | Nemunas | 42,800 | 1920–2010 | 837 | 747 |
| 31 | Neris | Vilnius | Nemunas | 15,200 | 1923–2010 | 365 | 305 |

Czech Republic, Poland, the UK, Switzerland, and some others. The performed analysis (Kundzewicz et al. 2013) shows an increasing trend during the 25-year period in the number of reported floods exceeding severity and magnitude thresholds. A much debated question is whether large floods have a clear rising trend.

The slightly earlier results of a global change detection study of annual maximum river flows (Kundzewicz et al. 2005) do not support the hypothesis of a ubiquitous increase of annual maximum river flows. Nevertheless, examining 70 time series for river discharge in Europe, Kundzewicz et al. (2005) found that the overall maxima (for the whole 1961–2000 period) occurred more

frequently (46 times) in the second 20-year sub-period, 1981–2000, than in the first 20-year sub-period, 1961–1980 (24 times). Mediero et al. (2015) analysed longest hydrological series from 102 gauging stations of 25 European countries. Five geographically distinct large-scale homogeneous regions were identified: the Atlantic region, the Continental region, the Scandinavian region, the Alpine region, and the Mediterranean region. The results of this study indicate that floods tend to cluster in the Atlantic and Continental regions during the given years. A decreasing trend in the magnitude of floods was observed mainly in the Continental region in the period of 1920–1999 with 22% of the catchments revealing such a trend, as well as a

decreasing trend in the timing of floods in the Alpine region in the period of 1900–1999, with 75% of the catchments revealing this trend.

However, most of the research (such as Bormann et al. 2011; Madsen et al. 2014) until now has failed to detect significant increasing trends in available observational data of maximum river flows in particular regions. The studies that deal with historical documentary sources in central Europe (Jacobeit et al. 2003; Glaser and Stangl 2004) reveal that there have always been periods with relatively high or low temperature, increased flooding or less frequent flooding, and that particular periods with frequent floods were associated with specific circulation patterns. Medium-term increases and decreases within a range of 30–100 years were normal within the past millennium (Glaser and Stangl 2004). Such findings are particularly important, as they concern a time when mankind did not influence climate change.

Many scientists from the Baltic region countries analysed flood peculiarities as well. For example, the trend analysis in Finland revealed that the spring peak has moved to an earlier date at over one-third of the sites. However, the magnitudes of spring high flow have not changed (Korhonen and Kuusisto 2010). Observations from the last 25 years of Sweden showed that spring floods may occur approximately 1 month earlier in the northern-central part of Sweden and more frequent high flows in the south due to lesser snow accumulation in the south and at low altitudes (Arheimer and Lindström 2015). It is very important to determinate not only floods changes, but also to find the best approaches for flood investigation. Researchers in Poland are focused on methods application for floods analysis (Strupczewski et al. 2011, 2016; Kochanek et al. 2011).

In Lithuania, river floods are generally classified into two types: spring floods and flash floods. However, traditionally, existing research on floods in Lithuania as well as the other Baltic States concentrates on spring flood phenomenon (Meilutytė-Barauskienė and Kovalenkoviėnė 2007; Reihan et al. 2012). This type of flood is a significant hydrological phase of rivers in Lithuania. In the spring season, snow melting and ice jams result in higher levels of upstream water. Spring flood in Lithuanian rivers is many times greater than the average annual discharge and several hundred times greater than its minimum values. Spring runoff makes up a significant part of the annual runoff (20–90%). Research in the Baltic States (Lithuania, Latvia, Estonia) showed that the maximum discharge of spring floods decreases in most of the regions (Reihan et al. 2007, 2012; Kriauciūnienė et al. 2012; Šarauskiėnė et al. 2015) and maximum discharges are observed earlier (because of warmer winters).

On the contrary, in reviewing the literature about the extremes of Lithuanian river discharges, the definition of flash floods is uncommon [except an entire chapter in the monograph by Gailiušis et al. (2001)], although increasing river water level during this kind of floods causes a lot of problems as well. Currently, flash floods are occurring with an increasing frequency, and there is a growing concern whether this growth has a clear tendency or the threatening flash floods are just accidental events following unusually large amounts of rainfall. This indicates the need to understand behaviour of this hydrological phenomenon. The aim of this research is to determine consistent patterns of spring and flash floods in Lithuanian rivers and compare these results with East-European hydrological region. Statistical methods were used for temporal and spatial flood variability.

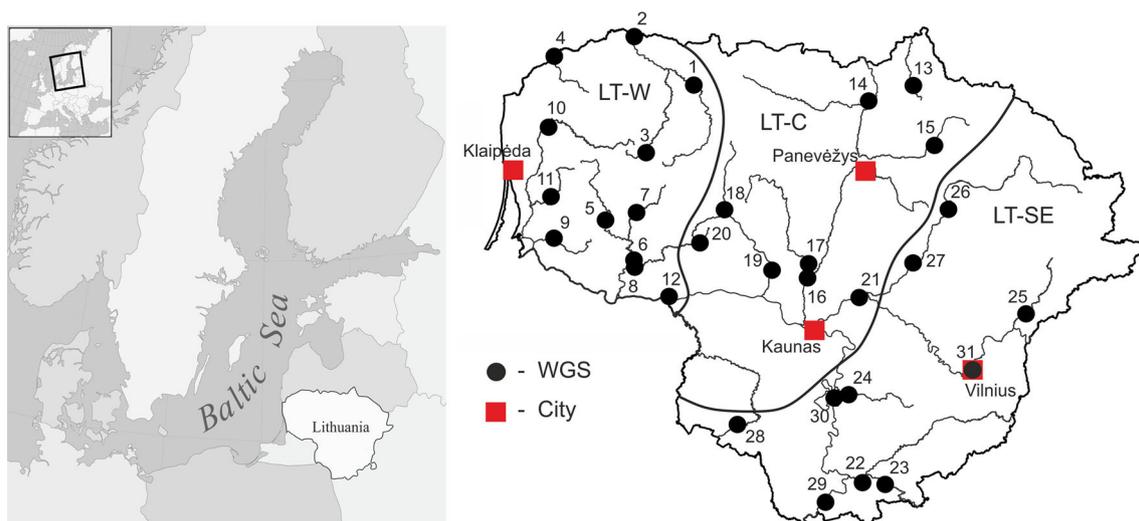


Fig. 1 Geographical position (*left*) and water gauging stations of three hydrological regions of Lithuania (*right*) (WGS numbering according by Table 1)

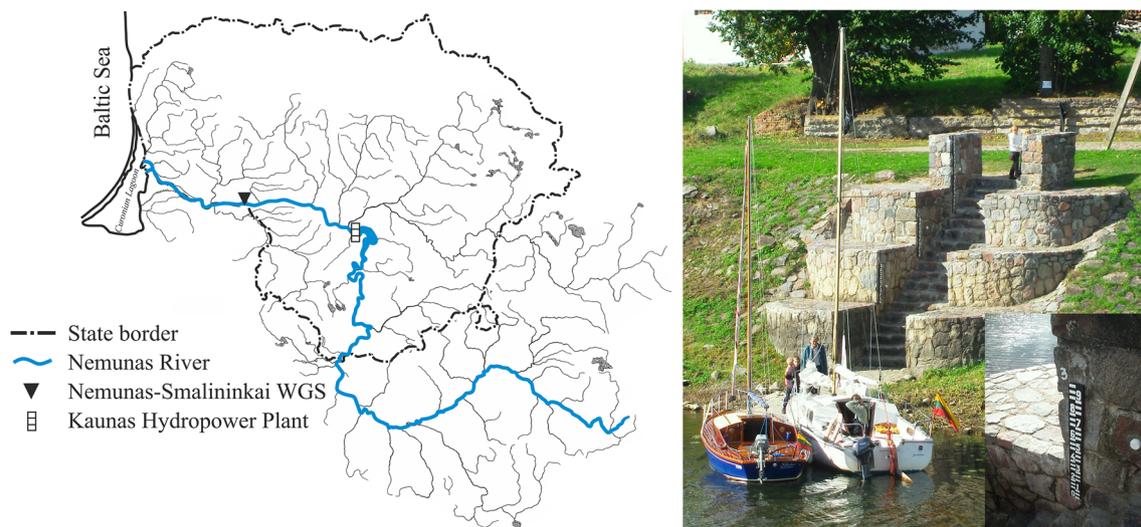


Fig. 2 The Nemunas River basin (*left*) and a photo of Smalininkai WGS at the Nemunas River (*right*)

Research object and data

Lithuania (total area of 65,200 km²) has over 22,000 rivers with a total length of more than 37,000 km. The annual river runoff fluctuates from 4.2 to 14.0 l/(s km²). The entire river runoff flows into the Baltic Sea: 72% of all river runoff via the Nemunas River, 13% via the Lielupe River, 8% via the Venta River, 3% via the Dauguva River, and 4% via other rivers.

The Nemunas River is a major Lithuanian river. The total length of the Nemunas is 937 km, while the basin area covers 98,200, 46,600 km² of which belongs to Lithuania (comprising 72% of the Lithuanian territory). This river delivers on average 21.8 km³ of water per year into the Baltic Sea through the Curonian Lagoon.

Climatic factors, soil structure, geology, type of land use and geomorphology affect hydrological regime of Lithuanian rivers. According to hydrological regime and river feeding type, Lithuanian territory is divided into three regions (Fig. 1): Western (1), Central (2), and Southeastern (3).

Long-term hydrological data sets analysed in this research were obtained from the Lithuanian Hydrometeorological Service. Previously, river observation data had been available from more than 150 water gauging stations (WGS), while the present hydrological station network consists of 50 WGS. Thirty-one water gauging stations were used in this research. They cover the entire territory of Lithuania (Table 1). The data of spring and flash floods were analysed separately. For the analysis of spring floods, the maximum discharges of the period from 1 March to 30 May were used, whereas for flash floods from June 1 to November 30. The maximum discharge data of four

periods (1922–2013, 1941–2013, 1961–2013, and 1991–2013) were compared with the data of the reference period (1961–1990).

The discharge data of the Nemunas River at Smalininkai WGS (1812–2013) were used in this research as well (Fig. 2). This data set is the longest in the country and one of the oldest in Europe (Gailiūšis et al. 2011). Such historical information is significant in the assessment of long-term variation of floods in Lithuania. Before the construction of Kaunas Hydropower plant (HPP) on the Nemunas River in 1960, the amplitude of water level fluctuation was about 10 m in the middle reach and 2–3 m in the lower reach. The construction of the HPP and the creation of the Kaunas reservoir (surface area—63.5 km²) reduced flooding downstream to Kaunas city.

Methods

Parametric and non-parametric methods (Zhang et al. 2006; Chen et al. 2007) are widely used for the detection of significance of trends in hydro-climatological time series. The Mann–Kendall (MK) and Spearman-Rho (SR) are non-parametric methods which have been applied in a large number of scientific researches (Yue et al. 2002; Kahya and Kalayci 2004; Li et al. 2008; Yaning et al. 2009). These two trend detection methods were chosen to investigate the variability of spring and flash flood data in Lithuania.

Mann–Kendall test

This non-parametric test was created by Mann (1945) and Kendall (1975) and used for detection of linear or non-

linear trends (Hisdal et al. 2001; Wu et al. 2008). In this test, the null (H_0) and alternative hypotheses (H) are equal to the non-existence and existence of a trend in the time series of the observational data, respectively. The related equations for calculating the MK test statistic S and the standardized test statistic Z_{MK} are as follows (Shadmani et al. 2012):

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{sgn}(X_j - X_i), \tag{1}$$

$$\text{sgn}(X_j - X_i) = \begin{cases} +1 & \text{if } (X_j - X_i) > 0 \\ 0 & \text{if } (X_j - X_i) = 0 \\ -1 & \text{if } (X_j - X_i) < 0 \end{cases} \tag{2}$$

$$\text{var}(S) = \frac{1}{18} \left[n(n-1)(2n+5) - \sum_{p=1}^q t_p(t_p-1)(2t_p+5) \right] \tag{3}$$

$$Z_{MK} = \begin{cases} \frac{S-1}{\sqrt{\text{var}(S)}} & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ \frac{S+1}{\sqrt{\text{var}(S)}} & \text{if } S < 0 \end{cases} \tag{4}$$

where X_i and X_j are the sequential data values of the time series in the years i and j , n is the length of the time series, t_p is the number of ties for the p th value, and q is the number of tied values. Positive values of Z_{MK} indicate increasing trends, while negative Z_{MK} values indicate

decreasing trends in the time series. When $|Z_{MK}| > Z_{1-\alpha/2}$, the null hypothesis is rejected and a significant trend exists in the time series. $Z_{1-\alpha/2}$ is the critical value of Z from the standard normal table, for 5% significance level the value of $Z_{1-\alpha/2}$ is 1.96 (Shadmani et al. 2012).

Spearman-Rho test

This test is non-parametric and very similar to the MK test. The SR test is a method applied to identify the lack of trends (Dahmen and Hall 1990; Tonkaz et al. 2007). In this test, the null hypothesis (H_0) is that all the data in the time series are independent and identically distributed, while the alternative hypothesis (H_1) is that increasing or decreasing trends exist (Yue et al. 2002). The SR test statistic D and the standardized test statistic Z_{SR} are expressed as follows (Shadmani et al. 2012):

$$D = 1 - \frac{6 \sum_{i=1}^n (R_i - 1)^2}{n(n^2 - 1)} \tag{5}$$

$$Z_{SR} = D \sqrt{\frac{n-2}{1-D^2}}, \tag{6}$$

where R_i is the rank of i th observation X_i in the time series, and n is the length of the time series. Positive values of Z_{SR} indicate upward trends, while negative Z_{SR} indicates downward trends in the time series. When $|Z_{SR}| < t_{(n-2, 1-\alpha/2)}$, the null hypothesis is rejected, and a significant trend exists in the time series. $t_{(n-2, 1-\alpha/2)}$ is the

Fig. 3 River feeding sources in different hydrological regions (LT-W Western Lithuania, LT-C Central Lithuania, LT-SE Southeastern Lithuania)

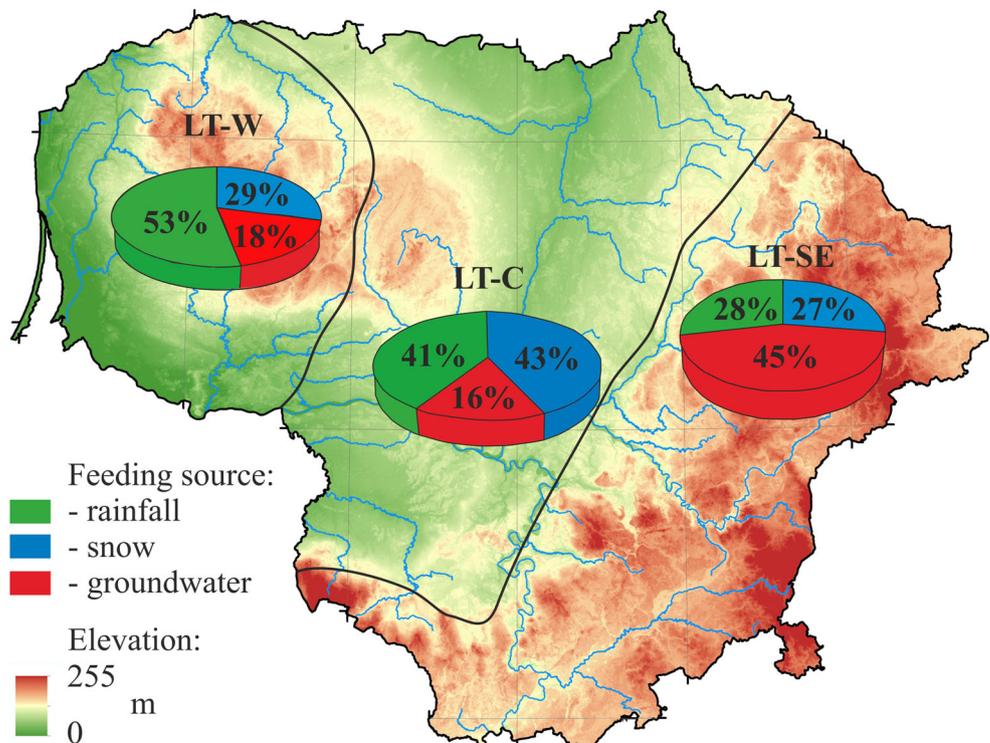


Fig. 4 Variability of spring and flash flood data in the different hydrological regions (the Minija River from LT-W, the Neris River from LT-C, the Merkys River from LT-SE)

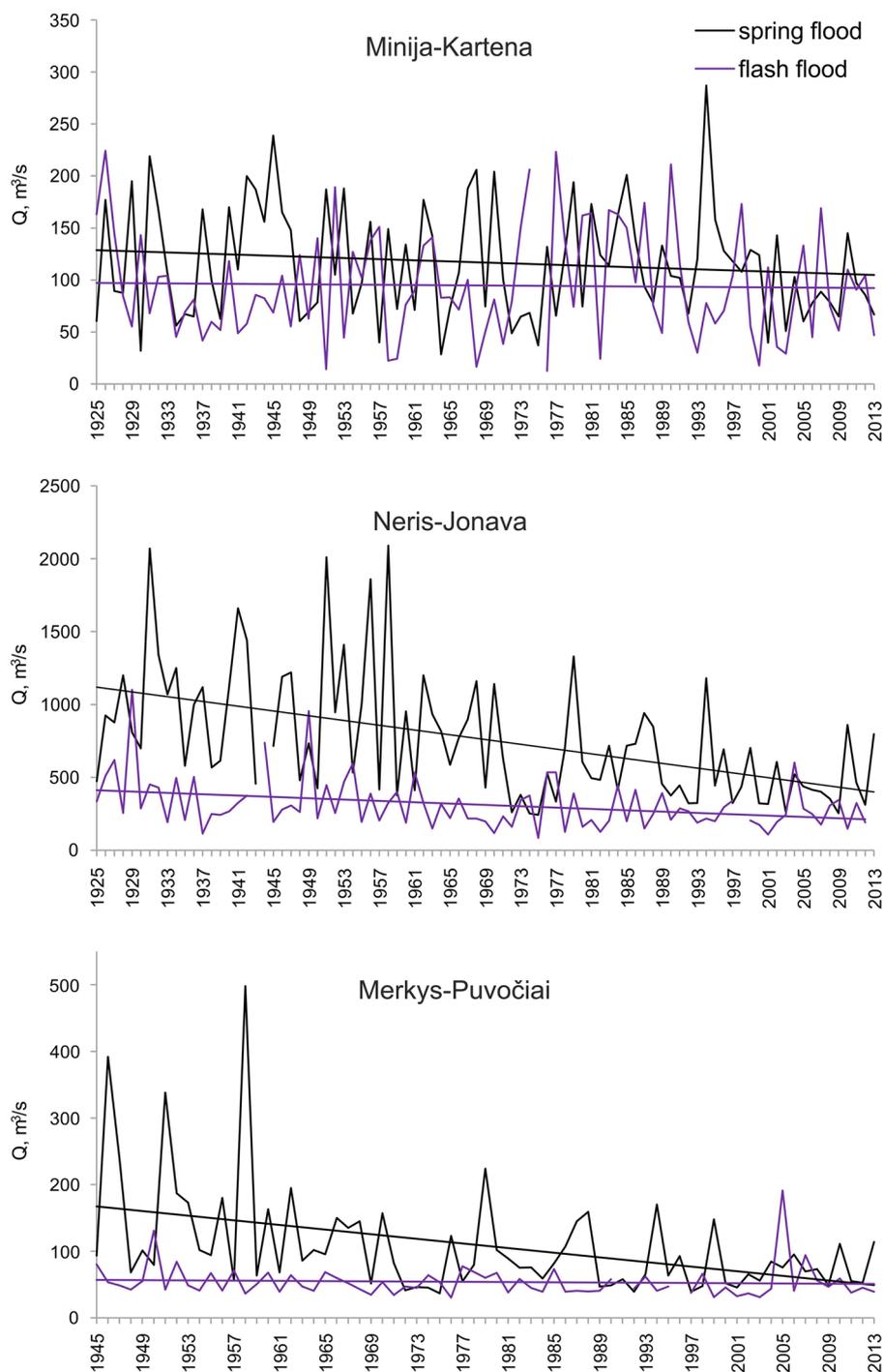


Table 2 Example of MK and SR tests for some rivers (flash flood)

| | Minija-Kartena | Jūra-Tauragė | Nemunas-Smalininkai |
|----|----------------|--------------|---------------------|
| MK | -2.14 | -1.19 | -2.47 |
| SR | -2.12 | -1.17 | -2.45 |

critical value of t from the t Student table, for 5% significance level (Shadmani et al. 2012).

Homogeneity tests were performed for the extreme data series of all hydrological stations. A test with a 5% significance level (recommended by WMO) was applied for

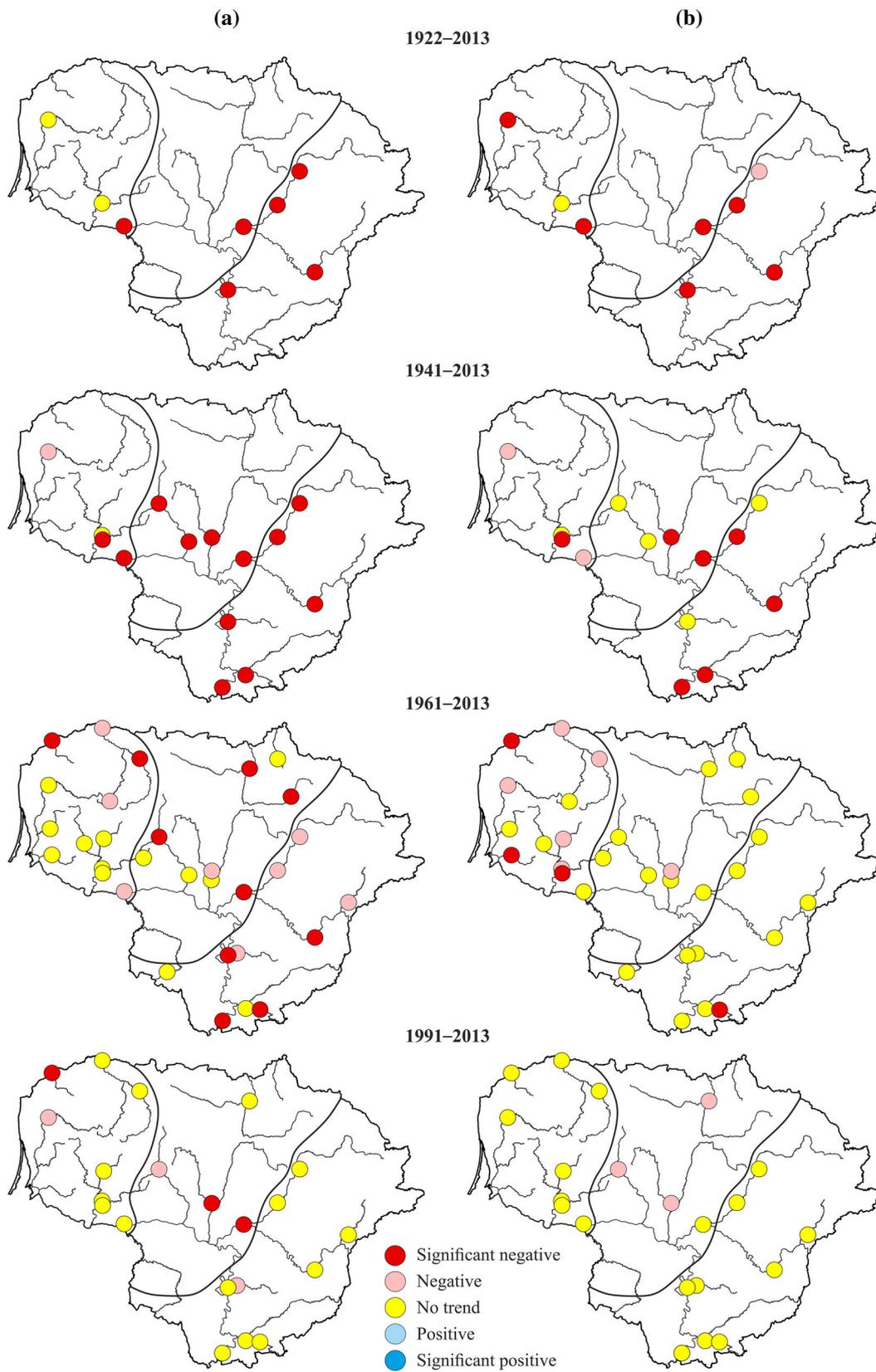


Fig. 5 Spatial distribution of trends of maximum discharges of spring (a) and flash (b) floods in different periods

the data series trend analysis. In addition, positive and negative trends at 30% level were applied.

Flood anomalies

Anomalies of maximum discharges of spring and flash floods (expressed in %) were calculated by dividing each member of series by the mean values of the reference period (1961–1990). Estimation of anomalies was made for decadal periods (1931–1940, 1941–1950, etc.) for the separate hydrological regions.

Integrated curves

In each hydrological region, periodic fluctuations and synchronical changes of the spring and flash floods were described by integrated curves. This curve ($\sum_{i=1}^n (k_i - 1)$) is the sum of variations of modular coefficients from the average value. Modular coefficient is $K_i = Q_i / \bar{Q}$, where Q_i is the flood discharge in year i , and \bar{Q} is the mean flood discharge for the entire period of observation.

Results

General characteristics of spring and flash floods in three different hydrological regions

The spring and flash flood patterns vary depending on hydrological regions. One of the main reasons for this behaviour is different sources of river feeding (Fig. 3). A marine type of climate dominates in the Western region, with the largest amount of precipitation, the highest winter temperature and the least number of days with snow cover (Kriaučiūnienė et al. 2012). Precipitation is the major source of river feeding in this region, exceeding 53%. The other sources include snowmelt (18%) and groundwater (29%). Due to frequent thaws in wintertime, rivers here often have “winter floods”, some of which are greater than spring floods. The continental type of climate is characteristic for Southeastern Lithuania: the snow cover here has the longest duration and the winters are the coldest. Sub-surface feeding dominates in the rivers of this region (45%). Permeable sandy soils, which are widespread, effectively absorb snowmelt and gradually release it later, supplying rivers in the low water period. The type of river feeding in Central Lithuania is mixed; the rivers get water mostly from two main sources: rainfall and snowmelt. A very irregular distribution of discharges during the year is the major feature of the rivers in this region.

Spring (Q_{Spring}) and flash (Q_{Flash}) flood data of the reference period (1961–1990) (Table 1) verify the regionalization of Lithuanian rivers. Q_{Flash} in LT-W are larger than

Q_{Spring} (in 67% of WGS); on the contrary, in both LT-C and LT-SE, Q_{Spring} tends to be greater than Q_{Flash} (in 89 and 80% of WGS, respectively).

The analysis of multi-annual changes in flood discharges of individual rivers confirms the presence of different flood models in three hydrological regions. The different variation of flood data of three “typical” rivers is presented in Fig. 4.

Temporal and spatial variation of spring and flash floods

The trend analysis of maximum discharges of spring and flash floods was performed using the MK and SR tests. Both tests revealed very similar results (Table 2).

The trends were estimated for four periods (1922–2013, 1941–2013, 1961–2013, and 1991–2013) (Fig. 5). No positive tendencies in data of both types of floods were detected during the entire investigated period. Trends of spring flood maximum discharges (Fig. 5a) of the longest period (1922–2013) showed decreasing tendency in the Central and Southeastern regions (two WGS had no trends, while six WGS had negative trends). Only negative trends were identified in spring flood data from 13 WGS for the period of 1941–2013, except for one station with no trend in the Western region of Lithuania. Behaviour of spring flood maximum discharges in the period of 1961–2013 was marginally different: significant negative or negative trends were found in the data of most rivers, except for a few WGS which had no trends (mainly in the Western region). The tests showed no significant trends for the period of 1991–2013 in the entire territory of Lithuania. Spring floods become smaller and earlier evidently due to the increasing air temperatures in winter season (Reihan et al. 2007, 2012).

The trends of flash floods in the periods of 1922–2013 and 1941–2013 were negative in all WGS, except for some WGS with no trends (Fig. 5b). Gradual decrease in flash flood data was observed only for the Western region in the period of 1961–2013, while the WGS data of the Central and Southeastern regions had no noticeable tendencies. In the period of 1991–2013, floods remained steady in the entire territory of Lithuania, except three WGS with negative trends in Central Lithuania.

In further analysis, Z values that indicate the strength and direction of trends were applied. Figure 6 shows prevailing negative Z values in the period of 1961–2013, i.e. the detected trends are significant negative or negative. In 1991–2013, some positive Z values emerge, but negative trends still dominate.

The calculated integrated curves of flood data are presented in Fig. 7. Upward and downward slopes in integrated curves indicate the periods in which the floods have increasing or decreasing tendencies. In the rivers of all

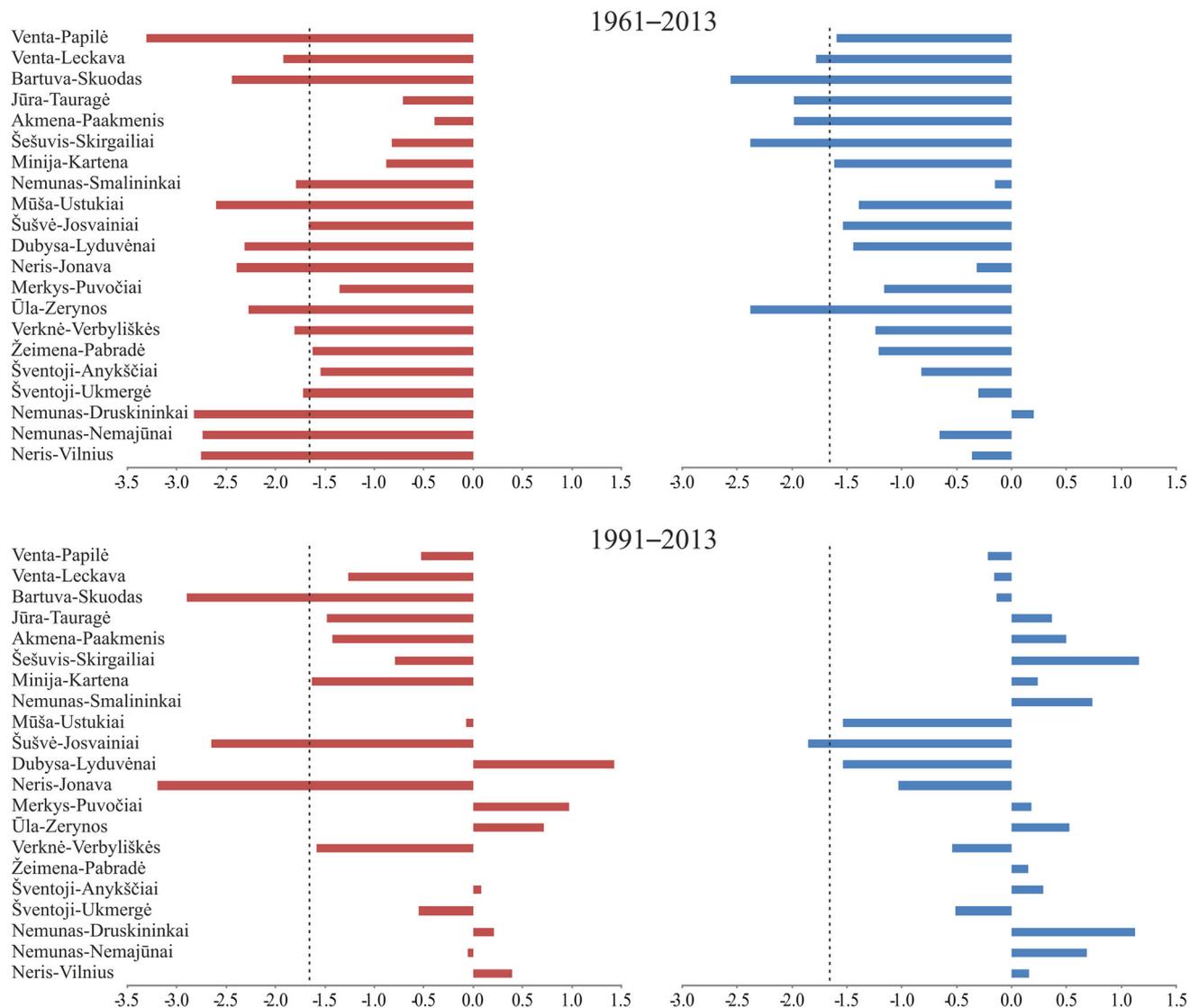


Fig. 6 The trends (for 1961–2013 and 1991–2013) of spring (a) and flash (b) floods of investigated Lithuanian rivers according to the test criterion (Z equals 1.96 at 5% significance level)

hydrological regions, the spring and flash floods tended to decline in the last decades.

Analysis of regional anomalies of decadal spring and flash floods (Table 3) indicates that in the last two decades, all floods were smaller than their mean values in the reference period. The lowest decrease of flash flood anomalies was identified in the Southeastern region. The largest and only negative anomalies, i.e. significantly diminished flood discharges, were found in the last period (2001–2010) in all hydrological regions.

Spring and flash flood differences according to long-term data of the Nemunas River (Smalininkai WGS)

The hydrological value of this long historical data set was already highlighted. The majority of earlier described and

investigated rivers from the Central and Southeastern hydrological regions are tributaries of the Nemunas River and belong to its basin.

Correlation coefficients of maximum discharges between the Nemunas River (Smalininkai WGS) and the rivers from the Southeastern region are 0.73–0.87, while the coefficients between the Nemunas River (Smalininkai WGS) and the rivers from the Central region are 0.60–0.90.

In the Nemunas River, over the long observation period spring floods decreased, whereas flash flood behaviour did not have any clear tendency (Fig. 8). The year 1960 marks significant changes in the runoff regime of the Nemunas River, since Kaunas Hydropower Plant on the river was built, and a large dam (Kaunas reservoir) was formed. Figure 8 indicates that until 1960, maximum discharges of spring and flash floods had a growing tendency, whereas

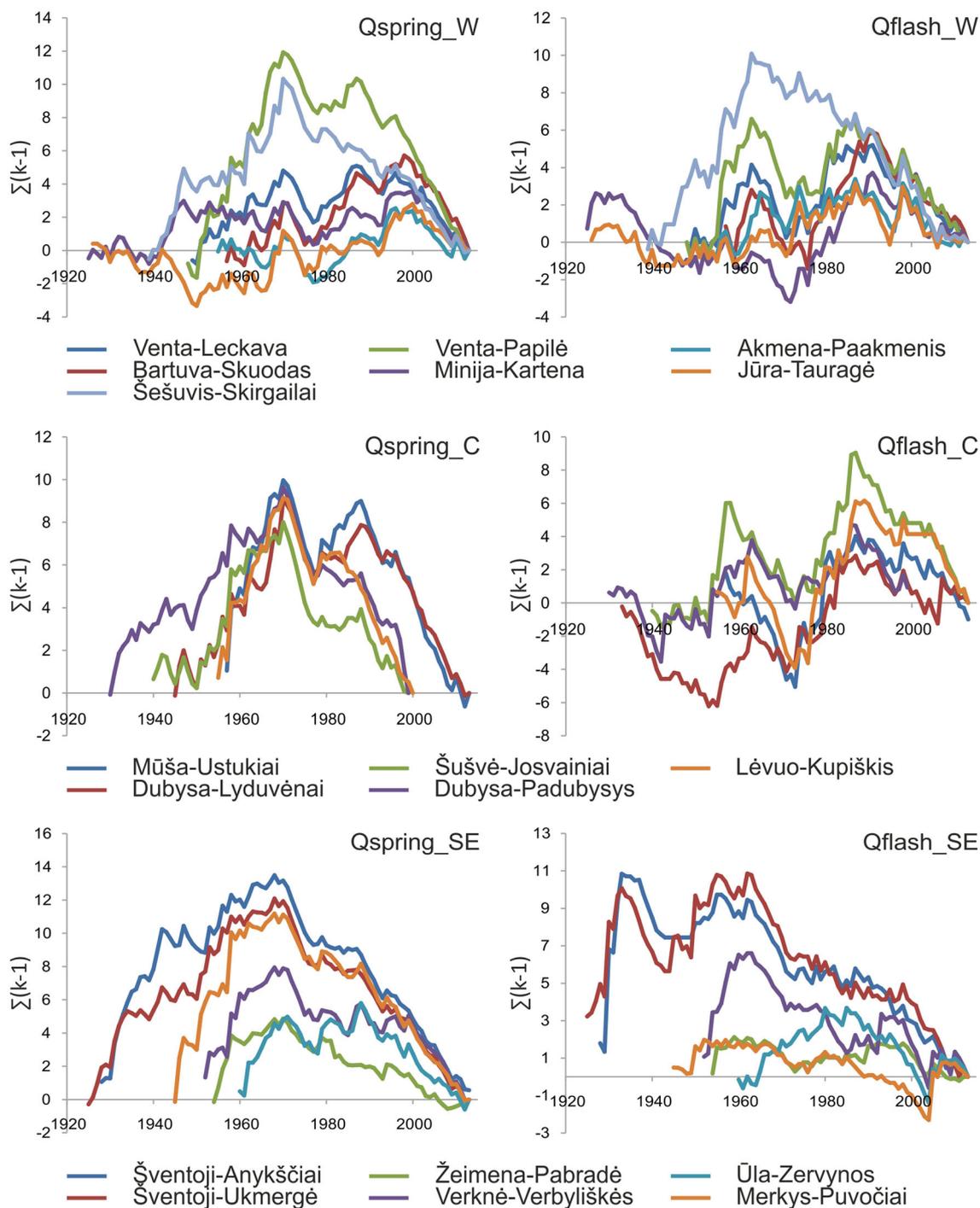


Fig. 7 Integrated curves of spring and flash floods in different regions (with 5 years moving average)

after the construction of the HPP, the natural variation of the Nemunas River runoff was disturbed. The river became safe from dangerous floods, but this human intervention resulted in irreversible changes of maximum discharges—they considerably decreased.

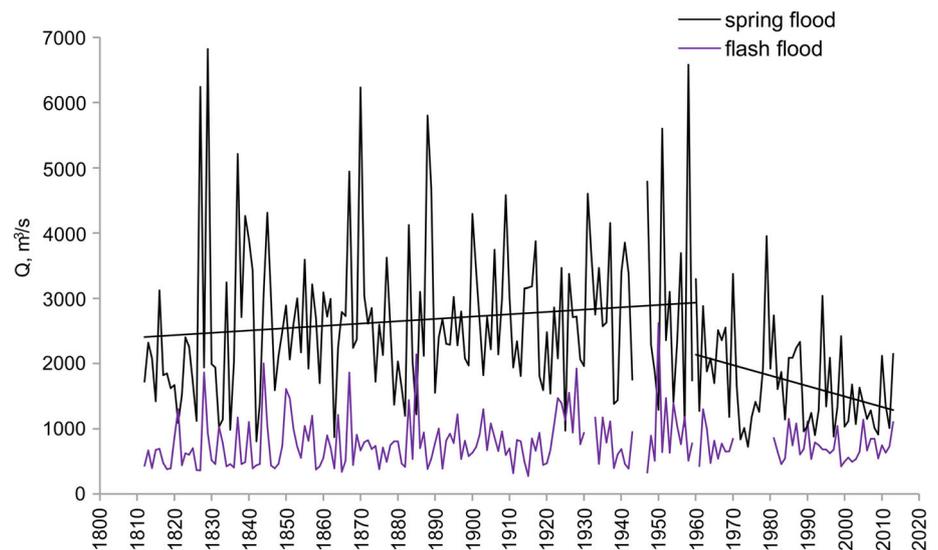
The analysis of spring and flash flood anomalies (Fig. 9) confirms that floods decreased in the last decades.

Anthropogenic activity (installation of Kaunas HPP) and climate change (warmer winters, thin snow cover, etc.) can be considered as two main reasons for the spring flood decrease.

A comparison of decadal data of maximum discharges of spring and flash floods of the Nemunas River at Smailininkai WGS with the average of the entire period

Table 3 Average anomalies of spring and flash floods

| Period | Average anomaly of spring flood, % | | | Average anomaly of flash flood, % | | |
|-----------|------------------------------------|-------|-------|-----------------------------------|-------|-------|
| | LT-W | LT-C | LT-SE | LT-W | LT-C | LT-SE |
| 1931–1940 | −4.8 | −28.6 | 64.5 | −24.6 | −51.6 | 17.1 |
| 1941–1950 | 6.4 | −0.4 | 43.1 | 1.0 | −19.0 | 25.0 |
| 1951–1960 | 14.7 | 59.2 | 80.8 | 6.8 | 6.9 | 31.6 |
| 1961–1970 | 24.6 | 36.2 | 21.6 | −10.3 | −36.0 | −5.2 |
| 1971–1980 | −27.6 | −37.1 | −20.6 | 4.5 | 10.7 | 2.4 |
| 1981–1990 | 2.9 | 1.4 | −6.8 | 5.1 | 18.9 | −0.6 |
| 1991–2000 | −3.5 | −31.5 | −14.7 | −25.3 | −35.6 | −10.0 |
| 2001–2010 | −33.4 | −45.2 | −26.8 | −25.5 | −34.5 | −1.9 |

Fig. 8 Trend of maximum discharges of spring and flash floods in the Nemunas River at Smalininkai WGS in the period 1812–2013

(1812–2013) was made (Table 4). The period of 1921–1930 distinguishes by considerably large flash floods. The largest increase (30%) of spring flood maximum discharge was in the period of 1951–1960 (the largest spring flood due to snow melting was observed in 1958), while the largest decrease (44%) was in the last decade (2001–2010).

Discussion

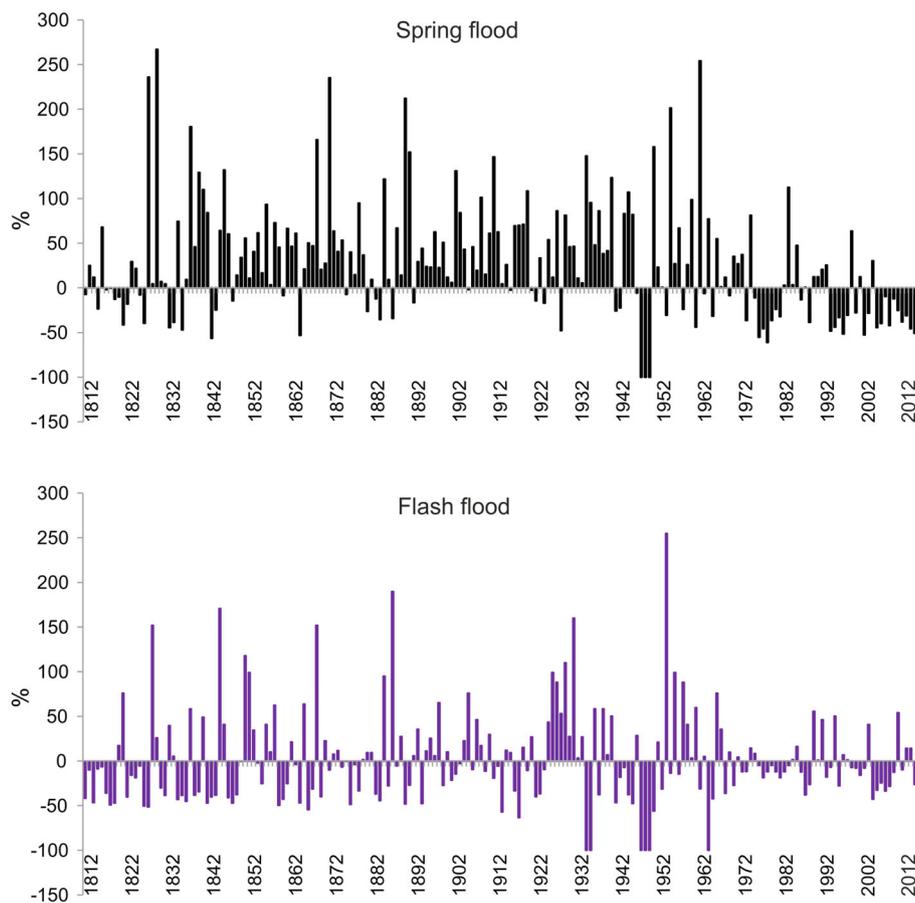
Lithuania, like the other Baltic States (Reihan et al. 2012), is divided into three hydrological regions: marine (LT-W), transitional (LT-C) and continental (LT-SE). In three hydrological regions, river floods depend on different pattern of river feeding. In Southeastern and Central regions of Lithuania, flash floods do not exceed the level of spring floods. An exception is the rivers in the Western region, where flash floods usually are larger than spring floods. These patterns could be explained by different river feeding sources: in the Western region of Lithuania, rivers are supplemented by rainfall, in the Southeastern region of

Lithuania by groundwater, while in Central Lithuania by snowmelt and rainfall.

Trend analysis of maximum discharges of spring floods showed gradual decrease during all analysed periods. Other researchers confirm the same tendency of spring floods. In the Baltic States, spring floods generally declined from 1922 to 2010 (Reihan et al. 2012; Šarauskienė et al. 2015). In Poland, spring floods decreased in the most of Poland's major rivers as well (Kaczmarek 2003). The smaller spring floods and earlier start of river flooding are evidently caused by increased air temperature in the winter season and earlier snowmelt as a consequence (Reihan et al. 2007; Kriauciuniene et al. 2012; Thorsteinsson and Björnsson 2012).

Significant negative (especially in LT-W), negative and insignificant trends of maximum discharges of flash floods were identified in all hydrological regions of Lithuania. Only insignificant trends were detected in the period of 1991–2013. The various tendencies of flash flood behaviour were identified in other European rivers. A general decreasing trend in flash floods was found in the Vltava and

Fig. 9 Anomalies of maximum discharges of spring and flash floods in the Nemunas River (Smalininkai WGS) in the period of 1812–2013



the Elbe rivers (Czech Republic) during the twentieth century (Yiou et al. 2006). In Slovakia (Solín 2008), the analysis of discharge of flash floods in small catchments (5–150 km²) showed an increasing tendency in the last two decades. In Catalonia, flash floods also increased during the period of 1982–2007 (Llasat et al. 2010); however, the floods assumed decreasing tendency when the data set was extended till 2010.

Conclusion

During all analysed periods (1922–2013, 1941–2013, 1961–2013, and 1991–2013), significant negative and negative trends dominated in maximum discharges of spring floods. These results confirm decrease of spring floods in Lithuania.

Significant negative, negative and insignificant trends of flash floods were detected during all analysed periods. The findings of two most recent periods (when the most observations data are available) revealed no clear tendencies in flash flood behaviour in the analysed rivers.

The analysis of maximum discharges (1812–2013) of the Nemunas River proved that both spring and flash floods

Table 4 Decadal period averages of maximum discharges of spring and flash floods in the Nemunas River (Smalininkai WGS)

| Period | Maximum discharges of spring flood, m ³ /s (deviation, %) | Maximum discharges of flash flood, m ³ /s (deviation, %) |
|-----------|--|---|
| 1812–1820 | 1958 (–19) | 551 (–29) |
| 1821–1830 | 2709 (13) | 769 (–1) |
| 1831–1840 | 2645 (10) | 676 (–13) |
| 1841–1850 | 2506 (4) | 796 (2) |
| 1851–1860 | 2605 (8) | 814 (4) |
| 1861–1870 | 3013 (25) | 793 (2) |
| 1871–1880 | 2451 (2) | 699 (–10) |
| 1881–1890 | 2745 (14) | 833 (7) |
| 1891–1900 | 2611 (9) | 769 (–1) |
| 1901–1910 | 2930 (22) | 844 (8) |
| 1911–1920 | 2533 (5) | 608 (–22) |
| 1921–1930 | 2373 (–1) | 1183 (52) |
| 1931–1940 | 3001 (25) | 798 (2) |
| 1941–1950 | 2746 (14) | 877 (13) |
| 1951–1960 | 3115 (30) | 933 (20) |
| 1961–1970 | 2178 (–9) | 747 (–4) |
| 1971–1980 | 1585 (–34) | 702 (–10) |
| 1981–1990 | 1811 (–25) | 789 (1) |
| 1991–2000 | 1556 (–35) | 670 (–14) |
| 2001–2010 | 1336 (–44) | 701 (–10) |

in Lithuania were getting smaller. The most significant decrease of maximum discharges of spring and flash floods was observed after 1960 (hydro power plant installation).

Acknowledgements This study is supported by National Research Programme “Sustainability of agro-, forest and water ecosystems” according to the Project “Impact Assessment of Climate Change and Other Abiotic Environmental Factors on Aquatic Ecosystems” (SIT-11/2015).

References

- Arheimer B, Lindström G (2015) Climate impact on floods: changes in high flows in Sweden in the past and the future (1911–2100). *Hydrol Earth Syst Sci* 19:771–784. doi:[10.5194/hess-19-771-2015](https://doi.org/10.5194/hess-19-771-2015)
- Barredo JI (2007) Major flood disasters in Europe: 1950–2005. *Nat Hazards* 42(1):125–148. doi:[10.1007/s11069-006-9065-2](https://doi.org/10.1007/s11069-006-9065-2)
- Bormann H, Pinter N, Elfert S (2011) Hydrological signatures of flood trends on German rivers: flood frequencies, flood heights and specific stages. *J Hydrol* 404:50–66. doi:[10.1016/j.jhydrol.2011.04.019](https://doi.org/10.1016/j.jhydrol.2011.04.019)
- Brazdil R, Glaser R, Pfister C, Stangl H (2002) Floods in Europe—a look into the past. *PAGES News* 10(3):21–23
- Chen H, Guo SL, Xu CY, Singh VP (2007) Historical temporal trends of hydro-climatic variables and runoff response to climate variability and their relevance in water resource management in the Hanjiang basin. *J Hydrol* 344(3–4):171–184. doi:[10.1016/j.jhydrol.2007.06.034](https://doi.org/10.1016/j.jhydrol.2007.06.034)
- Dahmen ER, Hall MJ (1990) Screening of hydrological data: tests for stationarity and relative consistency. ILRI Publication No. 49, Wageningen
- Gailiūšis B, Jablonskis J, Kovalenkoviėnė M (2001) The Lithuanian rivers. Hydrography and runoff. LEI, Kaunas (**in Lithuanian**)
- Gailiūšis B, Kriauciūnienė J, Jakimavičius D, Šarauskiėnė D (2011) The variability of long-term runoff series in the Baltic Sea drainage basin. *Baltica* 24(1):45–54
- Glaser R, Stangl H (2004) Climate and floods in Central Europe since AD 1000: data, methods, results and consequences. *Surv Geophys* 25:485–510. doi:[10.1007/s10712-004-6201-y](https://doi.org/10.1007/s10712-004-6201-y)
- Han D (2001) Flood risk assessment and management. University of Bristol, Bristol
- Hisdal H, Stahl K, Tallaksen LM, Demuth S (2001) Have streamflow droughts in Europe become more severe or frequent. *Int J Climatol* 21:317–333. doi:[10.1002/joc.619](https://doi.org/10.1002/joc.619)
- Jacobbeit J, Glaser R, Luterbacher J, Wanner H (2003) Links between flood events in Central Europe since AD 1500 and large-scale atmospheric circulation modes. *Geophys Res Lett* 30(4):211–214. doi:[10.1029/2002GL016433](https://doi.org/10.1029/2002GL016433)
- Kaczmarek Z (2003) The impact of climate variability on flood risk in Poland. *Risk Anal* 23:559–566. doi:[10.1111/1539-6924.00336](https://doi.org/10.1111/1539-6924.00336)
- Kahya E, Kalaycı S (2004) Trend analysis of streamflow in Turkey. *J Hydrol* 289:128–144. doi:[10.1016/j.jhydrol.2003.11.006](https://doi.org/10.1016/j.jhydrol.2003.11.006)
- Kendall MG (1975) Rank correlation methods. Griffin, London
- Kochanek K, Strupczewski WG, Bogdanowicz E (2011) On seasonal approach to flood frequency modelling. Part II: Flood frequency analysis of Polish rivers. *Hydrol Process* 26(5):717–730. doi:[10.1002/hyp.8178](https://doi.org/10.1002/hyp.8178)
- Korhonen J, Kuusisto E (2010) Long-term changes in the discharge regime in Finland. *Hydrol Res* 41(3–4):253–268. doi:[10.2166/nh.2010.112](https://doi.org/10.2166/nh.2010.112)
- Kriauciūnienė J, Meilutytė-Barauskiėnė D, Reihan A, Koltsova T, Lizuma L, Šarauskiėnė D (2012) Variability in temperature, precipitation and river discharge in Baltic States. *Boreal Environ Res* 17:150–162
- Kundzewicz ZW, Graczyk D, Maurer T, Pińskwar I, Radziejewski M, Svensson C, Szwed M (2005) Trend detection in river flow series: 1. Annual maximum flow. *Hydrol Sci J* 50(5):797–810. doi:[10.1623/hysj.2005.50.5.797](https://doi.org/10.1623/hysj.2005.50.5.797)
- Kundzewicz ZW, Pinskiwar I, Brakenridge GR (2013) Large floods in Europe, 1985–2009. *Hydrol Sci J* 58(1):1–7. doi:[10.1080/02626667.2012.745082](https://doi.org/10.1080/02626667.2012.745082)
- Li ZL, Xu ZX, Li JY, Li ZJ (2008) Shift trend and step changes for runoff time series in the Shiyang River Basin, Northwest China. *Hydrol Process* 22(23):4639–4646. doi:[10.1002/hyp.7127](https://doi.org/10.1002/hyp.7127)
- Llasat MC, Llasat-Botija M, Rodriguez A, Lindbergh S (2010) Flash floods in Catalonia: are current situation. *Adv Geosci* 26:105–111. doi:[10.5194/adgeo-26-105-2010](https://doi.org/10.5194/adgeo-26-105-2010)
- Madsen H, Lawrence D, Lang M, Martinkova M, Kjeldsen TR (2014) Review of trend analysis and climate change projections of extreme precipitation and floods in Europe. *J Hydrol* 519(D):3634–3650. doi:[10.1016/j.jhydrol.2014.11.003](https://doi.org/10.1016/j.jhydrol.2014.11.003)
- Mann HB (1945) Nonparametric tests against trend. *Econometrica* 13:245–259
- Mediero L, Kjeldsen TR, Macdonald N, Kohnova S, Merz B, Vorogushyn S, Wilson D, Alburquerque T, Blöschl G, Bogdanowicz E, Castellarin A, Hall J, Kobold M, Kriauciuniene J, Lang M, Madsen H, Onusluogul G, Perdigão RAP, Roald LA, Salinas JL, Toumazis AD, Veijalainen N, Öðinn Þ (2015) Identification of coherent flood regions across Europe by using the longest stream flow records. *J Hydrol* 528:341–360. doi:[10.1016/j.jhydrol.2015.06.016](https://doi.org/10.1016/j.jhydrol.2015.06.016)
- Meilutytė-Barauskiėnė D, Kovalenkoviėnė M (2007) Change of spring flood parameters in Lithuanian rivers. *Energetika* 2:26–33
- Reihan A, Koltsova T, Kriauciūnienė J, Lizuma L, Meilutytė-Barauskiėnė D (2007) Changes in water discharges of the Baltic States rivers in the 20th century and its relation to climate change. *Nord Hydrol* 38(4–5):401–412
- Reihan A, Kriauciūnienė J, Meilutyte-Barauskiėnė D, Kolcova T (2012) Temporal variation of spring flood in rivers of the Baltic States. *Hydrol Res* 43(4):301–314. doi:[10.2166/nh.2012.141](https://doi.org/10.2166/nh.2012.141)
- Šarauskiėnė D, Kriauciūnienė J, Reihan A, Klavis M (2015) Flood pattern changes in the rivers of the Baltic countries. *J Environ Eng Landsc Manag* 23(2):28–38. doi:[10.3846/16486897.2014.937438](https://doi.org/10.3846/16486897.2014.937438)
- Shadmani M, Marofi S, Roknian M (2012) Trend analysis in reference evapotranspiration using Mann–Kendall and Spearman’s Rho tests arid regions of Iran. *Water Resour Manag* 26:211–224. doi:[10.1007/s11269-011-9913-z](https://doi.org/10.1007/s11269-011-9913-z)
- Solín L (2008) Analysis of floods occurrence in Slovakia in the period 1996–2006. *J Hydrol Hydromech* 56:95–115
- Strupczewski WG, Kochanek K, Bogdanowicz E, Markiewicz I (2011) On seasonal approach to flood frequency modelling. Part I: Two-component distribution revisited. *Hydrol Process* 26(5):705–716. doi:[10.1002/hyp.8179](https://doi.org/10.1002/hyp.8179)
- Strupczewski WG, Kochanek K, Bogdanowicz E, Markiewicz I, Feluch W (2016) Comparison of two nonstationary flood frequency analysis methods within the context of the variable regime in the representative Polish rivers. *Acta Geophys* 64(1):206–236. doi:[10.1515/acgeo-2015-0070](https://doi.org/10.1515/acgeo-2015-0070)
- Thorsteinsson T, Björnsson H (2012) Climate change and energy systems, impacts, risks and adaptation in the Nordic and Baltic countries. Nordic Council of Ministers, Copenhagen, Denmark
- Tonkaz T, Çetin M, Kâzım T (2007) The impact of water resources development projects on water vapour pressure trends in a semi-arid region, Turkey. *Clim Chang* 82(1):195–209. doi:[10.1007/s10584-006-9160-0](https://doi.org/10.1007/s10584-006-9160-0)
- Wu H, Soh LK, Samal A, Chen XH (2008) Trend analysis of streamflow drought events in Nebraska. *Water Resour Manag* 22(2):145–164. doi:[10.1007/s11269-006-9148-6](https://doi.org/10.1007/s11269-006-9148-6)

- Yaning C, Changchun X, Xingming H, Weihong L, Yapeng C, Chenggang Z, Zhaoxia Y (2009) Fifty-year climate change and its effect on annual runoff in the Tarim River Basin, China. *Quat Int* 208(1–2):53–61
- Yiou P, Ribereau P, Naveau P, Nogaj M, Brázdil R (2006) Statistical analysis of floods in Bohemia (Czech Republic) since 1825. *Hydrol Sci J* 51(5):930–945. doi:[10.1623/hysj.51.5.930](https://doi.org/10.1623/hysj.51.5.930)
- Yue S, Pilon P, Cavadas G (2002) Power of the Mann–Kendall and Spearman’s rho test for detecting monotonic trends in hydrological series. *J Hydrol* 259:254–271. doi:[10.1016/S0022-1694\(01\)00594-7](https://doi.org/10.1016/S0022-1694(01)00594-7)
- Zhang Q, Liu C, Xu CY, Xu YP, Jiang T (2006) Observed trends of water level and streamflow during past 130 years in the Yangtze River basin, China. *J Hydrol* 324(1–4):255–265. doi:[10.1016/j.jhydrol.2005.09.023](https://doi.org/10.1016/j.jhydrol.2005.09.023)

Improved petrographic-coded model and its evaluation to determine a thermal conductivity log

Nina Gegenhuber¹ · Markus Kienler¹

Received: 9 January 2017 / Accepted: 16 January 2017 / Published online: 30 January 2017
© The Author(s) 2017. This article is published with open access at Springerlink.com

Abstract Thermal conductivity is one of the crucial properties for thermal modelling as well as tunnelling or geological modelling. Available data are mainly from laboratory measurements. Therefore, additional ways, such as correlations with other properties to derive the petro-physical parameter, will be an advantage. The research presented here continues and improves the petrographic-coded model concept with an increased set of data, including a variety of lithologies, and, furthermore, the correlations, including the electrical resistivity. Input parameters are no longer taken from the literature, but are derived directly from measurements. In addition, the results are compared with other published approaches. Results show good correlations with measured data. The comparison with the multi-linear regression method shows acceptable outcome, in contrast to a geometric-mean method, where data scatter. In summary, it can be said that the improved model delivers for both correlation (compressional wave velocity and electrical resistivity with thermal conductivity) positive results.

Keywords Thermal conductivity · Model calculations · Magmatic and metamorphic rocks · Log

Introduction

Energy from geothermal resources is considered regenerative and sustainable. Thermal conductivity is one of the crucial properties for thermal modelling as well as tunnelling or geological modelling. It is also important for hydrocarbon maturation processes. Available data are mainly from laboratory measurements, because it is difficult to determine thermal conductivity in a borehole. However, cores are rare, expensive and represent point information only. Therefore, additional ways, such as correlations with other properties (measurable in the borehole by logging) to describe and derive the petro-physical parameter, will be an advantage to provide data cheaper and faster. Hartmann et al. (2005) published, for example, empirical relationships of thermal conductivity and compressional wave velocity, as well as density, taken from laboratory and log data for shaly sandstones and marls. They also noted that these correlations are valid for local conditions only.

Barry-Macaulay et al. (2013) made laboratory measurements on rocks and soils from the area around Melbourne (Australia) and stated that thermal conductivity varies with the moisture content, density, mineral composition, and particle size. They measured sandstone, siltstone, and basalt samples, and the results showed that thermal conductivity rises with an increase in density in dry conditions. Saturated samples of sandstone and siltstone did not show a correlation between density and thermal conductivity, but for basalt, a linear increase could be observed. These effects were related to anisotropy and mineralogy of the samples.

Abdulagatova et al. (2009) show the dependence of the effective thermal conductivity (ETC) with pressure and temperature. The ETC increases in porous rocks with

✉ Nina Gegenhuber
nina.gegenhuber@unileoben.ac.at

¹ Department of Applied Geophysics, Montanuniversitaet Leoben, Leoben, Austria

increasing pressure and is high in the pressure area of 1–100 MPa. If pressure over 100 MPa is applied, this effect gets weaker due to the bridging of microcracks or a better grain contact. Furthermore, the ETC decreases with rising temperature.

A general relationship between thermal conductivity and density was published by Sundberg et al. (2009) for igneous rocks. They additionally used density logs for the correlation. Oezkahraman et al. (2004) described the derivation of thermal conductivity from p-wave velocity for building rock types. Kukkonen and Peltioniemi (1998) related thermal conductivity, density, magnetic susceptibility, and compressional wave velocity for 2705 different rock types (plutonic rocks, dykes, volcanic rocks, sedimentary, and metamorphic rocks) from Finland. They concluded that there is no general trend between thermal conductivity and other petrophysical properties.

Popov et al. (2003) divided a collection of data into six subcategories, from different silt and sandstone, to limestone and granite as well as gneiss and amphibolite samples. Correlations display again the general trends for thermal conductivity and porosity, electrical resistivity, and permeability. Depending on the rock type, regression curves fit well to the data for porosity and electrical resistivity.

In this paper, improvements and further development of the “petrographic-coded model concept” (Gegenhuber and Schön 2012) for the derivation of a thermal conductivity log are demonstrated. Therefore, correlations between thermal conductivity and other petrophysical properties, which can be derived with standard logs, such as a sonic or resistivity log, are used. Furthermore, the results are compared with other published works, such as a multi-linear regression method (Gąsior and Przelaskowska 2014) and mixing rules (geometric-mean model) for log data (Fuchs 2013). In his Ph.D. Thesis, Fuchs (2013) described the well-log-based determination of rock thermal conductivity with mixing rules in the North German Basin. The main rock types are sandstone and siltstone. In the thesis, three published papers are combined, where the measurements on samples in the laboratory, different mixing rules for thermal conductivity calculation, and a well-log-based prediction of thermal conductivity are discussed. Further equations for the calculation of thermal properties for sedimentary rock are presented by Fuchs et al. (2015) using again the well-log data.

Furthermore, the comparison with a multi-linear regression method, as presented by Gąsior and Przelaskowska (2014), is introduced, as verification for our model correlated thermal conductivity log. For the multi-linear regression method, other input parameters were

used, due to the fact that some of the logs used by Gąsior and Przelaskowska are not available for the used boreholes.

The research presented here continues and improves the petrographic-coded model concept with an increased set of data, including a variety of lithologies, and, furthermore, new correlations, including electrical resistivity. The input parameters are no longer taken from the literature, but are derived directly from measurements. In addition, the results are compared with other published approaches. Data are sorted in the following lithology groups: granite/gneiss, phyllite, mica schist, sandstone, and basalt. In a second step, petrophysical parameters are calculated with respect to the mineral composition out of measured data and the correlations are described with model calculations. Next, the respective models are applied on the log data to determine a fast evaluation of the thermal conductivity in boreholes. Used are, therefore, the elastic wave velocity and specific electrical resistivity due to the fact that these are the standard logs. The last part is the evaluation of the models and the comparison of various approaches to derive a thermal conductivity log.

Methods and samples

Samples

Samples are selected from a database available at the Chair of Applied Geophysics (Montanuniversitaet Leoben, Austria). Selected are different lithologies (sedimentary, magmatic, and metamorphic rock types) from Austria. The database contains all measured petrophysical data from the last couple of years, including geothermal projects. The lithologies used are granite and gneiss, phyllite, various mica schist samples and sandstone as well as basalt. Details and data can be found in the “Appendix”.

For the derivation and testing of the model concept on log data and the comparison with other approaches, data from the continental deep drilling project (KTB) in Germany are used. These are available online (<http://www-icdp.icdp-online.org/sites/ktb/welcome.html>).

Laboratory measurements

Thermal conductivity

Thermal conductivity relates the heat flow density q and the temperature gradient $\text{grad } T$. The SI unit is $\text{Wm}^{-1} \text{K}^{-1}$. Igneous rocks show high thermal conductivity for acid or felsic and lower values for basic or mafic rocks. In sedimentary rocks, the quartz content is essential. Metamorphic

rocks can show anisotropy in their thermal conductivity parallel and perpendicular to the schistosity (Schoen 2011). The thermal conductivity is determined with a non-steady-state (transient) method. The thermal conductivity meter TK04 (1990, from TeKa, Berlin) is used for the measurements. As heat source serves, a needle encased in a cylinder (half-spaced line-source). The needle is pressed onto the sample with 15 bar, and a contact agent is used to establish an optimal heat flow. A defined heating power (here 3 Wm^{-1}) is used. The thermistor in the middle of the needle measures the temperature as a function of time. In the end, a heating cycle is recorded and analysed. To consider an anisotropic effect, the needle is rotated in 45° steps for each measurement.

Electrical resistivity

Specific electrical resistivity ρ (Ωm) and its reverse, the electrical conductivity, are intrinsic material properties. The electrical resistivity of rocks is controlled by the rock type, porosity, fluid type in the pore space, clay content, and metallic content. Most of the rock forming minerals show high resistivity. The resistivity is high in dry rocks, but conductivity rises in water saturated, porous rocks because of electrolytic conductivity and solid fluid interactions (Schoen 2015). To determine the specific electric resistivity, a four-point-light instrument (Type: LF 325 from WTW, Germany) is used.

Compressional wave velocity

Compressional wave velocity (v_p) is described with bulk modulus k , the shear modulus μ , and the bulk density ρ . The SI unit is ms^{-1} . Wave velocity is related to the solid rock skeleton, the pore volume and cracks, the grain contact, pressure and temperature, the fluid saturation, and the type of pore filling. If the porosity increases, the velocity decreases. In metamorphic rocks, anisotropy can be recognized. Igneous rocks show an increase in velocity from acid to basic minerals. The velocity in sedimentary rocks is influenced by porosity the most (Schoen 2015). The compressional wave velocity is determined with a bench-top ultrasonic device (main parts from 2012). The plug (diameter = 1 inch) is fixed in the device at a pressure of 5 bar. A transmitter produces a mechanical pulse, moving through the sample and recorded at the receiver. To get a good contact between the sample and the transmitters, a contact agent is used. The measured signals are recorded and saved on the computer. To pick the first arrivals and calculate the velocity, a self-written program with MATLAB is applied (Gegenhuber and Steiner-Luckabauer 2012).

The models used for correlations and derivation of the solid parameters

Thermal conductivity depends mainly on mineral composition and porosity or fractures. In this study, laboratory data are described with three equation types to find correlations between thermal conductivity and compressional wave velocity as well as resistivity. To link the matrix and fluid properties, the inclusion model after Budiansky and O'Connell (1976) (elastic properties) and Clausius–Mosotti model (Berryman 1995) (thermal properties) is used. To calculate the formation factor (resistivity), the Archie' equations (Archie 1942) are applied. The following paragraphs describe that the three model types are used. These models are used, because first results showed good fit with the data.

The Archie equation, derived from experimental data, combines the cementation factor m (–), the resistivity of the formation R_0 (Ωm), the effective porosity Φ (–), the resistivity of water R_w (Ωm), and the formation factor F (–)

$$F = \frac{1}{\Phi^m} = \frac{R_0}{R_w} \quad (1)$$

Samples with a low cementation factor show flat or jointed pores. Spherical pores show a higher cementation factor. The formation factor (F) is independent of the rock type (all rock building minerals are isolators). Just the shape and the volume of the pore space have an influence on this parameter if the pore space is filled with water and the samples are clay free.

The inclusion model estimates penny-shaped pores, and therefore, Budiansky and O'Connell (1976) developed equations for the elastic properties. The approach assumes high frequencies (ultrasonic laboratory measurements) for saturated rocks and idealizes ellipsoidal inclusions, isotropic, and linear elastic rock matrix, and that cracks are isolated with respect to fluid flow:

$$k_{sc} = k_s \left[1 - \frac{16}{9} \frac{1 - v_{sc}^2}{1 - 2v_{sc}} \varepsilon \right], \quad (2)$$

$$\mu_{sc} = \mu_s \left[1 - \frac{32}{45} \frac{(1 - v_{sc})(5 - v_{sc})}{2v_{sc}} \varepsilon \right], \quad (3)$$

where k_{sc} is the calculated bulk modulus, k_s is the bulk modulus of the host material, μ_{sc} is the calculated shear modulus, and μ_s is the shear modulus of the host material.

Bulk and shear moduli of the host material are determined from laboratory measurements by plotting the measured bulk or shear modulus versus porosity for each lithology and extrapolating the linear correlation to zero porosity. This method delivers optimal input values for the model calculations:

$$\varepsilon = \left(\frac{\Phi}{\alpha}\right) \left(\frac{3}{4\pi}\right), \quad (4)$$

where ε is the crack density parameter, α is the aspect ratio ($\alpha = a/c$), a and c are axes of ellipsoid, v_{sc} is the effective Poisson's ratio:

$$v_{sc} \approx v_s \left[1 - \frac{16}{9} \varepsilon\right], \quad (5)$$

and v_s is the Poisson ratio of the host material.

To calculate the velocity of the compressional wave v_p , also the bulk density ρ_b is needed:

$$\rho_b = (1 - \Phi)\rho_s + \Phi\rho_{fluid}, \quad (6)$$

$$v_p = \left(\frac{k_{sc} + \frac{4}{3}\mu_{sc}}{\rho_b}\right)^{1/2}, \quad (7)$$

where ρ_s is the grain density (measured in the laboratory), and ρ_{fluid} is the density of the fluid.

Thermal conductivity is calculated with the equation of Clausius–Mossotti (Berryman 1995):

$$\lambda_{CM} = \frac{1 - 2\Phi R_{mi}(\lambda_S - \lambda_{fl})}{1 + \Phi R_{mi}(\lambda_S - \lambda_{fl})}, \quad (8)$$

$$R_{mi} = \frac{1}{9} \left(\frac{1}{L_{a,b,c} \lambda_{fl} + (1 - L_{a,b,c}) \lambda_S}\right), \quad (9)$$

where λ_S is the thermal conductivity of the matrix, λ_{fl} is the thermal conductivity of the inclusion, and R_{mi} is the function of depolarization exponents L_a , L_b , and L_c . In this study, the shape of the pores is idealized as plate-like objects ($a = b \gg c$). The model assumes randomly arranged inclusions.

The thermal conductivity of the matrix λ_S is also determined by plotting thermal conductivity versus effective porosity, and the linear trend is extrapolated to zero porosity. To eliminate the influence of pores, thermal conductivity is modelled for correlations (Clausius–Mossotti) under consideration of different aspect ratios. The experimental data should range between higher and lower aspect ratio.

By summarizing, the following steps are carried out for each lithology:

- Plotting thermal conductivity and bulk/shear modulus of measured data versus porosity for each rock type (petrographic code influence).
- Extrapolating the linear correlation to zero porosity for derivation of the solid host parameters (input values for the model calculations).
- Plotting thermal conductivity versus compressional wave velocity and formation factor for the correlations.
- Data are described with model calculations and different aspect ratios for an optimal result.
- The calculated correlations are described with one equation for the further application on log data.

Models for comparison and evaluation

Geometric-mean approach (Fuchs 2013)

Lichtenecker (1924) introduced the geometric-mean model and Fuchs (2013) applied this model to calculate matrix and bulk thermal conductivity:

$$\lambda_m = \sum_{i=1}^n \lambda_i V_i, \quad (10)$$

where λ_m is the matrix thermal conductivity [$\text{Wm}^{-1} \text{K}^{-1}$], λ_i is the thermal conductivity each component [$\text{Wm}^{-1} \text{K}^{-1}$], and V_i is the volume fraction each component [–].

For this model, the porosity is predicted from the neutron log. The thermal conductivity values for the used minerals are literature values. The KTB provides data from XRD, and therefore, the volume fractions of the main minerals are known. Used minerals are quartz, potassium feldspar, biotite, white mica, amphibolite, chlorite, garnet, and plagioclase.

Multi-linear regression

The relationships between several independent variables can be analysed with the statistical technique of the multiple regression. Gašior and Przelaskowska (2014) used empirical models based on the relationship between thermal conductivity and other petrophysical parameters on well-log data from Meso-Paleozoic rocks from Tarnów–Dębica (Poland). They distinguished between siliciclastic and carbonate rocks and a determination coefficient at minimum >0.74 indicates the correctness of their models (Gašior and Przelaskowska 2014). The program Interactive Petrophysics (Senergy) allows the user to predict a result curve from different input logs. The curve to predict (in this case thermal conductivity data from the cores) is defined, and these data are used to create the model. With the different input logs, the “new log” will be calculated and the percentage used of each log is given. The input curves are the corrected gamma ray (CGR), the Laterolog deep (LLD), the inverse velocity of the compression wave (DTCO), and the neutron porosity (NPHI).

The derived regression formula is

$$\lambda = f(v_p, \text{CGR}, \text{LLD}, \text{NPHI}), \quad (11)$$

$$\lambda = 10(-1.016 + 0.097 \text{CGR} + 0.56 \text{DTCO} + 0.004 \text{LLD} - 0.075 \text{NPHI}). \quad (12)$$

Results

Laboratory data

The model calculations are used for five different lithologies. For every lithology, the properties of the host rock are determined by laboratory data and these are the basis for

Table 1 Overview of the host properties (n = number of samples for F = formation factor and v_p = compressional wave velocity, k_s = compressional modulus, μ_s = shear modulus, ρ_{matrix} = grain density,and λ_s = thermal conductivity) and aspect ratios α and cementation factor m for different lithologies

| Rock type | n (F/v_p) | k_s (GPa) | μ_s (GPa) | ρ_{matrix} (g cm ⁻³) | λ_s (Wm ⁻¹ K ⁻¹) | α | m |
|----------------|-----------------|-------------|---------------|--|---|----------|-----|
| Granite/gneiss | 47/25 | 32 | 29 | 2.74 | 3.4 | 0.1 | 1.3 |
| Phyllite | 61/35 | 56 | 43 | 2.79 | 4 | 0.2 | 1.9 |
| Mica schist | 57/25 | 48 | 38 | 2.77 | 4.3 | 0.02 | 1.3 |
| Sandstone | 33/21 | 43 | 30 | 2.73 | 4.5 | 0.1 | 1.9 |
| Basalt | 17/11 | 84 | 68 | 2.98 | 3.2 | 0.2 | 1.3 |

the calculations. Table 1 gives an overview of the input data for the inclusion models for different lithologies. The differentiation in the various lithologies and derivation of the input parameter are the petrographic part in the model concept. Detailed interpretation of the lithologies is important to derive excellent results.

Figure 1 shows the results of the correlations between thermal conductivity and compressional wave velocity for different lithologies derived from laboratory data. Points are measured data and lines are calculated with the inclusion models. For each porosity, one value for thermal conductivity is calculated with the inclusion model by Clausius–Mossotti (Berryman 1995) and one value for compressional wave velocity is calculated with the inclusion model by Budiansky and O’Connell (1976). These values are furthermore plotted in Fig. 1 as correlation lines. The result of calculated lines cannot be applied directly. Therefore, trend lines are added to describe the lines and to make the correlations easily applicable in the form of $\lambda = f(v_p \text{ or } F)$. These equations can be found in Table 2.

Granite and gneiss as well as phyllite and mica schist samples can show an anisotropic effect. To make the approach easily applicable, this effect is neglected within the model calculations. Focusing on the singular plots, granite shows better results for the correlation with v_p , where there are some outliers for the correlation with resistivity. The higher thermal conductivities are “Fasergneiss” samples and have higher quartz content. The low velocities are an anisotropic effect.

The basalt and diorite show three outliers with higher thermal conductivity because of lower porosity. They are all from the same stone pit.

Data of mica schist and phyllite show a big variation due to anisotropic effects and variations in porosity data as well as mineralogical effects. The anisotropic effect is neglected, because it is rarely analysed in the borehole too.

Sandstone samples show good correlations for both properties and provide highest thermal conductivities. The highest thermal conductivities can be found for quartzite samples, which are outliers for correlation with formation factor. This is may be a result of their low porosity.

Table 2 summarizes the derived equations for further application on log data. The lines which result from the two models cannot be applied directly on log data; therefore, the resulting correlation lines out of the calculations are additionally described with trend lines, which are presented in Table 2. These equations make the model concept applicable, because the form results in $\lambda = f(v_p \text{ or } F)$.

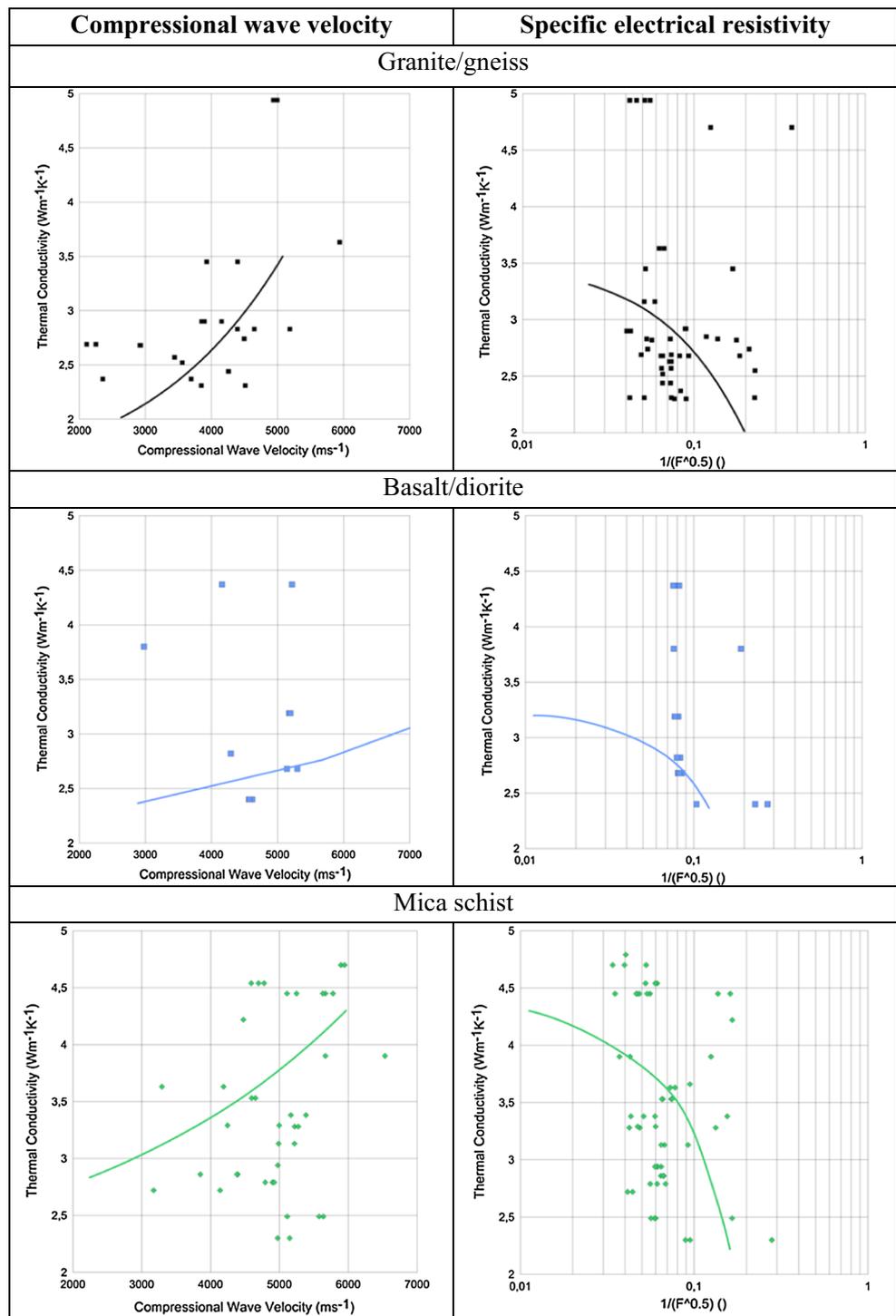
Application on log data

The following figures will show the results for two selected sections of the continental deep drilling project, one for gneiss and one for metabasite. Shown are various combinations with other approaches to set them all in context to each other. Figure 2 shows the gneiss section. In the middle of the figure, the histograms show the variation of the different approaches, starting in the left upper corner with the measured thermal conductivity of the cores, followed by the geometric-mean model, the multi-linear regression model, and thermal conductivity with the equations presented here for the sonic and the resistivity log. The geometric-mean model delivers a broad variation of thermal conductivity values, possibly because the model is developed for sedimentary rock types. The multi-linear regression model and the models presented here deliver good results and the data do not scatter.

The other figures show various correlations between the five thermal conductivity values. Due to the fact that core data show a broader variation, the comparison with these shows also some scattering. This may be the result of anisotropy within the gneiss samples. Best result is delivered by the model using the sonic log. The mean value delivered by the resistivity log fits well to the core data but does not show such a broad range of data as the cores do. A strong correlation can be observed for the approach using the sonic log and the multi-linear regression method.

In the same plots for a metabasite section, the core data show no big variation within the set of data and a lower thermal conductivity than the gneiss samples. The geometric mean again scatters too much to deliver a good result. The multi-linear regression delivers good results

Fig. 1 Plots of v_p and $1/(F^{0.5})$ versus thermal conductivity for the different lithologies: granite/gneiss, basalt, mica schist, and phyllite; dots are measured data from the laboratory, and lines are the result of calculation of thermal conductivity with the inclusion model by Clausius–Mossotti (Berryman 1995) and compressional wave velocity with the inclusion model by Budiansky and O’Connell (1976)



with only a minor overestimation of about $0.5 Wm^{-1} K^{-1}$. Both model concepts also overestimate the core data, data from the sonic log with about $0.5 Wm^{-1} K^{-1}$, and the resistivity log data with $1 Wm^{-1} K^{-1}$. Here, again, a strong correlation between the approach using the sonic log and the multi-linear regression method can be observed.

Figure 3 shows additionally two log sections (2900–3400: gneiss and 3600–4000: metabasite) with the results in a log plot. The dots in the plots are measured samples. In red is the result with the geometric mean and in black the multi-linear regression approach. The fourth track shows the result for λ for basalt from the resistivity (turquoise), λ for basalt from the compressional wave velocity

Fig. 1 continued

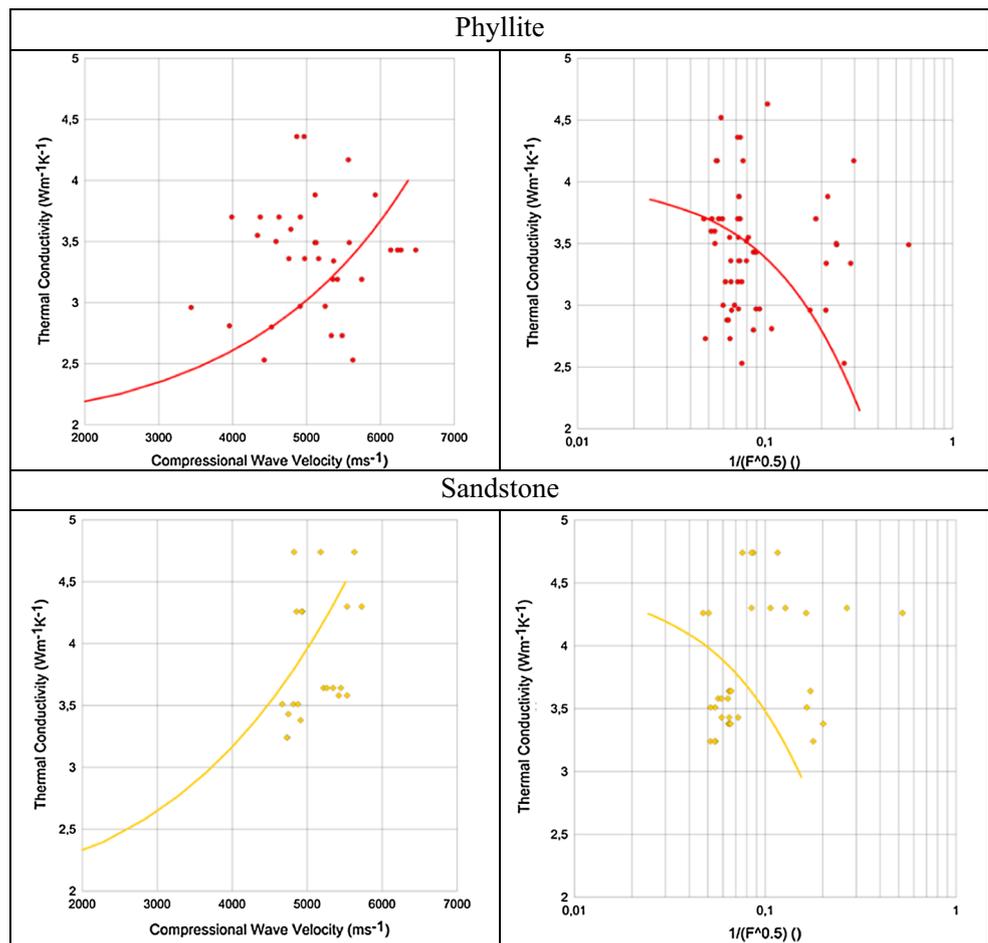


Table 2 Regression equations for a further application on log data derived from the calculated correlation lines [thermal conductivity with the inclusion model by Clausius–Mossotti (Berryman 1995) and

compressional wave velocity with the inclusion model by Budiansky and O’Connell (1976)]

| Rock type | Sonic log | Resistivity log |
|----------------|---|---|
| Granite/gneiss | $\lambda = 1E - 07 v_p^2 - 0.0003v_p + 1.98$ | $\lambda = -3.73(1/F^{0.5})^2 - 3.5 (1/F^{0.5}) + 3.59$ |
| Phyllite | $\lambda = 9E - 08 v_p^2 - 0.0003v_p + 2.48$ | $\lambda = -5.85 (1/F^{0.5}) + 3.98$ |
| Mica schist | $\lambda = 5E - 08v_p^2 - 2E - 05 v_p + 2.63$ | $\lambda = -10.2 (1/F^{0.5})^2 - 12.4(1/F^{0.5}) + 4.5$ |
| Sandstone | $\lambda = 1E - 07v_p^2 - 0.0003v_p + 2.41$ | $\lambda = 5.28 (1/F^{0.5})^2 - 10.93 (1/F^{0.5}) + 4.52$ |
| Basalt | $\lambda = 2E - 08v_p^2 - 1E - 06 v_p + 2.23$ | $\lambda = -4.73 (1/F^{0.5})^2 - 7.69 (1/F^{0.5}) + 3.34$ |

(violet), λ for gneiss from the resistivity (blue), and λ for gneiss from the compressional wave velocity (orange).

Summary and conclusion

In most drilling projects, sonic log and resistivity log are available. Measuring thermal conductivity in a borehole is difficult and time consuming. Therefore, petrographic-

coded models for an indirect estimation are calculated and tested on real data of the continental deep drilling project in Germany (KTB). In addition, the data are compared with other approaches, such as the geometric-mean model and the multi-linear regression model.

The inclusion model is a good basis for the correlation of formation factor, respectively, compressional wave velocity with thermal conductivity. In general, the calculated regressions for the model deliver in general a good fit

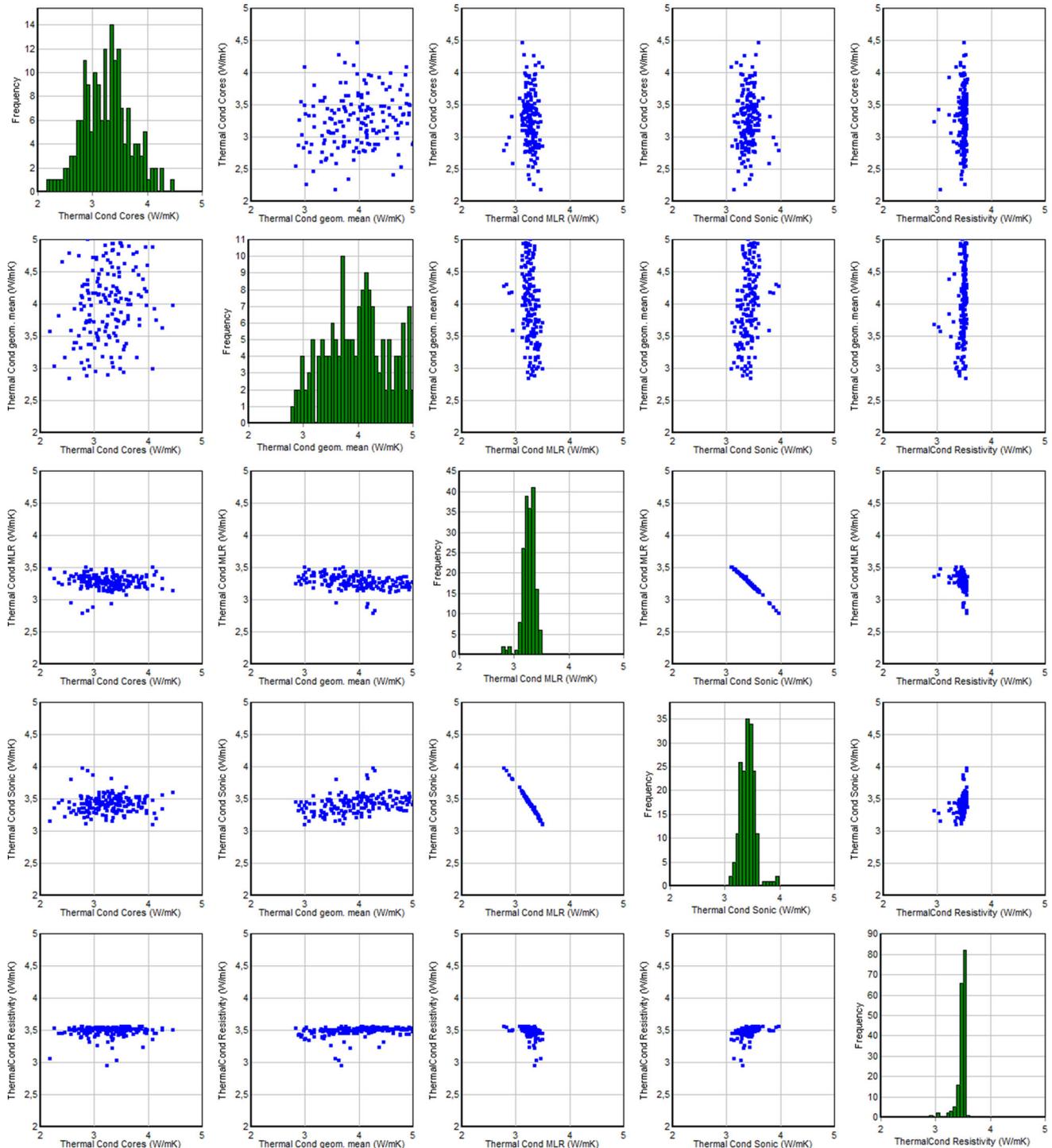


Fig. 2 Thermal conductivity for granite section, thermal conductivity from cores versus the improved derived correlation equations (sonic and resistivity log), the multi-linear regression, and the geometric-mean approach. In green are the histogram data for the various approaches

with the laboratory data for the five selected lithologies: granite/gneiss, phyllite, mica schist, basalt/diorite, and sandstone. In the next step, the models for granite/gneiss and basalt are applied on the log data. The sonic log provides the values for the compressional wave velocity and is

the basis for the first model estimation. The compressional wave velocity shows a strong correlation with the thermal conductivity, and therefore, the calculated data fits the real data excellent.

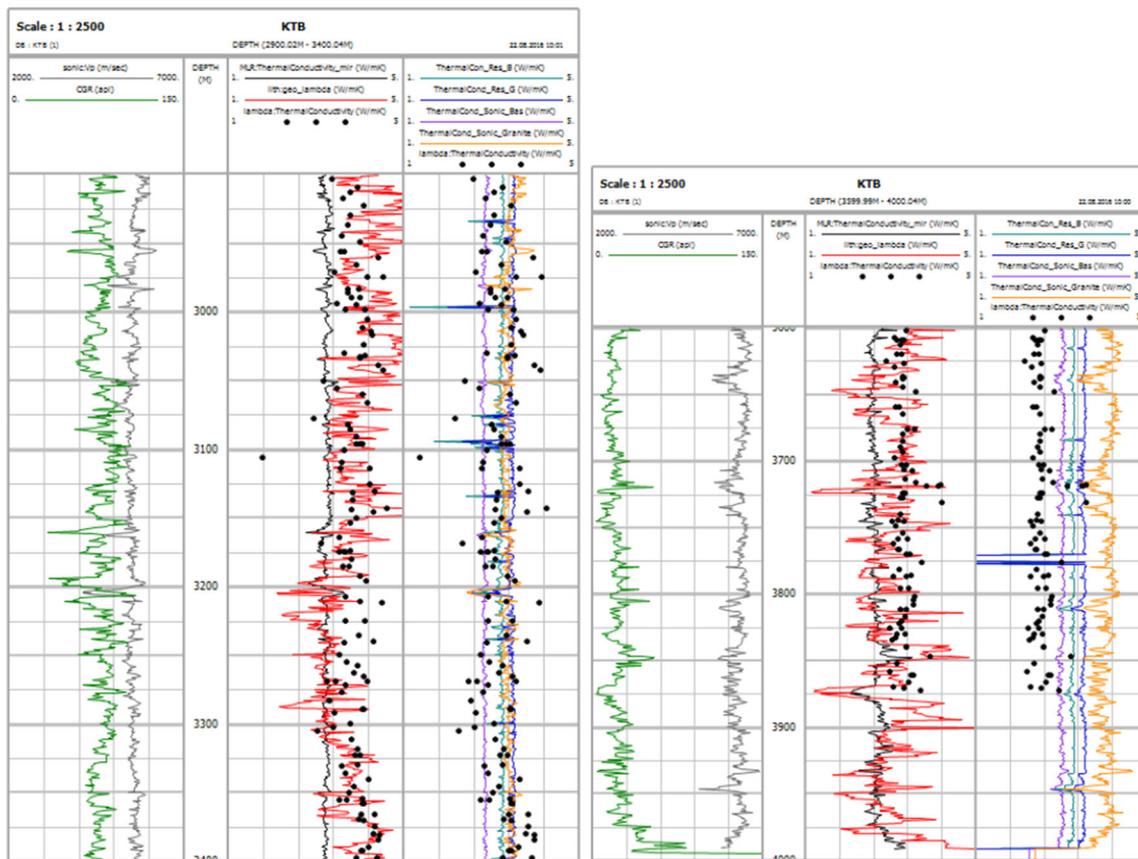


Fig. 3 Log plot for the results of thermal conductivity for 2900–3400 m (gneiss) and 3800–4000 m (metabasite). The first track in both plots shows compressional wave velocity (*green*) and gammaray (*grey*), the third track shows multi-linear regression (black) and the geometric mean (*red*) and plus the measured data

With the resistivity log, the formation factor can be calculated and the second approach can be applied. This model is influenced mostly by the pore shape and their fillings. For the lithology granite/gneiss, it works well, but in the case of basalt, thermal conductivity is slightly overestimated. The calculated models show:

- no depth influence;
- strong correlation with the real data;
- dependence on lithology and pores/fractures (aspect ratio).

The petrographic-coded model concept covers the influence of lithology and delivers good results for the derivation of thermal conductivity with the standard logs (sonic and resistivity log). In comparison with other approaches, data do not scatter, such as the geometric-mean model. This may result from the fact that the model is developed for sedimentary rock types. The multi-linear regression method shows a strong correlation for both gneiss and metabasite sections with the derived thermal conductivity out of the sonic log. For the presented models,

from the samples as dots, the fourth track shows the result for λ for basalt from the resistivity (turquoise), λ for basalt from the compressional wave velocity (violet), λ for gneiss from the resistivity (blue), and λ for gneiss from the compressional wave velocity (orange), and the measured λ from the samples as dots

only the lithology needs to be known and a sonic or resistivity log needs to be available. Especially, the sonic log delivers good results and derived thermal conductivity data can be used for example in geothermal projects. This method is better than using only data from the literature due to the fact that various lithologies can have a broad range of thermal conductivity values.

Acknowledgement Open access funding provided by Montanuniversität Leoben.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix

See Table 3.

Table 3 Data

| Rock type | λ (W/mK) | v_p dry (m/s) | F (-) | Φ_{eff} (-) |
|----------------------------------|------------------|-----------------|---------|-------------------------|
| Granite/gneiss (rich in biotite) | 2.52 | 3559.7 | 232.60 | 0.126 |
| Gneiss (rich in biotite) | 2.37 | 3691.9 | | 0.340 |
| Gneiss (rich in biotite) | 2.37 | 2357.0 | 143.81 | 0.353 |
| Flasergneiss | 4.94 | 4944.0 | 325.94 | 0.940 |
| Flasergneiss | 4.94 | 4936.9 | 563.40 | 0.600 |
| Flasergneiss | 4.94 | 4995.0 | 376.80 | 0.630 |
| Flasergneiss | 4.94 | 488.7 | 469.39 | 0.730 |
| Gneiss | 3.63 | 532.0 | 254.98 | 0.267 |
| Gneiss | 3.63 | 5941.0 | 223.59 | 0.126 |
| Gneiss | 4.70 | | 63.93 | 0.279 |
| Gneiss | 4.70 | | | 0.234 |
| Gneiss | 4.70 | | 7.26 | 0.229 |
| Gneiss | 1.69 | | | 0.494 |
| Gneiss | 1.69 | 3472.0 | 72.45 | 0.443 |
| Gneiss | 3.84 | | | 0.600 |
| Gneiss | 3.30 | | | 0.285 |
| Gneiss | 3.30 | | | 0.332 |
| Gneiss | 4.36 | | | 0.211 |
| Gneiss | 4.36 | | | 0.169 |
| Granite | 2.90 | 3854.3 | 613.69 | 0.480 |
| Granite | 2.90 | 3891.7 | 579.41 | 0.570 |
| Granite | 2.90 | 4155.2 | 552.92 | 0.380 |
| Granite/gneiss | 2.69 | 2248.9 | 185.27 | 0.130 |
| Granite/gneiss | 2.69 | 2111.1 | 416.48 | 0.132 |
| Granite/gneiss | 2.30 | | 123.87 | 0.153 |
| Granite/gneiss | 2.30 | | 169.69 | 0.133 |
| Granite/gneiss | 2.82 | | 31.93 | 0.134 |
| Granite/gneiss | 2.82 | | 311.24 | 0.100 |
| Granite/gneiss | 2.44 | 4255.4 | 233.57 | 0.126 |
| Granite/gneiss | 2.44 | 4261.3 | 188.49 | 0.970 |
| Granite/gneiss (fine grained) | 2.92 | | 124.59 | 0.144 |
| Granite/gneiss (fine grained) | 2.92 | | 127.21 | 0.144 |
| Granite/gneiss (fine grained) | 2.68 | | 115.63 | 0.131 |
| Granite/gneiss (fine grained) | 2.68 | | 243.40 | 0.132 |
| Granite/gneiss (coarse grained) | 2.68 | | 147.78 | 0.165 |
| Granite/gneiss (coarse grained) | 2.55 | | 19.45 | 0.137 |
| Granite | 2.85 | | 72.17 | 0.163 |
| “Knollkopf”-gneiss | 2.68 | 2916.9 | 232.40 | 0.118 |
| “Knollkopf”-gneiss | 2.68 | 2928.8 | 29.29 | 0.150 |
| Migmatit gneiss | 5.90 | 497.8 | 312.29 | 0.520 |
| Migmatit Granite | 2.74 | 455.0 | 349.19 | 0.112 |
| Migmatit Granite | 2.74 | 4495.7 | 22.81 | 0.126 |
| Orthogneiss | 3.16 | 242.9 | 381.89 | 0.117 |
| Orthogneiss | 3.16 | | 286.47 | 0.126 |
| Para(Bi-)gneiss | 2.31 | 3846.7 | 565.77 | 0.870 |
| Para(Bi-)gneiss | 2.31 | 4512.9 | 382.46 | 0.710 |
| Gneiss | 2.63 | | 193.72 | 0.171 |
| Gneiss | 2.63 | | 184.65 | 0.188 |

Table 3 continued

| Rock type | λ (W/mK) | v_p dry (m/s) | F (-) | Φ_{eff} (-) |
|------------------------------------|------------------|-----------------|---------|-------------------------|
| Gneiss (rich in quartz) | 3.45 | 3929.1 | 369.78 | 0.650 |
| Gneiss (rich in quartz) | 3.45 | 4393.8 | 35.31 | 0.720 |
| Gneiss (rich in pyrite) | 2.31 | | 185.23 | 0.133 |
| Gneiss (rich in pyrite) | 2.31 | | 19.62 | 0.146 |
| Gneiss | 2.57 | 3442.7 | 239.89 | 0.137 |
| Gneiss | 2.57 | | 185.61 | 0.180 |
| Granite | 2.83 | 5186.9 | 357.48 | 0.830 |
| Granite | 2.83 | 4392.4 | 189.63 | 0.930 |
| Granite | 2.83 | 4647.3 | 52.98 | 0.820 |
| Phyllite (rich in chlorite) | 3.43 | 6477.1 | 124.27 | 0.630 |
| Phyllite (rich in chlorite) | 3.43 | 6138.9 | 127.99 | 0.520 |
| Phyllite (rich in chlorite) | 3.43 | 6276.7 | 133.94 | 0.630 |
| Phyllite (rich in chlorite) | 3.43 | 6232.7 | 124.48 | 0.520 |
| Phyllite (rich in chlorite) | 3.70 | 3986.0 | 313.26 | 0.840 |
| Phyllite (rich in chlorite) | 3.70 | 379.1 | 285.37 | 0.620 |
| Phyllite (rich in chlorite) | 3.70 | 4372.8 | 449.57 | 0.620 |
| Phyllite (rich in chlorite) | 3.70 | 454.2 | 369.77 | 0.640 |
| Phyllite (rich in chlorite) | 4.17 | 5566.7 | 172.26 | 0.520 |
| Phyllite (rich in chlorite) | 4.17 | 499.3 | 11.34 | 0.630 |
| Phyllite (rich in chlorite) | 4.17 | 549.3 | 326.14 | 0.430 |
| Phyllite (rich in chlorite) | 4.17 | 53.5 | 332.33 | 0.420 |
| Phyllite (rich in chlorite) | | 4421.2 | 223.19 | 0.620 |
| Phyllite (rich in chlorite) | | 4134.3 | 118.97 | 0.720 |
| Phyllite (rich in chlorite) | | 4557.6 | 191.49 | 0.720 |
| Phyllite (rich in chlorite) | 3.19 | 5746.0 | 196.20 | 0.470 |
| Phyllite (rich in chlorite) | 3.19 | 5355.0 | 265.33 | 0.490 |
| Phyllite (rich in chlorite) | 3.19 | 5418.8 | 231.29 | 0.590 |
| Phyllite (rich in chlorite) | 3.19 | 533.0 | 177.95 | 0.490 |
| Phyllite (rich in chlorite) | 4.36 | 4866.5 | 183.90 | 0.750 |
| Phyllite (rich in chlorite) | 4.36 | 4965.2 | 197.20 | 0.660 |
| Phyllite (rich in chlorite) | 4.63 | | 94.30 | 0.116 |
| Phyllite (rich in chlorite) | 2.73 | 5483.5 | 236.84 | 0.430 |
| Phyllite (rich in chlorite) | 2.73 | 5335.1 | 435.42 | 0.420 |
| Phyllite (rich in chlorite quartz) | 4.52 | | 295.80 | 0.132 |
| Phyllite (rich in graphite) | 3.60 | 32.9 | 344.75 | 0.780 |
| Phyllite (rich in graphite) | 3.60 | 4787.0 | 376.30 | 0.580 |
| Phyllite (rich in graphite) | 2.88 | 442.0 | 255.28 | 0.970 |
| Phyllite (rich in graphite) | 2.88 | 43.8 | 247.24 | 0.880 |
| Phyllite (rich in graphite) | 3.49 | 5575.9 | 17.25 | 0.880 |
| Phyllite (rich in graphite) | 3.49 | 5112.4 | 17.29 | 0.970 |
| Phyllite (rich in graphite) | 3.49 | 5128.6 | 2.96 | 0.170 |
| Phyllite (rich in graphite) | 3.34 | 566.4 | 12.28 | 0.126 |
| Phyllite (rich in graphite) | 3.34 | 5364.3 | 22.25 | 0.780 |
| Phyllite (rich in graphite) | 3.50 | 321.4 | 343.30 | 0.850 |
| Phyllite (rich in graphite) | 3.50 | 4586.0 | 17.44 | 0.131 |
| Phyllite (rich in graphite) | 3.55 | 426.1 | 151.10 | 0.900 |
| Phyllite (rich in graphite) | 3.55 | | 193.11 | 0.890 |
| Phyllite (rich in graphite) | 3.55 | 4334.7 | 239.84 | 0.880 |

Table 3 continued

| Rock type | λ (W/mK) | v_p dry (m/s) | F (-) | Φ_{eff} (-) |
|------------------------------------|------------------|-----------------|---------|-------------------------|
| Phyllite (rich in graphite) | 2.96 | 3437.3 | 227.19 | 0.980 |
| Phyllite (rich in graphite) | 2.96 | 451.5 | 33.22 | 0.790 |
| Phyllite (rich in graphite) | 2.96 | 468.6 | 22.50 | 0.780 |
| Phyllite | 2.81 | 3956.6 | 85.27 | 0.197 |
| Phyllite | 3.00 | | 281.50 | 0.143 |
| Phyllite | 3.00 | | 212.20 | 0.168 |
| Phyllite | 2.97 | | 125.30 | 0.950 |
| Phyllite | 2.97 | 5252.6 | 114.99 | 0.190 |
| Phyllite | 2.97 | 4911.3 | 192.34 | 0.860 |
| Phyllite | 2.53 | 4427.0 | 176.90 | 0.970 |
| Phyllite | 2.53 | 5627.8 | 14.37 | 0.110 |
| Phyllite | 3.70 | 453.0 | 186.30 | 0.800 |
| Phyllite | 3.70 | 499.6 | 28.80 | 0.790 |
| Phyllite | 3.70 | 4914.4 | 184.62 | 0.790 |
| Phyllite | 3.70 | 4626.1 | 196.93 | 0.790 |
| Phyllite | 2.80 | 459.9 | 134.16 | 0.790 |
| Phyllite | 2.80 | 4525.5 | 133.21 | 0.100 |
| Phyllite | 3.52 | | 158.65 | 0.170 |
| Phyllite | 3.88 | 66.1 | 189.33 | 0.730 |
| Phyllite | 3.88 | 5115.4 | 191.82 | 0.120 |
| Phyllite | 3.88 | 5931.1 | 21.53 | 0.130 |
| Phyllite | 3.36 | 5163.2 | 187.10 | 0.113 |
| Phyllite | 3.36 | 4972.6 | 191.24 | 0.920 |
| Phyllite | 3.36 | 4758.3 | 232.64 | 0.710 |
| Phyllite | 3.36 | | 158.44 | 0.920 |
| Black phyllite/quartzite | | 4338.8 | 147.73 | 0.580 |
| Black phyllite/quartzite | | | 96.45 | 0.940 |
| Black phyllite/quartzite | | | 133.64 | 0.770 |
| Green phyllite (rich in anhydrite) | 5.42 | 2673.0 | 43.83 | |
| Mica schist (rich in biotite) | 2.72 | 3171.7 | 584.12 | 0.470 |
| Mica schist (rich in biotite) | 2.72 | 4135.8 | 511.14 | 0.580 |
| Mica schist (rich in chlorite) | 3.28 | 5222.0 | 557.18 | 0.520 |
| Mica schist (rich in chlorite) | 3.28 | 5275.1 | 56.69 | 0.740 |
| Mica schist (rich in chlorite) | 3.28 | 573.2 | 422.61 | 0.620 |
| Mica schist | 3.66 | | 111.99 | 0.240 |
| Mica schist | 3.54 | | | 0.133 |
| Mica schist | 3.38 | 498.9 | 41.72 | 0.520 |
| Mica schist | 3.38 | 5165.8 | 282.90 | 0.510 |
| Mica schist | 3.38 | 535.4 | 381.51 | 0.410 |
| Mica schist | 3.38 | 5385.9 | 535.26 | 0.620 |
| Mica schist | 2.94 | 4975.2 | 239.91 | 0.510 |
| Mica schist | 2.94 | 493.7 | 267.40 | 0.510 |
| Mica schist | 2.94 | 481.7 | 262.46 | 0.510 |
| Mica schist | 2.94 | 492.4 | 282.96 | 0.520 |
| Mica schist | 5.46 | 3995.9 | 276.60 | 0.740 |
| Mica schist | 5.46 | 649.1 | 168.70 | 0.960 |
| Mica schist | 4.54 | 4588.8 | 267.39 | 0.620 |
| Mica schist | 4.54 | 4694.8 | 362.56 | 0.510 |

Table 3 continued

| Rock type | λ (W/mK) | v_p dry (m/s) | F (-) | Φ_{eff} (-) |
|--------------------------------|------------------|-----------------|---------|-------------------------|
| Mica schist | 4.54 | 4775.6 | 281.46 | 0.530 |
| Mica schist | 5.98 | 4199.6 | 314.50 | 0.590 |
| Mica schist | 5.98 | 4149.2 | 294.88 | 0.680 |
| Mica schist | 5.98 | 4196.9 | 53.85 | 0.570 |
| Mica schist | 5.98 | 436.1 | 335.35 | 0.580 |
| Mica schist | 4.45 | 5775.7 | 38.43 | 0.380 |
| Mica schist | 4.45 | 5248.8 | 457.60 | 0.480 |
| Mica schist | 4.45 | 5668.4 | 323.81 | 0.390 |
| Mica schist | 4.45 | 577.0 | 347.89 | 0.390 |
| Mica schist | 4.45 | 566.2 | 427.64 | 0.280 |
| Mica schist | 4.45 | 5111.1 | 464.14 | 0.390 |
| Mica schist | 4.45 | 65.6 | 816.42 | 0.390 |
| Mica schist | 4.45 | 5628.3 | 53.32 | 0.390 |
| Mica schist | 3.29 | 4995.3 | 278.85 | 0.390 |
| Mica schist | 3.29 | 4245.0 | 447.93 | 0.780 |
| Mica schist | 2.49 | 5636.4 | 278.23 | 0.720 |
| Mica schist | 2.49 | 535.3 | 286.22 | 0.510 |
| Mica schist | 2.49 | 5576.3 | 36.63 | 0.610 |
| Mica schist | 2.49 | 5111.4 | 314.88 | 0.620 |
| Mica schist | 3.63 | 4188.5 | 188.57 | 0.111 |
| Mica schist | 3.63 | 3292.4 | 165.80 | 0.134 |
| Mica schist (rich in chlorite) | 2.79 | 4898.5 | 212.90 | 0.420 |
| Mica schist (rich in chlorite) | 2.79 | 4922.1 | 321.16 | 0.420 |
| Mica schist (rich in chlorite) | 2.79 | 4792.7 | 267.46 | 0.420 |
| Mica schist (rich in chlorite) | 2.86 | 4392.4 | 242.60 | 0.600 |
| Mica schist (rich in chlorite) | 2.86 | 3851.1 | 224.18 | 0.800 |
| Mica schist (rich in chlorite) | 2.86 | 4382.2 | 227.72 | 0.590 |
| Mica schist (rich in garnet) | 3.90 | 666.4 | 64.18 | 0.290 |
| Mica schist (rich in garnet) | 3.90 | 6533.7 | 725.86 | 0.270 |
| Mica schist (rich in garnet) | 3.90 | 5669.0 | 547.86 | 0.350 |
| Mica schist (rich in garnet) | 4.70 | 598.8 | 356.80 | 0.360 |
| Mica schist (rich in garnet) | 4.70 | 5892.9 | 867.23 | 0.460 |
| Mica schist (rich in garnet) | 4.70 | 5946.6 | 634.71 | 0.270 |
| Mica schist (rich in garnet) | 4.79 | | 613.41 | 0.450 |
| Mica schist (rich in garnet) | 4.79 | | | 0.470 |
| Green schist | | 482.8 | 166.35 | 0.860 |
| Green schist | | 5436.2 | 171.31 | 0.170 |
| Green schist | | 5353.9 | 157.14 | 0.980 |
| Green schist | | 579.1 | 144.43 | 0.116 |
| Mica schist (rich in quartz) | 4.22 | 4474.8 | 36.46 | 0.860 |
| Mica schist (rich in quartz) | 3.13 | 5218.7 | 219.56 | 0.740 |
| Mica schist (rich in quartz) | 3.13 | 4987.4 | 239.95 | 0.740 |
| Mica schist (rich in quartz) | 3.13 | 431.7 | 117.59 | 0.159 |
| Mica schist (rich in quartz) | 3.53 | 416.9 | 182.75 | 0.160 |
| Mica schist (rich in quartz) | 3.53 | 4652.8 | 231.41 | 0.770 |
| Mica schist (rich in quartz) | 3.53 | 447.8 | 235.13 | 0.680 |
| Mica schist (rich in quartz) | 3.53 | 4597.0 | 179.92 | 0.116 |
| Mica schist (rich in Ca) | 2.30 | 491.5 | 12.79 | 0.630 |

Table 3 continued

| Rock type | λ (W/mK) | v_p dry (m/s) | F (-) | Φ_{eff} (-) |
|------------------------------|------------------|-----------------|---------|-------------------------|
| Mica schist (rich in Ca) | 2.30 | 4972.6 | 111.40 | 0.127 |
| Mica schist (rich in Ca) | 2.30 | 5148.4 | 125.95 | 0.640 |
| Mica schist | 3.33 | | 283.40 | 0.950 |
| Mica schist | 3.45 | 2941.9 | 334.74 | 0.161 |
| Mica schist | 3.45 | 213.0 | 36.18 | 0.280 |
| Mica schist (rich in quartz) | 4.65 | 3554.0 | 337.12 | 0.187 |
| Mica schist (rich in quartz) | 4.65 | 3449.8 | 248.32 | 0.168 |
| Mica schist (rich in quartz) | 4.65 | 3755.5 | 452.41 | 0.144 |
| Mica schist (rich in quartz) | 3.83 | 4529.5 | 252.87 | 0.270 |
| Mica schist (rich in quartz) | 3.83 | 4615.0 | 257.60 | 0.140 |
| Mica schist (rich in quartz) | 3.83 | 4714.6 | 252.64 | 0.145 |
| Mica schist | 2.85 | 3811.0 | 162.73 | 0.398 |
| Mica schist | 2.85 | 3727.1 | 127.36 | 0.438 |
| Mica schist | 2.85 | 393.5 | 221.98 | 0.252 |
| Mica schist | 2.85 | 379.3 | 82.87 | 0.477 |
| Mica schist | 3.33 | | 423.73 | 0.152 |
| Metaarenite | 3.58 | 552.6 | 312.12 | 0.300 |
| Metaarenite | 3.58 | 5533.7 | 287.34 | 0.290 |
| Metaarenite | 3.58 | 5421.3 | 245.53 | 0.390 |
| Metaarenite | 3.64 | 5216.7 | 239.72 | 0.390 |
| Metaarenite | 3.64 | 5263.7 | 234.86 | 0.490 |
| Metaarenite | 3.64 | 5449.2 | 33.67 | 0.390 |
| Metaarenite | 3.64 | 5346.7 | 228.36 | 0.490 |
| Metaarenite | 3.43 | 518.2 | 286.72 | 0.840 |
| Metaarenite | 3.43 | 4748.8 | 239.28 | 0.530 |
| Metaarenite | 3.43 | 589.3 | 191.92 | 0.730 |
| Metaarenite | 3.38 | 578.9 | 244.00 | 0.490 |
| Metaarenite | 3.38 | 533.2 | 234.84 | 0.490 |
| Metaarenite | 3.38 | 526.3 | 24.74 | 0.490 |
| Metaarenite | 3.38 | 4912.4 | 231.11 | 0.690 |
| Metaarenite | 4.30 | 68.5 | 61.56 | 0.190 |
| Metaarenite | 4.30 | 5531.3 | 139.60 | 0.690 |
| Metaarenite | 4.30 | 5725.8 | 13.96 | 0.990 |
| Metaarenite | 4.30 | 588.7 | 87.86 | 0.980 |
| Metaarenite | 4.74 | 4822.7 | 173.69 | 0.690 |
| Metaarenite | 4.74 | 5178.7 | 132.96 | 0.790 |
| Metaarenite | 4.74 | 5629.6 | 74.31 | 0.900 |
| Metaarenite | 4.74 | | 139.34 | 0.790 |
| Metaarenite | 6.80 | 4635.5 | 324.80 | 0.530 |
| Metaarenite | 6.80 | 4683.3 | 448.93 | 0.540 |
| Metaarenite | 6.80 | 4863.9 | 353.85 | 0.530 |
| Metaarenite | 3.51 | 4666.7 | 36.63 | 0.530 |
| Metaarenite | 3.51 | 4815.5 | 335.81 | 0.530 |
| Metaarenite | 3.51 | 4876.8 | 374.50 | 0.630 |
| Metaarenite | 4.26 | 4938.1 | 37.13 | 0.420 |
| Metaarenite | 4.26 | 545.7 | 391.93 | 0.630 |
| Metaarenite | 4.26 | 4857.8 | 448.16 | 0.420 |
| Metaarenite | 4.26 | 4923.6 | 3.64 | 0.530 |

Table 3 continued

| Rock type | λ (W/mK) | v_p dry (m/s) | F (-) | Φ_{eff} (-) |
|-----------------------------|------------------|-----------------|---------|-------------------------|
| Metaarenite | 3.24 | 517.7 | 333.84 | 0.530 |
| Metaarenite | 3.24 | 491.2 | 375.60 | 0.530 |
| Metaarenite | 3.24 | 4731.1 | 339.39 | 0.630 |
| Metaarenite | 3.24 | 4729.9 | 31.40 | 0.640 |
| Sandstone | 2.18 | 1887.7 | 46.98 | 0.1141 |
| Quartzite | 4.60 | 5464.5 | 261.90 | 0.186 |
| Quartzite | 4.60 | 5484.5 | 265.99 | 0.178 |
| Quartzite | 4.60 | 5484.5 | 259.40 | 0.151 |
| Quartzite | 4.60 | 55.0 | 331.90 | 0.151 |
| Metaarenite | 3.51 | 494.3 | 314.72 | 0.221 |
| Sandstone (with clay. marl) | 2.27 | 2759.2 | 7.61 | 0.752 |
| Sandstone (with clay. marl) | 2.27 | 321.0 | 78.94 | 0.688 |
| Sandstone (with clay. marl) | 2.27 | 31.0 | 78.90 | 0.693 |
| Sandstone (with clay. marl) | 2.10 | 4271.0 | 96.20 | 0.726 |
| Sandstone (with clay. marl) | 2.10 | 4277.6 | 86.75 | 0.734 |
| Sandstone (with clay. marl) | 2.10 | 4315.6 | 85.77 | 0.740 |
| Metabasalt | 2.40 | 4614.0 | 13.21 | 0.125 |
| Metabasalt | 2.40 | 4564.2 | 92.50 | 0.127 |
| Metabasalt | 2.40 | 4571.1 | 18.50 | 0.125 |
| Metabasite | 3.19 | 5173.6 | 151.82 | 0.730 |
| Metabasit | 3.19 | 494.6 | 166.48 | 0.720 |
| Metabasit | 3.19 | 5192.9 | 165.67 | 0.860 |
| Metabasit | 2.68 | 5298.9 | 137.49 | 0.730 |
| Metabasit | 2.68 | 5144.3 | 152.86 | 0.730 |
| Metabasit | 2.68 | 563.5 | 147.17 | 0.830 |
| Metabasit | 2.82 | 439.0 | 156.30 | 0.880 |
| Metabasit | 2.82 | 4292.3 | 144.25 | 0.970 |
| Metabasit | 2.82 | 434.8 | 157.77 | 0.970 |
| Metabasit | 1.72 | 4568.8 | 91.96 | 0.147 |
| Metabasit | 1.72 | 4473.1 | 94.83 | 0.147 |
| Metabasit | 1.72 | 455.2 | 78.53 | 0.157 |
| Metabasit | 3.80 | 2979.1 | 171.78 | 0.880 |
| Metabasit | 3.80 | 359.9 | 27.40 | 0.880 |
| Metabasit | 4.37 | 45.7 | 149.34 | 0.158 |
| Metabasit | 4.37 | 4158.2 | 168.27 | 0.120 |
| Metabasit | 4.37 | 5218.1 | 173.5 | 0.128 |

References

- Abdulagatova Z, Abdulagatov IM, Emirov VN (2009) Effect of temperature and pressure on the thermal conductivity of sandstone. *Int J Rock Mech Min Sci* 46:1055–1071. doi:[10.1016/j.ijrmms.2009.04.011](https://doi.org/10.1016/j.ijrmms.2009.04.011)
- Archie G (1942) The electrical resistivity log as an aid in determining some reservoir characteristics. *Trans Am Inst Min Metal Eng* 146:54–62
- Barry-Macaulay D, Bouazza A, Singh RM, Wang B, Ranjith PG (2013) Thermal conductivity of soils and rocks from the Melbourne (Australia) region. *Eng Geol* 164:131–138. doi:[10.1016/j.enggeo.2013.06.014](https://doi.org/10.1016/j.enggeo.2013.06.014)
- Berryman JG (1995) Mixture theories for rock properties. In: Ahrens TJ (ed) *Rock physics and phase relations: a handbook of physical constants*. Am Geophys Union Washington, DC. 205–228, doi:[10.1029/RF003p0205](https://doi.org/10.1029/RF003p0205)
- Budiansky B, O'Connell RJ (1976) Elastic moduli of a cracked solid. *Int J Solids Struct* 12:81–97
- Fuchs S (2013) Well-log based determination of rock thermal conductivity in the North German Basin, PhD thesis. University Potsdam, Germany
- Fuchs S, Balling N, Foerster A (2015) Calculation of thermal conductivity, thermal diffusivity and specific heat capacity of sedimentary rocks using petrophysical well logs. *Geophys J Int* 203:1977–2000. doi:[10.1093/gji/ggv403](https://doi.org/10.1093/gji/ggv403)

- Gašior I, Przelaskowska A (2014) Estimating thermal conductivity from core and well log data. *Acta Geophys* 62(4):785–801. doi:[10.2478/s11600-014-0204-y](https://doi.org/10.2478/s11600-014-0204-y)
- Gegenhuber N, Schön J (2012) New approaches for the relationship between compressional wave velocity and thermal conductivity. *J Appl Geophys* 76:50–55. doi:[10.1016/j.jappgeo.2011.10.005](https://doi.org/10.1016/j.jappgeo.2011.10.005)
- Gegenhuber N, Steiner-Luckabauer C (2012) Vp/Vs automatic picking of ultrasonic measurements and their correlation of petrographic coded carbonates from Austria. In: 74th EAGE conference and exhibition, Copenhagen, Denmark
- Hartmann A, Rath V, Clauser C (2005) Thermal conductivity from core and well log data. *Int J Rock Mech Min Sci* 42(7–8):1042–1055. doi:[10.1016/j.ijrmms.2005.05.015](https://doi.org/10.1016/j.ijrmms.2005.05.015)
- Kukkonen IT, Peltioniemi S (1998) Relationship between thermal and other petrophysical properties of rocks of Finland. *Phys Chem Earth* 23:341–349
- Lichtenecker K (1924) Der elektrische Leitunswiderstand künstlicher und natürlicher Aggregate. *Physikal Zeitsch* 25(8):169–233
- Oezkahraman HT, Selver R, Işık EC (2004) Determination of the thermal conductivity of rock from P-wave velocity. *Int J Rock Mech Min Sci* 41(4):703–708. doi:[10.4236/gm.2013.34018](https://doi.org/10.4236/gm.2013.34018)
- Popov Y, Tertchnyi V, Romushkevich R, Korobkov D, Pohl J (2003) Interrelations between thermal conductivity and other physical properties of rocks: experimental data. *Pure Appl Geophys* 160:1137–1161
- Schoen JH (2011) *Physical properties of rocks: a workbook, handbook of petroleum exploration and production*, vol 8. Elsevier, Amsterdam
- Schoen JH (2015) *Physical properties of rocks: fundamentals and principles of petrophysics, developments in petroleum science*, vol 65. Elsevier, Amsterdam
- Sundberg J, Back P-E, Ericsson LO, Wrafter J (2009) Estimation of thermal conductivity and its spatial variability in igneous rocks from in situ density logging. *Int J Rock Mech Min Sci* 46(6):1023–1028. doi:[10.1016/j.ijrmms.2009.01.010](https://doi.org/10.1016/j.ijrmms.2009.01.010)

Magnetic investigation and $2\frac{1}{2}$ D gravity profile modelling across the Beattie magnetic anomaly in the southeastern Karoo Basin, South Africa

Christopher Baiyegunhi¹ · Oswald Gwavava¹

Received: 17 January 2017 / Accepted: 20 January 2017 / Published online: 1 February 2017
© Institute of Geophysics, Polish Academy of Sciences & Polish Academy of Sciences 2017

Abstract The southeastern Karoo Basin is considered to be one of the most prospective areas for shale gas exploration in South Africa. An interesting magnetic anomaly, the Beattie magnetic anomaly (BMA), and geologic intrusions are seen on the magnetic map. To date, the source of the BMA and interconnectivity of the igneous intrusions are not well understood. In this study, we investigate the interconnectivity of the igneous intrusions and possible location of the source of the BMA using gravity and magnetic methods. The gravity model results showed that igneous intrusions are interconnected at depth, which probably pose threat by increasing the risk of fracking the Karoo for shale gas exploration. The magnetic results revealed that the BMA becomes stronger with depth. The average depths to the top of the shallow and deep magnetic sources were estimated to be approximately 0.6 and 15 km, respectively.

Keywords Magnetic anomaly · Igneous intrusions · Gravity · Models · Karoo Basin

Introduction

Two distinctive geophysical anomalies, the Beattie magnetic anomaly (BMA) (Beattie 1909) and the Southern Cape Conductive Belt (SCCB) (De Beer et al. 1982) cut across the southeastern Karoo Basin. The BMA is the

largest known magnetic anomaly in the southern Karoo Basin; it was first discovered by Beattie (1909) and later documented by several researchers like Gough et al. (1973), Pitts et al. (1992), Weckmann et al. (2007a, b), Lindeque et al. (2007), and Stankiewicz et al. (2007). A regional aeromagnetic survey conducted by Fugro Airborne Surveys in 1972 over the Eastern Cape Province of South Africa also picked up this set of anomalies that possibly extend into Antarctica and southern South America. The magnitude of the anomaly ranges from about 200–500 nT, thus making it one of the world's largest magnetic anomalies (Weckmann et al. 2007a). The result of the geo-electrical and gravity geophysical investigation of the BMA by De Beer and Meyer (1983, 1984) revealed that the anomaly marks the southern boundary of the Namaqua–Natal Mobile Belt (NNMB). It is confined to the continental crust (Du Plessis and Simpson 1974). The upper and lower boundaries of the anomaly are due to two crustal discontinuities in the region that were identified at depths of about 7–11 and 17–19 km, respectively (Harvey et al. 2001).

The BMA spatially coincides with the central part of the 100–200 km wide electrically conductive zone known as the SCCB (Gough et al. 1973; Lindeque et al. 2007). They suggested that the BMA and SCCB have a common source and that the conductive material that trends east–west beneath the southern Karoo Basin and Cape Fold Belt lies in the crust. Both the BMA and SCCB anomalies are parallel to the tectonic edge or margin of the NNMB and the Cape Fold Belt (CFB). Harvey et al. (2001) and Lindeque et al. (2007) documented that the tectonic structures resulted from the continual addition or deposition of sediments over time. Schreiber-Enslin et al. (2013) also emphasized the importance of understanding the source of the BMA to ascertain or create a precise potential model of

✉ Christopher Baiyegunhi
201201530@ufh.ac.za; cbaiyegunhi@yahoo.com

¹ Department of Geology, Faculty of Science and Agriculture, University of Fort Hare, Alice, Eastern Cape Province, South Africa

the Karoo Basin and to know the evolutionary history of the Karoo Basin. With the use of Curie isotherm calculations, De Beer and Gough (1980) suggested that the BMA is possibly due to a source at about 25–38 km depth. Pitts et al. (1992) proposed that the BMA may be due to a magnetic body that extends from depths of 7–30 km and infer that serpentinised oceanic crust could be the source of the BMA. A recent magnetotelluric survey shows no indication or sign for such a body (Weckmann et al. 2007a).

Geophysical surveys (i.e., geo-electrical and gravity (De Beer and Meyer 1983), seismic reflection (Hälbich 1983; Hälbich 1993), near vertical seismic (Lindeque et al. 2007) and magnetotelluric (Weckmann et al. 2007a, b) across the western part of the BMA shows an anomaly source depth of 10–15 km within the mid-crust. It was earlier assumed that the BMA is due to the partially serpentinised oceanic lithosphere probably connected to a suture zone. However, Lindeque et al. (2011) as well as Schreiber-Enslin et al. (2013) suggested that the BMA is part of the tectono-metamorphic NNMB and associated shear zones. To date, due to insufficient geophysical information, the precise location of the source as well as the possible source for the BMA is still debatable. In this paper, we present the result of magnetic derivatives, power spectrum, depth slicing and simple gravity profile models that revealed the configuration of the basin and how geologic structures, such as dolerite sills and dykes, are interconnected at depth. In addition, we critically examined the gravity models to see any possible anomalous features that can be inferred to be caused by the BMA and SCCB sources.

Geological setting

The main Karoo Basin in South Africa covers up to 700,000 km² and represents about 100 Ma of sedimentation spanning from 280 to 180 Ma (Johnson et al. 2006). The deposition of Karoo sedimentary rocks that cover almost two-third of the area of South Africa occurred during the Late Carboniferous across the Gondwana, and span until the breakup of the Gondwana supercontinent during the Middle Jurassic (Catuneanu et al. 1998; Geel et al. 2013). The Karoo Basin is believed to have evolved from two distinct tectonic regimes sourced from the southern and the northern margin of Gondwana (Catuneanu et al. 1998, 2005). The southern tectonic regimes are believed to be related to the processes of subduction and orogenesis along the Panthalassan (paleo-Pacific) margin of Gondwana. In addition, it resulted in the formation of a retro-arc foreland system known as the “main Karoo Basin” in association with the primary subsidence mechanism represented by flexural and dynamic loading. The northern event was associated with extensional stresses that

propagated southwards into the supercontinent from the divergent Tethyan margin of Gondwana. Superimposed on the tectonic control on basin development, climate fluctuations also left a mark on the stratigraphic record. In addition, it shows evidence of a general shift from cold and semi-arid conditions during the Late Carboniferous—Earliest Permian interval, to warmer and eventually hot climates with fluctuating precipitation during the rest of the Karoo time (Keyser 1996 in Catuneanu et al. 2005).

The importance of the tectonic control on Karoo Basin development and sedimentation was first proposed by Rust (1973) and subsequently modified by several researchers (i.e., Tankard et al. 1982; Smith et al. 1993; Veevers et al. 1994; Johnson et al. 1996). However, Visser (1987) and Catuneanu et al. (1998, 2002, 2005) suggested the evolution of the retro-arc foreland Karoo Basin in relation to fold-thrust belt inboard of magmatic arc. Several authors like De Wit and Ransome (1992), Veevers et al. (1994), Visser (1995) and Catuneanu et al. (1998) envisaged that the Karoo Basin is a retro-arc foreland basin. However, Tankard et al. (2009) gave an alternative interpretation for the tectono-sedimentary evolution of the Karoo Basin by dividing it into a pre-foreland phase and a foreland phase. The Cape Supergroup (Cape basin) spans from the Early Ordovician to Early Carboniferous (Veevers et al. 1994), and consists of up to 8 km of shallow marine, deltaic and fluvial deposits that thicken southwards into an east–west trending depo-axis (Rust 1973; Turner 1999).

According to Tankard et al. (1982), these sediments were believed to have been derived from a cratonic source to the north. The overlying Karoo Supergroup (Karoo Basin) that spans from the Late Carboniferous to Early Jurassic period and consists of about 12 km of deep marine to fluvial deposits. Visser and Praekelt (1996) documented that the Karoo-Falklands basin, the Sierra Australise Colorado basins (Argentina), as well as the Central Antarctic Mountains basin were developed as extensional back-arc basins, in relation to the oblique subduction of the paleo-Pacific plate beneath the western Gondwana. Several major strike-slip systems begin to develop, with the Southern Trans-African Shear system, within the Damara mobile belt. According to Stollhofen et al. (2000), all the Karoo basins in southern Africa can be attributed to the extensional intra-cratonic rifts, in relation to the N–S trending basement shear zones. The detailed record for southern Africa is still less complete when compared to South America and Antarctica, as a result of the major Mesozoic regional strike-slip and extensional tectonics that removed most of the late Palaeozoic record to the south of the present day South Africa (Flint et al. 2011).

Structural elements/features, such as lineaments, faults and intrusions, are present in the study area, as depicted in Fig. 1. Chevallier and Woodford (1999) allude that the

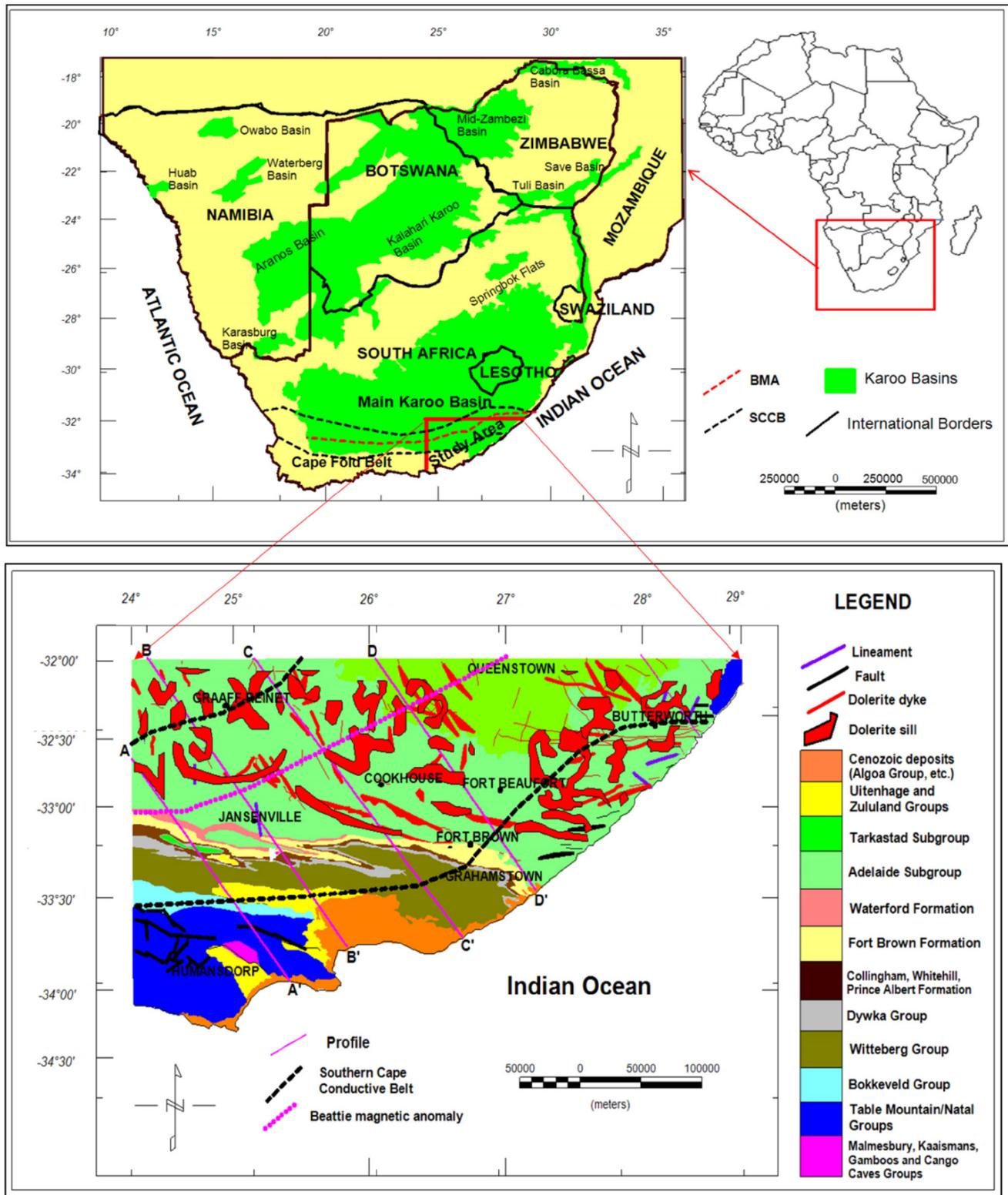


Fig. 1 Geological map of the study area. The map also shows dolerite intrusions (sills and dykes), faults, part of the BMA and SCCB (after Council for Geoscience 1995)

Karoo dolerite, which comprises several petrological facies (ranging from a leucogabbro to a dolerite-pegmatite), consists of interconnected networks of dykes and sills and it is very difficult to single out any particular intrusive or tectonic event due to the fact that, an individual sill can be fed by many dykes of different orientations or a dyke can act as a feeder to two different sills or more. Chevallier and Woodford (1999) and Svensen et al. (2007) concluded that during the intrusive phases, the molten magma concurrently filled in the numerous fractures, and that the dolerite intrusive network probably behaved as a shallow stockwork-like reservoir or storage system where molten magma of different viscosities intruded the fractures. It appears that there is a lithological control on the emplacement of dykes within the Western Karoo Basin, as the bulk of the dykes are strata bound and concentrated in the Upper Ecca and Beaufort Group (Woodford and Chevallier 2002). According to Chevallier et al. (2001), three major structural domains that are indicated by dyke distribution have been identified in the Main Karoo Basin (Fig. 2). These domains are:

- The Western Karoo Domain: It extends from Calvinia to Middelburg and is characterised by two distinctive structural features, east west trending zone of long and thick dykes associated with right lateral shear deformation and north northwest dykes.

- The Eastern Karoo Domain: It extends from Middelburg to East London and comprises two major dyke swarms, namely, a major curvi-linear swarm of extensive and thick dykes diverging from a point offshore of East London and minor north–northeast trending dykes.
- The Transkei–Lesotho–Northern Karoo Domain: It consists of two swarms, northwest trending dykes in the Transkei Region, curving to east–west in the Free State and northeast trending dykes mainly occurring within and alongside the Lesotho basalt.

Stratigraphically, the Karoo Supergroup can be subdivided into five groups, namely, in their ascending stratigraphic order, the Dwyka (Late Carboniferous), Ecca (Early Permian), Beaufort (Late Permian—Middle Triassic), Stormberg (Late Triassic—Early Jurassic), and Drakensberg Groups (Middle Jurassic) (Johnson et al. 2006; Table 1). Johnson et al. (2006) documented that prior to the Middle Jurassic period, the environmental conditions in the Karoo Basin changed rapidly, coupled with the large movement of the Earth's mantle, the crust experienced a large-scale lifting, and eruption of massive volumes of basaltic lava that constitute the Drakensberg Group. The outpourings of the basaltic lava spread across much of Gondwana

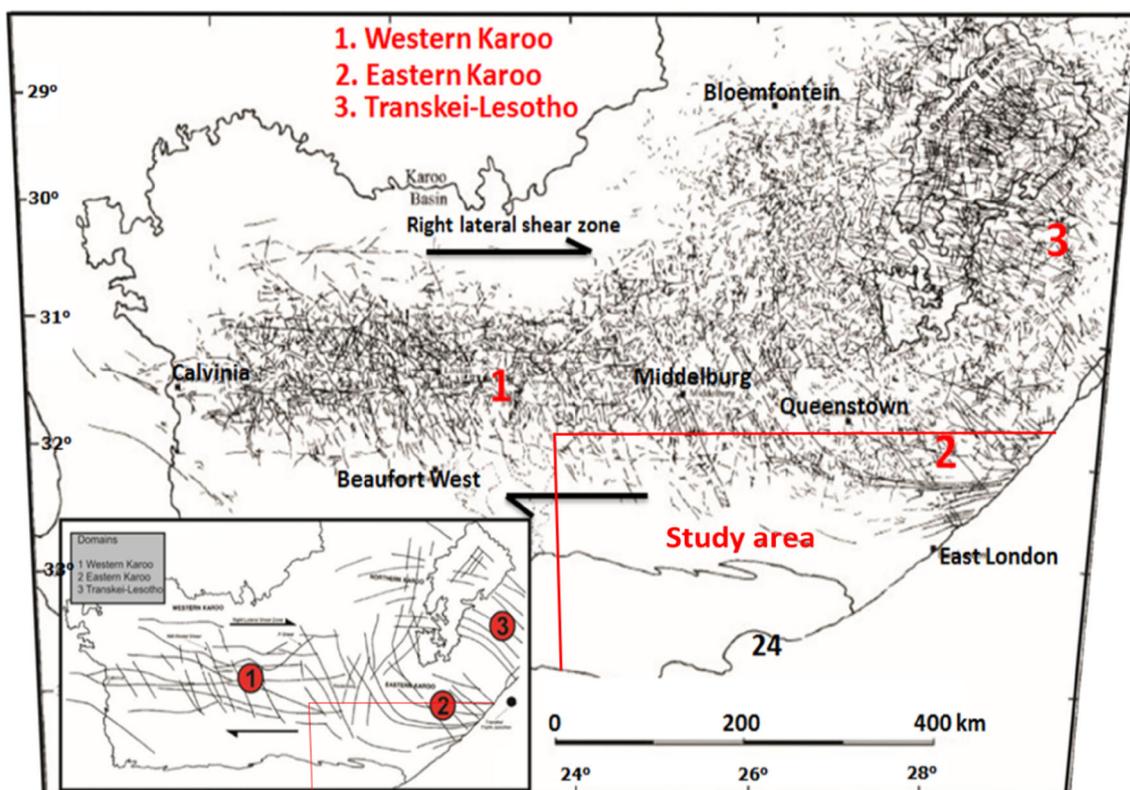


Fig. 2 Dolerite dykes of the main Karoo Basin. *Inset* is a simplified structural map showing the three structural domains (after Woodford and Chevallier 2002)

Table 1 Lithostratigraphy of the Karoo Supergroup in the Eastern Cape Province (Johnson et al. 2006)

| Supergroup | Group | Subgroup | Formation | Member | Lithology | Maximum thickness (m) | | |
|------------|--------------------------------------|----------------------|----------------------|---------------------------|----------------------|--------------------------------|----------------------|----|
| Karoo | Beaufort | Tarkastad | Drakensberg | | Basalt | 1400 | | |
| | | | | | Pyroclastic deposits | | | |
| | | | Clarens | Sandstone | 300 | | | |
| | | | | Elliot | Red mudstone | 500 | | |
| | | | Molteno | Sandstone | | | | |
| | | Coarse sandstone | | 450 | | | | |
| | | Grey and khaki shale | | | | | | |
| | | ECCA | Tarkastad | Burgersdorp | | Red mudstone | 1000 | |
| | | | | | | Sandstone light grey sandstone | | |
| | | | | Katberg | | Grey shale | | |
| | | | | | Light grey sandstone | 900 | | |
| | | | | Red mudstone | | | | |
| | | | | Grey shale | | | | |
| | Adelaide | | | Balfour | Palingkloof | | Red mudstone | 50 |
| | | | | | | | Light grey sandstone | |
| | | | | | Elandsberg | Sandstone | 700 | |
| | | | | Siltstone | | | | |
| | Barberskrans | | Light grey sandstone | 100 | | | | |
| | | | Khaki shale | | | | | |
| | Daggaboersnek | | Grey shale | 1200 | | | | |
| | | Sandstone | | | | | | |
| | | Siltstone | | | | | | |
| | | Oudeberg | Light grey sandstone | 100 | | | | |
| | | Khaki shale | | | | | | |
| | Middleton | Grey and black shale | 1500 | | | | | |
| | | Light grey sandstone | | | | | | |
| | | Red mudstone | | | | | | |
| | Koonap | Grey sandstone shale | 1300 | | | | | |
| | Waterford (not present in ECCA Pass) | Sandstone shale | 800 | | | | | |
| | Fort brown | Shale sandstone | 1500 | | | | | |
| | Ripon | Sandstone shale | 1000 | | | | | |
| | Collingham | Grey shale | 30 | | | | | |
| | | Yellow claystone | | | | | | |
| | Whitehill | Black shale chert | 70 | | | | | |
| | Prince Albert | Khaki shale | 120 | | | | | |
| | DWYKA | | | Diamicite, tillite, shale | | | | |

about 180 million years ago, indicating the start of Gondwana breakup. In this paper, the stratigraphy, geological and tectonic setting of the southeastern Karoo Basin are not reported/documentated in detail but can be found in Hälbich (1993), (Hälbich 1983), de Wit and Ransome (1992), Veevers et al. (1994), Visser (1995), Catuneanu et al. (1998, 2002, 2005), Johnson et al. (1996, 2006), De Wit and Horsfield (2006), Tankard et al. (2009, 2012), and Pángaro and Ramos (2012).

Materials and methods

Locations within the Eastern Cape Province where the Karoo Supergroup (Dwyka, Eccla, and Beaufort Groups) (Fig. 1; Table 1) as well as the Cape Supergroup (Table Mountain, Bokkeveld and Witteberg Groups) (see Fig. 1) outcrop were visited for field investigation and sampling. A total of 258 samples were collected on road cut exposures from the geologic formations (Fig. 1) that cover the study area for density measurements.

Determination of dry density

The dry densities of rock samples were obtained from laboratory measurements utilizing the buoyancy determined volume that uses Archimedes' principle. The rock samples were left in the sun for about two weeks, with an average daily temperature of about 25 °C. The samples were dried long enough to remove any moisture from the voids. An average water density value of 1.022 g/cm³ was obtained and applied when calculating the dry density of the rock samples. The water density values depend on its temperature and pressure, i.e., decrease in temperature results in the water molecules to be well and closely packed together, thereby leading to increase in the density of water and vice versa, but this relationship is not linear. The temperature of water was monitored throughout the experiment and the effect was small or negligible, giving no appreciable change in the density. It could be inferred that the temperature and pressure were fairly constant throughout the experiment period. The dry density of the rocks was calculated from the expression:

$$\rho_d = \left[\frac{dM_a}{dM_a - dM_b} \right] \times \rho_w \quad (1)$$

where ρ_d = dry density, dM_a = mass of dry sample in air, dM_b = mass of sample in water and ρ_w = density of water (determined using a density bottle).

The average density value for each geologic group was obtained from the measured dry density.

Available geophysical data

Aeromagnetic and gravity data in the format of Geosoft files were supplied by Fugro Airborne Surveys and Council for Geoscience, respectively. The supplied magnetic data had all the temporal variation and IGRF removed, to leave the crustal magnetic field. Similarly, the gravity data had been reduced to Bouguer gravity values.

Geophysical data enhancement and filtering

The acquired magnetic original space domain grid was prepared for filtering and enhanced using Geosoft Oasis montaj. The Fast Fourier Transform (FFT) was applied to transform the square and periodic space domain to the wave number domain, and then filters were applied to enhance the effects of shallow magnetic sources (such as removal of first-order trend) or enhance deeper magnetic sources. The inverse FFT was applied, enhancements (i.e., vertical derivative, horizontal derivative, analytical signal and radially averaged spectrum) were calculated, and the results are presented in the form of geophysical maps

(“Reduction to the pole (RTP)”, “Vertical derivative”, “Total horizontal derivative”, “Analytical signal”, “Power spectrum”, sections). Magnetic depth slicing was done using the GETECH GETGRID (“Depth slicing” section). The gridded gravity data was enhanced to specifically remove the first order trend in geosoft. No filtering was carried out on the gravity data; the map is presented in “Gravity” section.

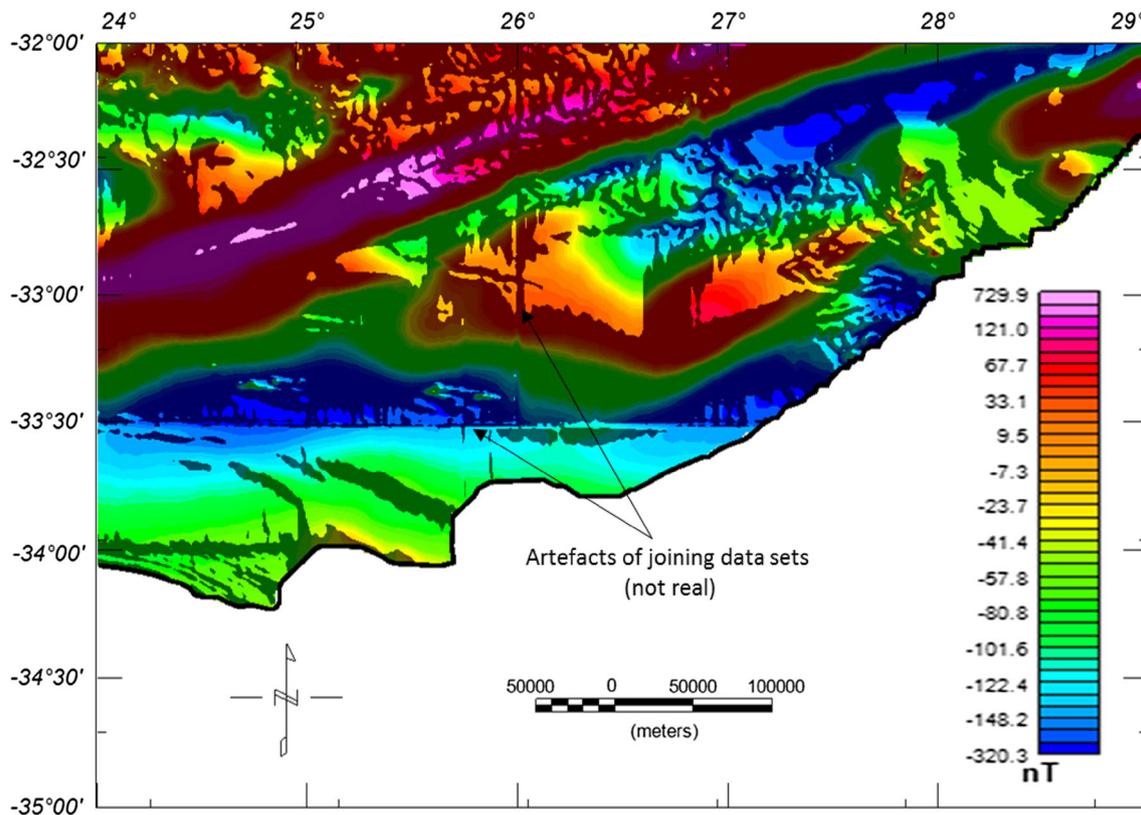
½ D gravity profile modelling

Four profiles (A–D) that are roughly equally spaced to cover the study area were selected (Fig. 1). The geological map of the study area was used to measure the actual profile lengths and positions along the profiles to each geologic group (succession) outcrop within the study area and at infinity (outside the study area but along the profile). This was done to set constraints to the model such that the model will perfectly fit or agree with what is obtained in the real-world. These profiles were modelled and for each profile, 3 models were produced using the minimum, average and maximum density values giving a total of 12 models that were analysed. A starting GM-SYS model was created from a map profile (e.g., A–A') by importing the gravity grid and elevation grid. The estimated thickness (length) and position of each group from the geological map were used to develop the starting model. Each starting model was extended to infinity at both ends (i.e., a long distance of ±500 km in this case) to eliminate edge effects. The geologic models for the selected profiles were created and real-time calculations of the gravity response of a specific Earth model were performed. The modelling was constrained by the stratigraphic thickness (measured in the field and data from boreholes CR1/68, SP1/69, SC1/67 and SC3/67, Table 2), elevation and average density of rocks from various formations of the Karoo Supergroup that outcrop in the study area.

Several authors, like Tedla et al. (2011), Stankiewicz and de Wit (2013) have indicated a Moho depth in the range of 44–50, 40–50, and 40–45 km, respectively, in the study area. To cover the depth envisaged for the Moho by the authors, the depth to the Moho was originally set to vary from about 35 up to 60 km during the modelling process, such that the model will establish the best fit depth for the Moho. The thickness of the modelled dolerite intrusions were estimated from literature (Chevallier and Woodford 1999; Chevallier et al. 2001) and the measured thickness of the intrusions from the boreholes were also used to constrain the models. The modelled dolerite sills and dykes were originally set in horizontal and vertical positions, and allowed to vary during the modelling process to have the best fit model. Since a total of 12 models were

Table 2 Summary of borehole data extracted from Scheiber-Enslin et al. (2014)

| Well | Thickness (m) | | | |
|--------|----------------|------------|-------------|-----------------|
| | Beaufort group | Ecce group | Dwyka group | Cape supergroup |
| CR1/68 | 3600 | 1700 | 450 | >850 |
| SP1/69 | 750 | 1250 | 70 | – |
| SC1/67 | 2300 | 1600 | >100 | – |
| SC3/67 | 1900 | 1550 | 45 | – |

**Fig. 3** Reduced to the pole magnetic residual anomaly map

obtained, only those of the average density values are presented in this paper. The alternative models were used to check the sensitivity of the models with respect to change in layer densities and thicknesses.

Results and discussion

Reduction to the pole (RTP)

RTP was performed on magnetic data to remove the asymmetry caused by the inclined main field or convert it to a symmetrical shape (Fig. 3).

The RTP magnetic map (Fig. 3) was overlain on the digitized geological map of the area (Fig. 4) to correlate dolerite intrusions (i.e., dykes and sills) and any other magnetic sources with the magnetic anomalies.

Figure 5 shows three main magnetic anomalies. The first anomaly of up to 729.9 nT extends from west of Jansenville (long. 24°E, lat. 33°S) to north of Fort Beaufort (long. 27°E, lat. 32°S). It is part of the BMA and could be due to a buried body with high magnetic susceptibility minerals. The centre of the first anomaly also shows that the body divides into two, but still trend is in the NE–SW direction. The second anomaly, in the north-eastern part of Butterworth (long. 29°E, lat. 33.5°S) is of relatively small extent with maximum magnetic value of about 227.5 nT. It could be due to dolerite intrusions with associated faulting that are more prominent in the northeastern part of the map. The third anomaly is the “bean shape anomaly” and it coincides with a continuous, regional band of high magnetic susceptibility seen in the centre-eastern part of the map.

Fig. 4 Digitized geological map of the study area

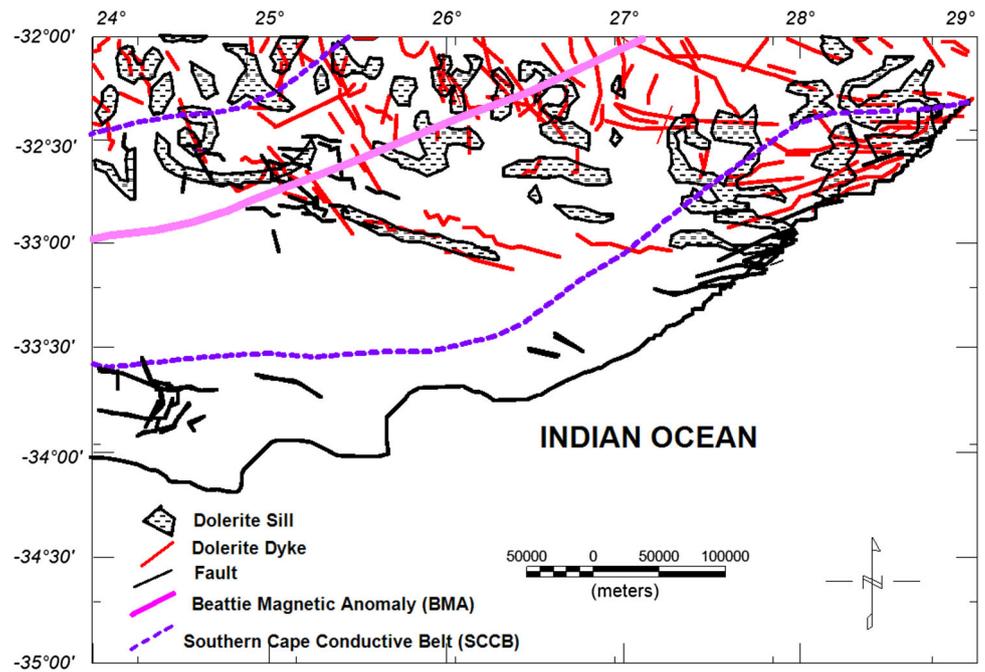
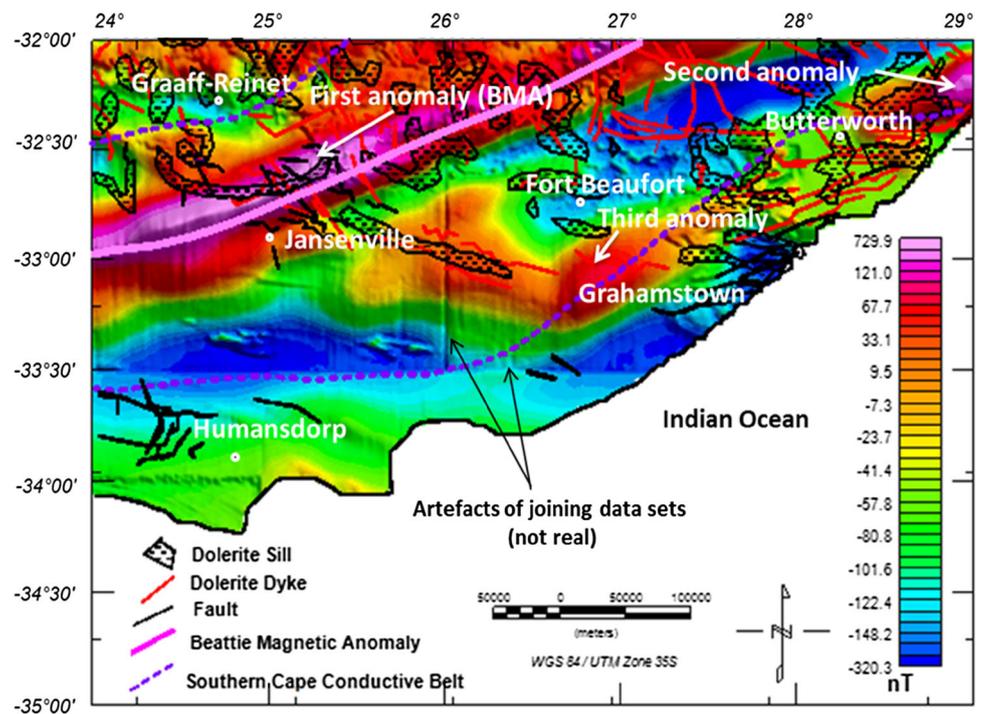


Fig. 5 Reduced to the pole (RTP) magnetic residual anomaly map overlain on the geology



Vertical derivative

The first vertical derivative (Fig. 6) was calculated to enhance the effects of near-surface geology (shallow magnetic sources), suppress anomalies caused by deeper sources, thereby giving a better resolution of closely spaced sources.

The first vertical derivative magnetic map (Fig. 6) was overlain on the digitized geological map (Fig. 4) as depicted in Fig. 7. The splitting of the BMA from long. 25.5°E, lat. 32.5°S to long. 27°E, lat. 32°S is now clearly visible. In addition, the ring structures and linear features due to the dolerite dykes and sills are clear in Fig. 7.

Fig. 6 First vertical derivative magnetic map

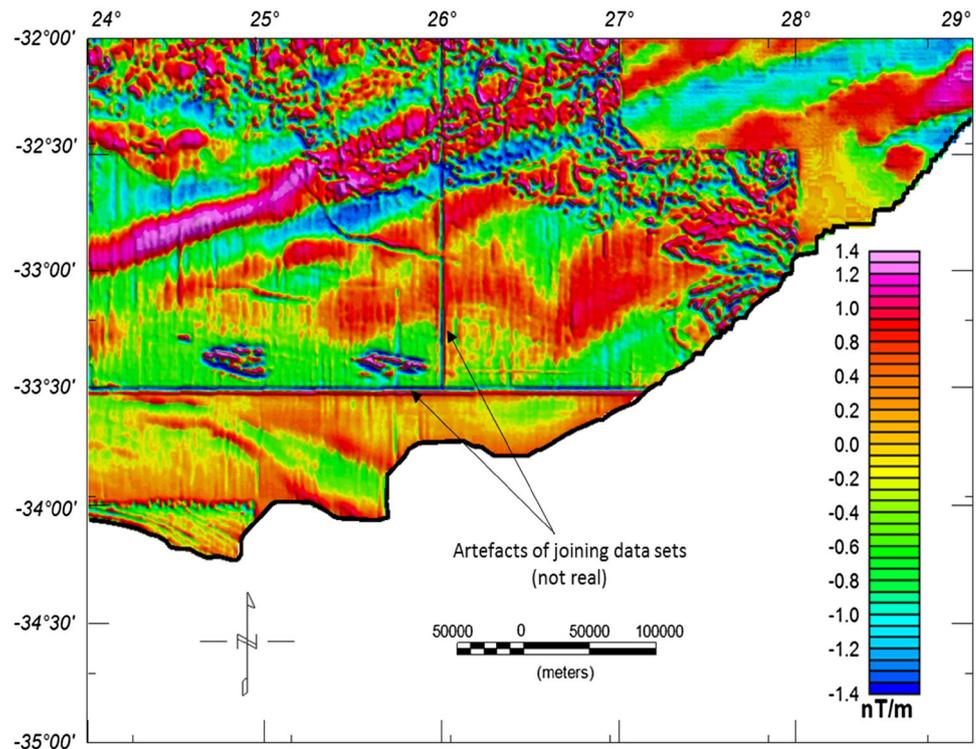
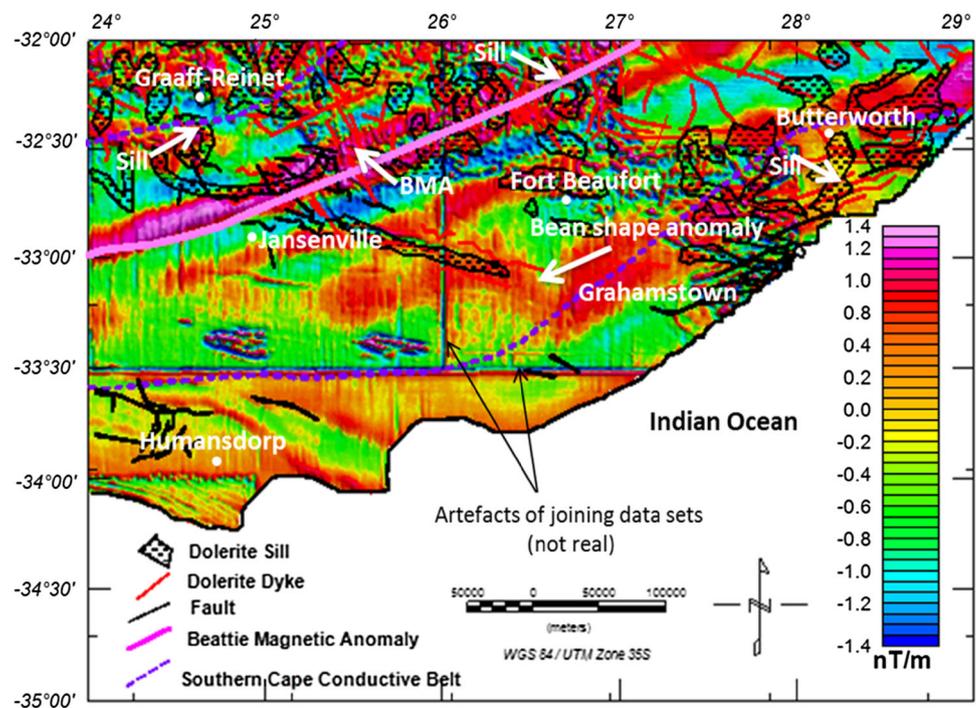


Fig. 7 First vertical derivative magnetic map overlain on the geology

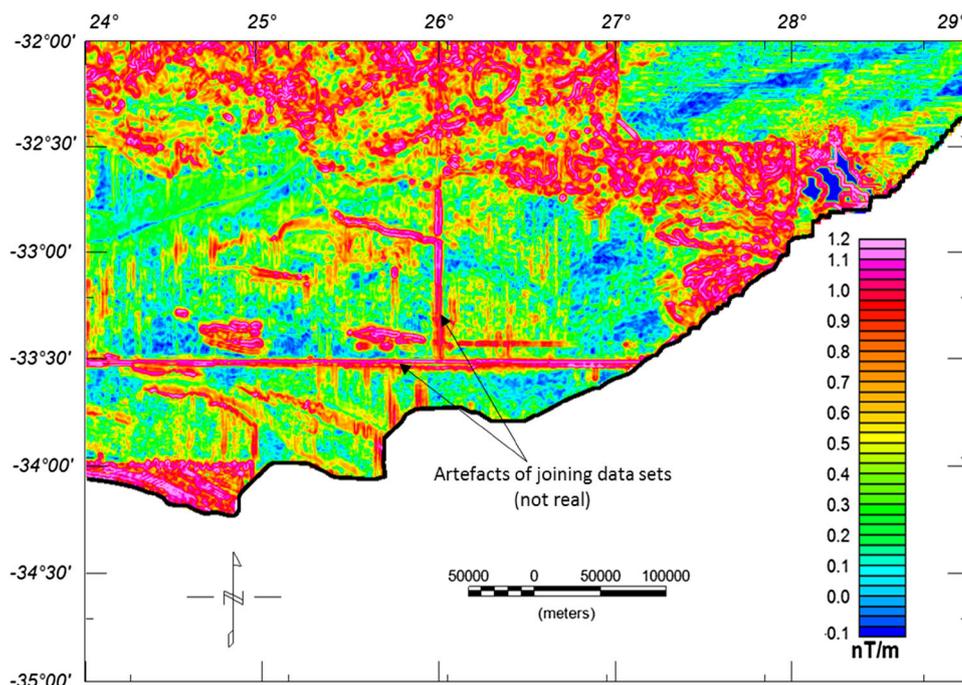


Total horizontal derivative

The first total horizontal derivative of the magnetic field data was calculated in the space domain (Fig. 8). It gives an indication of the boundary structure and this will be

accurate for pole reduced data. The first vertical derivative magnetic map (Fig. 8) was also overlain on the digitized geological map (Fig. 4) in order to correlate dolerite intrusions (i.e., dykes and sills) and any other magnetic sources with the magnetic anomalies. The overlain first

Fig. 8 First total horizontal derivative magnetic map



total horizontal derivative map (Fig. 9) shows interesting ring-like and linear features that extend from north of Graaff-Reinet to the northeastern part of Grahamstown. These features also coincide with the dolerite intrusions on the geological map.

Analytical signal

The analytical signal magnetic map (Fig. 10) was calculated to produce features that are independent of the inclination of the main field and magnetisation direction (either induced or remanent magnetisation direction). This implies that if the geometry of the magnetic sources or bodies is the same, then they will have the same analytical signals. Li (2006) stated that the features produced in 3D case are dependent on the magnetisation direction but when the magnetic source is more than one, the features that are produced on the analytical signal map are effects of shallow sources.

Figure 11 shows highs that clearly outline the BMA, dolerite intrusions and a number of faults. The dolerite intrusions are concentrated in the northern part, while most of the faults are concentrated toward the centre and eastern part (edge) of the map.

Power spectrum

Depth estimation from the spectrum of potential field data (magnetic or gravity) works on the principle that the measured potential field data at the surface is an integral

part of the source signature from different depths (Rabeh et al. 2008). The radially averaged power spectrum was calculated in the wavenumber domain and implemented in Geosoft. The radial power spectrum is averagely calculated such that the effect of noise is minimised, ensuring a more accurate result than other commonly used methods (Spector and Grant 1970). The radially averaged power spectrum of the field increases with decrease in depth (h) by a factor that is proportional to exponential ($-4\pi hk$), where k is the wavenumber. Maus and Dimri (1996) stated that, if the depth factor dominates the shape of the power spectrum, then the logarithm of the power spectrum must be proportional to $-4\pi h$. Thus, the slope of the log of the radially averaged spectrum can be used directly to estimate the depth to source.

Quite often, the power spectrum has a linear segment and, for each segment, a depth can be calculated since the slope is $-4\pi h$. Two linear segments can be fitted in Fig. 12 and this corresponds to a depth of 0.6 km (shallow sources) and 15 km (deeper sources).

Depth slicing

Depth slicing entails the use of linear filters to isolate (i.e., based on wavelength criteria) anomaly contributions to a map derived from source bodies in a certain depth range. It creates separate magnetic maps for a sequence of layers at a particular depth, thus allowing the comparison of amplitude response and recognition of the disappearance and arrival of magnetic units at different depth. Depth

Fig. 9 First total horizontal derivative magnetic map overlain on the geology

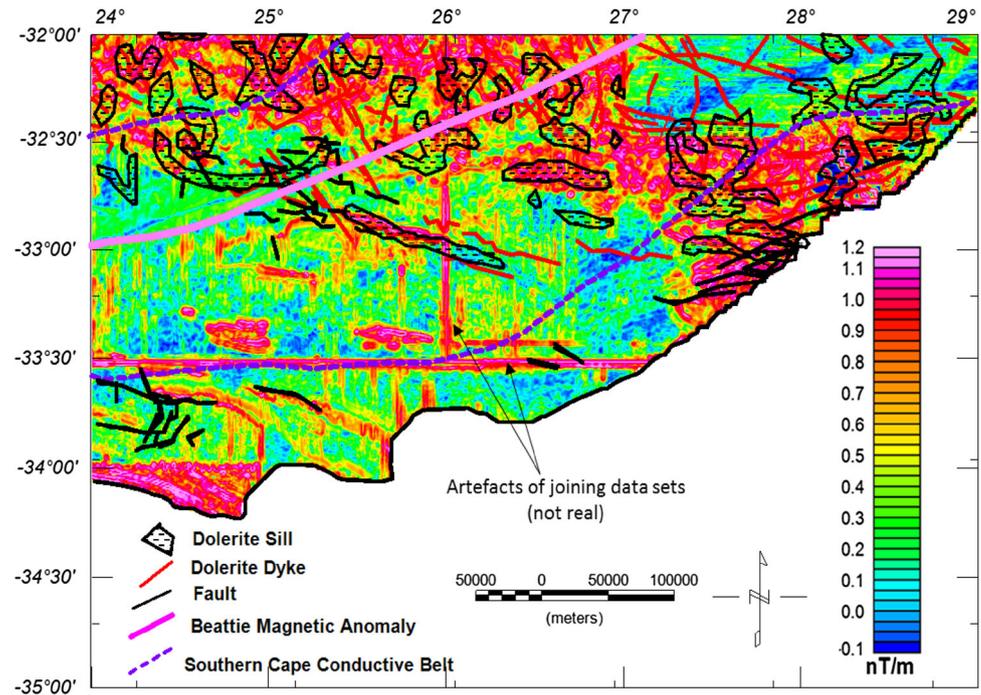
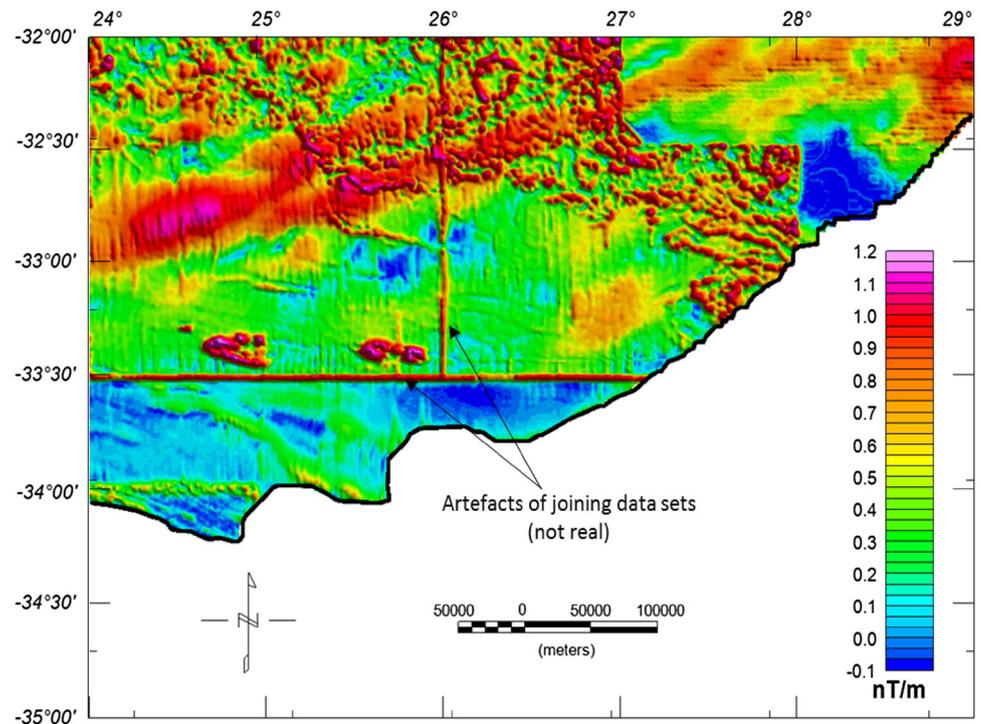


Fig. 10 Analytical signal magnetic map of the study area



slicing works on the principle of Wiener filtering, which assumes that potential field signals (e.g. magnetic and gravity) results from two or more uncorrelated arbitrary processes (Fuller 1967). Fuller (1967) stated that the various components of the data can be separated by highlighting the effects of shallow sources from deeper sources. Most magnetic anomalies at the surface usually originate at

a shallow depth, while obscure magnetic anomalies usually originate from a deeper depth (Gunn 1972).

Depth slices (Fig. 13) were used to compare the anomaly amplitude responses at different depths. Thus, anomalous features at a particular depth can be isolated and recognised. Depth slice results (slices 1–3) clearly show structural trends that coincide with dolerite intrusions on

Fig. 11 Analytical signal magnetic map overlain on the geology

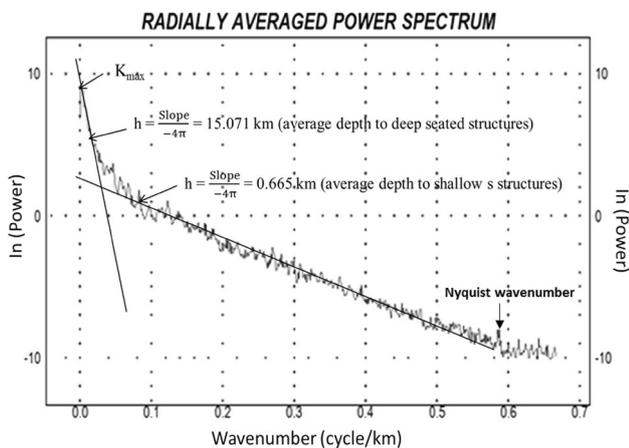
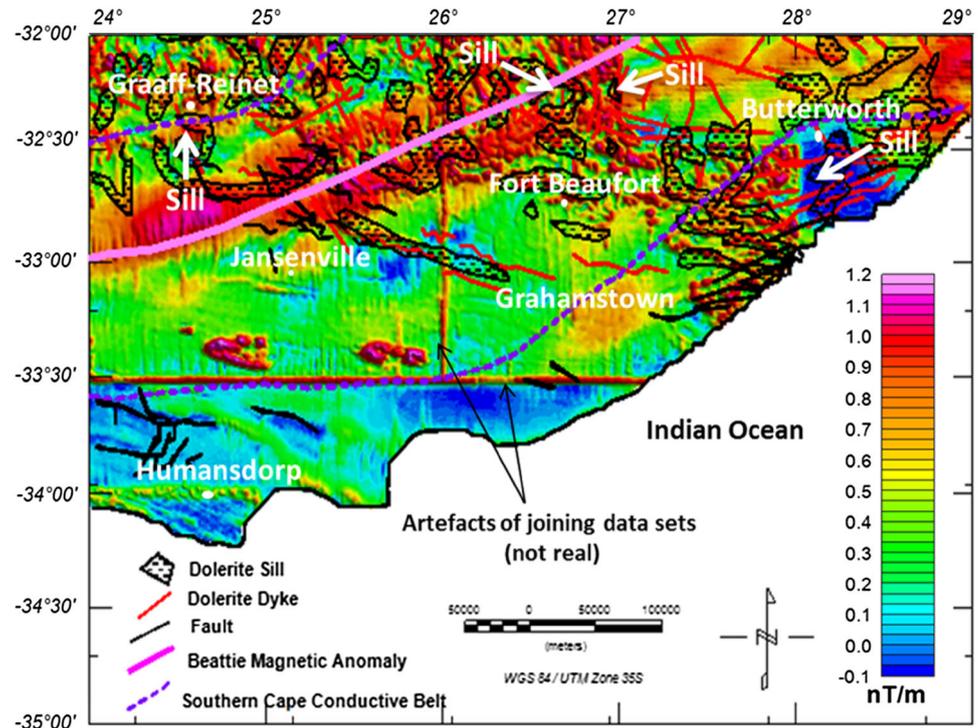


Fig. 12 Depth estimates from the radially averaged power spectrum

the geological map of the area. It shows that the mapped geologic structures such as dolerite intrusions have a response which is clearly visible down to a depth of about 3.3 km. The depths slice 4 (15.2 km) shows no anomalies due to intrusions when compared with the other slices (slice 1–3). A ring structure (inferred dolerite sill) seen in the northern part of the map becomes more prominent or stronger up to a depth of about 3.3 km, but thereafter it disappears at a depth of 15.2 km. The BMA appears in all the depth slices (0–15.2 km) and becomes broader and stronger up to a depth of 15.2 km. This could be an indication that the source is deeper, possibly within the basement. It could be a buried body in the basement. The bean

shape anomaly (BSA) also behaves in a similar manner such as the BMA, so it could possibly have source(s) at similar depths to the BMA.

Gravity

The Bouguer gravity anomaly map (Fig. 14) shows an increase in gravity values from about -141 mGal inland (southwestern and northern side of Jansenville) to high gravity values of up to 62 mGal in coastal areas (western part of Humansdorp and continues to the northeast of Butterworth along the coastline). The dominant gravity highs seen along the coastal areas are of long wavelength. This is possibly due to a deeper source/interface inland that shallows southward towards the coast, e.g., the basement and/or Moho (this is seen on models presented in Figs. 15, 16, 17, 18).

Dry density

The calculated densities of the rock samples are listed in Table 3.

The average dry density for each geologic group falls within the range (2.3 – 2.8 g/cm³) which is similar to the values obtained by several researchers, like Van der Voort (2001) and Johnson et al. (2006) that investigated the Karoo Basin in South Africa. The dry densities of rocks in the upper mantle (below the Moho) were extracted from literature (Hynes and Snyder 1995; Cook et al. 2010;

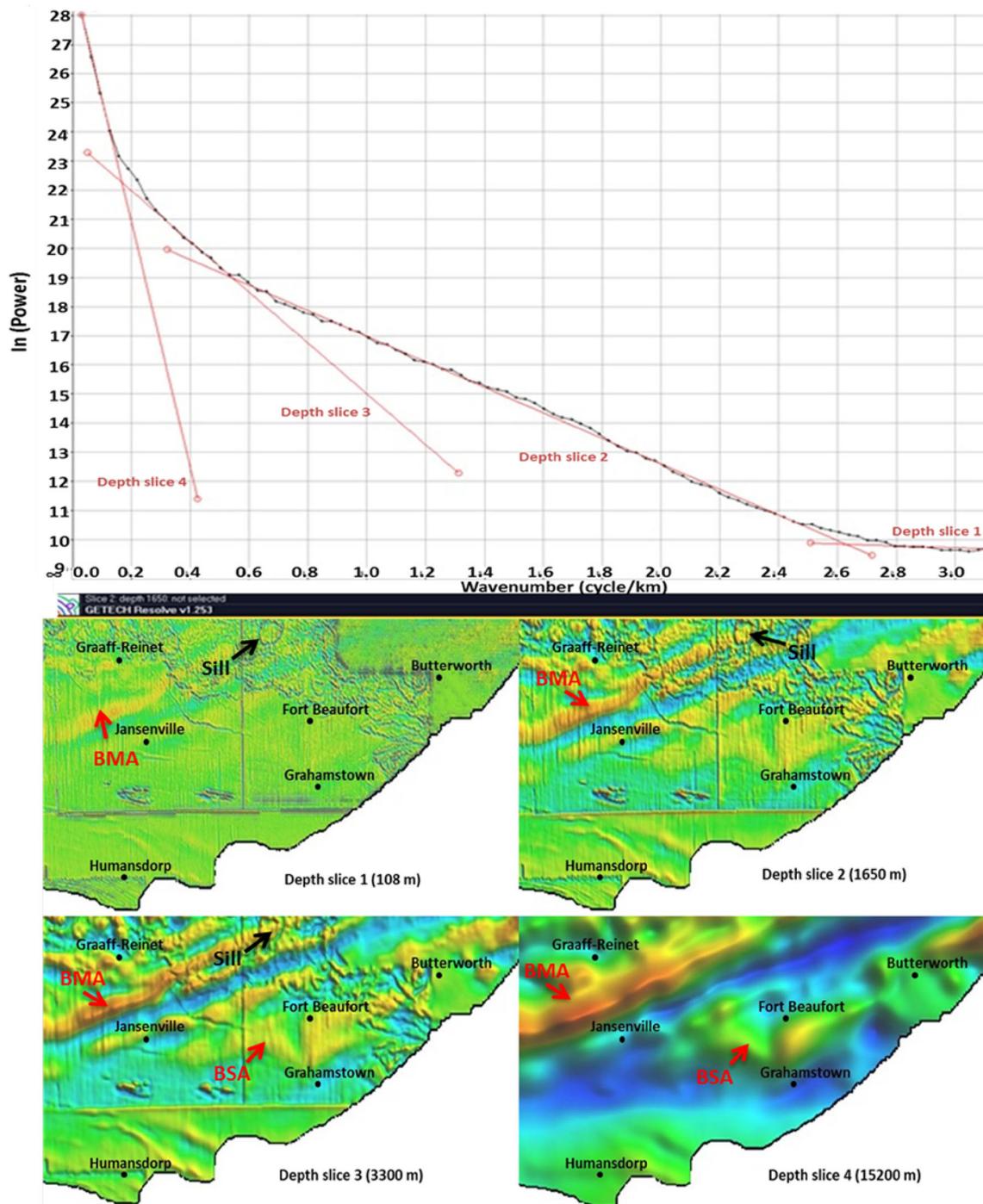


Fig. 13 Depth slice showing changes in anomaly features with depth. *Note* BMA is the Beattie Magnetic Anomaly and BSA is the bean shape anomaly

Mjelde et al. 2013; Thybo and Artemieva 2014) and the density values range from 3.14 to 3.4 g/cm³.

½ D gravity profile models

The gravity modelling result of profile A–A' is shown in Fig. 15. There is a relatively low gravity zone in the

northwestern–southeastern part of the map, as shown along the profile. The minimum gravity value along this profile is about –190 mGal, whilst the maximum value is –24 mGal. The thickness of the Beaufort, Ecca and Dwyka Groups varies along this profile. The Karoo (Dwyka, Ecca, and Beaufort Groups) and Cape Supergroup (Witteberg, Bokkeveld, and Table Mountain Group) have

Fig. 14 Bouguer gravity anomaly map on the geology

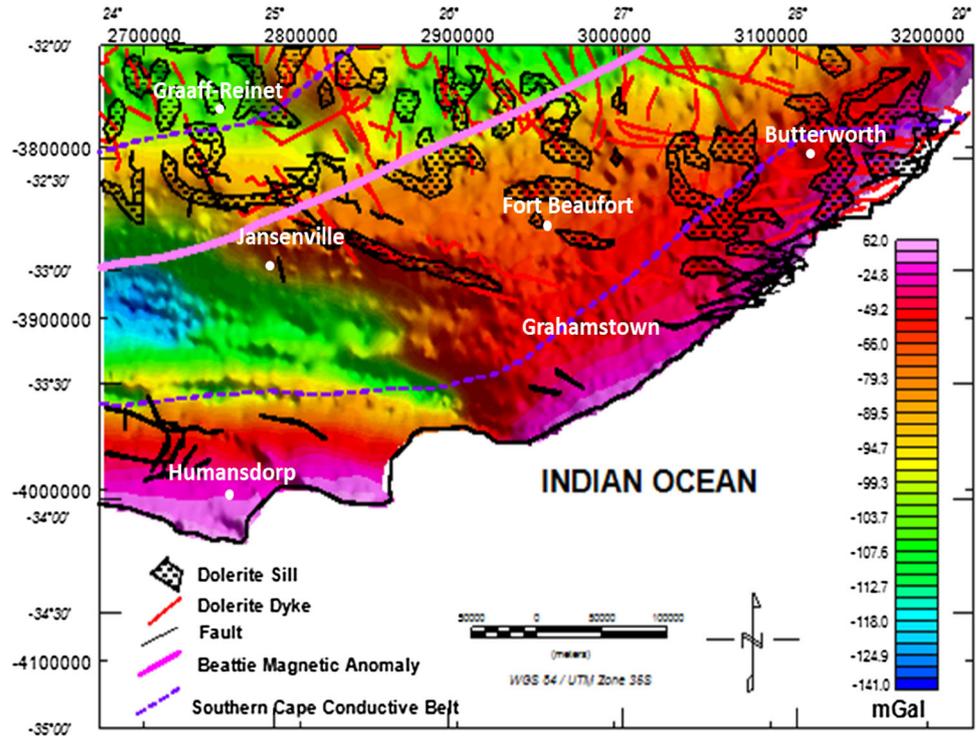


Fig. 15 Gravity modelling of profile A–A'. VE = 0.8; initial and final RMS errors are 64.4 and 1.7, respectively. Note: the shape of the BMA is unknown

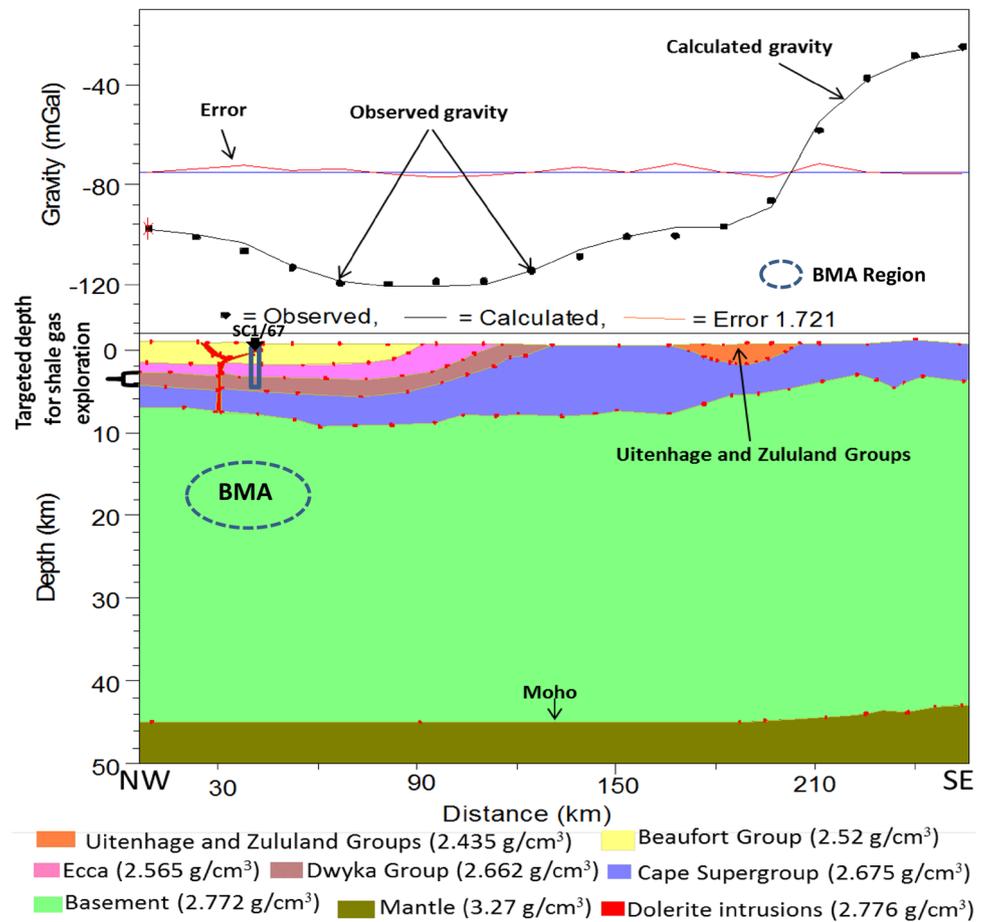
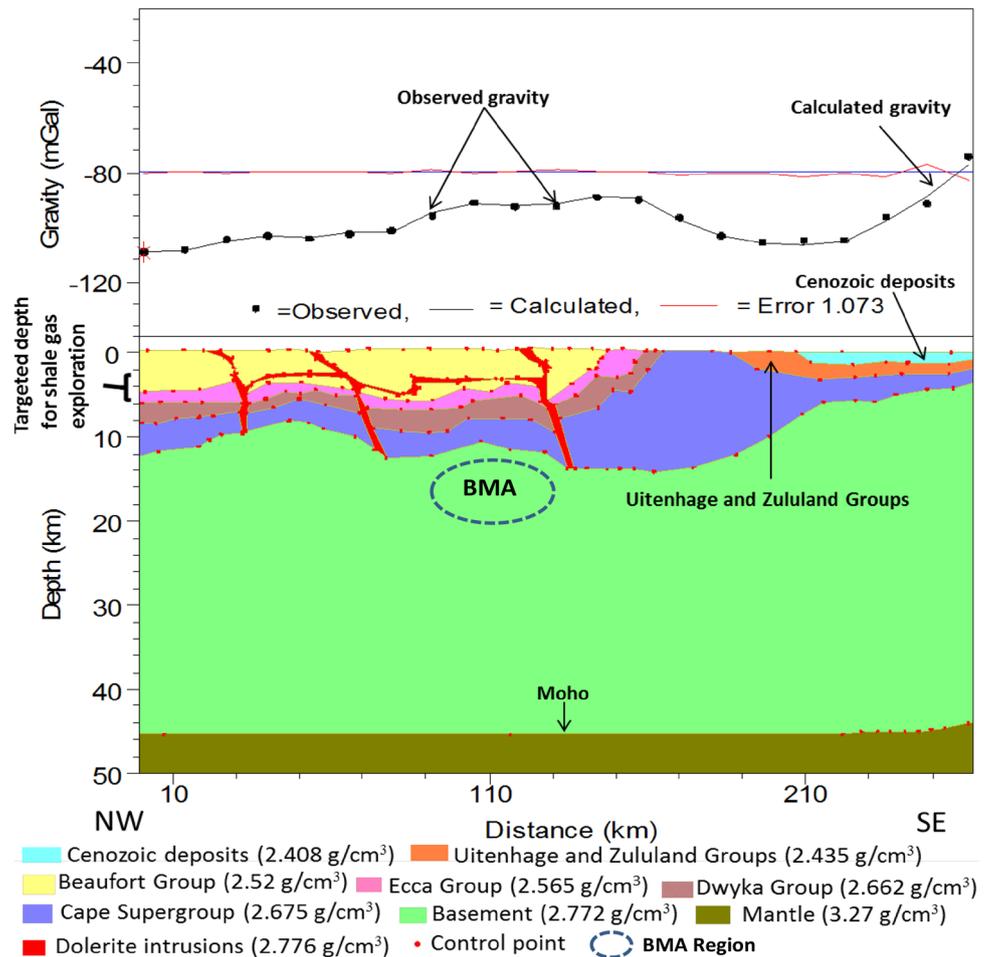


Fig. 16 Gravity modelling of profile B–B'. VE = 1.4; initial and final RMS errors are 110 and 1.2, respectively. Note: the shape of the BMA is unknown



been intruded by a dolerite dyke as shown in Fig. 15 which probably served as feeder to the sill that is seen in the Beaufort Group as well as outcropping on the surface. On the geology map, there are two places where dolerite intrusions are seen and these outcrops, as indicated in the model, merged at a depth of about 1.8 km. In addition, the Uitenhage and Zululand Groups (outcropping on the surface) are encountered along this profile between the distances of about 160–200 km, and they extend to about 1.8 km below the sea level.

The Bouguer gravity profile in Fig. 12 shows a remarkable gravity low towards the southeastern end of the profile (i.e., from 150 to 250 km). The layers in the model show undulations inferred to be due to deformation. In addition, the geologic sequence has been intruded by dolerites in the form of dykes and sills. The dykes probably served as feeders to the intrusions that outcrop on the surface. The Cenozoic deposits as well as the Uitenhage and Zululand Groups were seen along the profile (between the distance of 190 to about 240 km) and their thicknesses (i.e., up to 2.5 km thick) also varies across the profile. The surface expression (region) of the BMA along this profile is

located between the distances of about 90–130 km. In the BMA region, there is no significant gravity anomaly noticeable on the profile. In addition, there is no significant departure of the Bouguer anomaly curve from the general background values that can be inferred to be due to the BMA. This could be an indication that the BMA has little or no density contrast with the basement.

The gravity model for profile C–C' is shown in Fig. 17. The Bouguer gravity value increases steadily from the northwestern–southeastern direction of the map. The thickness of the Karoo (Beaufort, Ecca, and Dwyka) and Cape Supergroup varies along this profile with noticeable undulations which may be due to deformation. The dolerite intrusions that outcrop on the surface are seen in the model extending to depth cutting through the Cape, Dwyka, Ecca, and Beaufort sequence in the form of dykes and sills (Fig. 17). Lateral displacement (faulting) is also seen cutting through the Cape, Dwyka, and Ecca Groups between distances of about 20 and 100 km. Along profile C–C', the surface expression of the BMA is located between the distances of about 92–150 km. But again, the BMA does not have any appreciable gravity anomaly.

Fig. 17 Gravity modelling of profile C–C'. VE = 1.5; initial and final RMS errors are 65.4 and 1.2, respectively. Note: the shape of the BMA is unknown

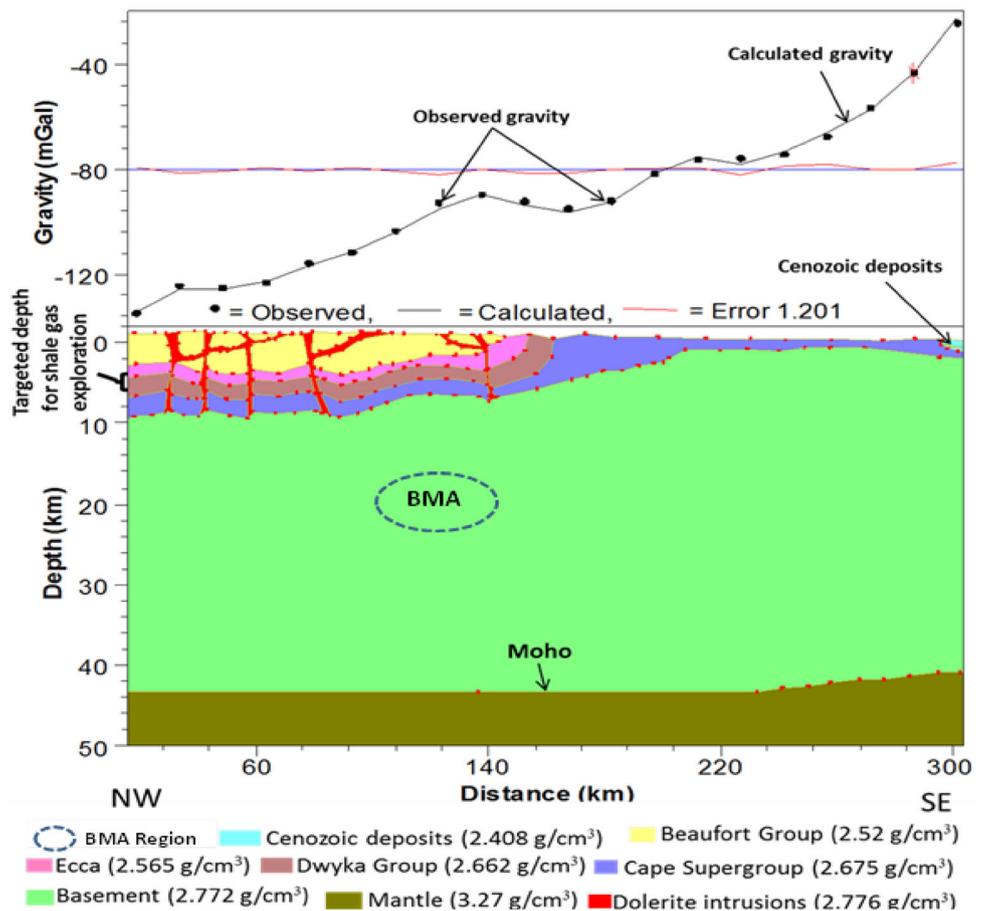


Figure 18 shows model of profile D–D'. The minimum gravity value along this profile is about -136 mGal, whilst the maximum gravity value is -22.1 mGal. Generally, the thickness of the Karoo (Beaufort, Eccca, and Dwyka Groups) and Cape (Witteberg, Bokkeveld, and Table Mountain Groups) Supergroups varies along this profile probably due to deformation as depicted in the model by undulating surfaces. The geologic sequences have been intruded by dolerites that extend into the basement.

Discussion

In the study area, the total magnetic field residual anomalies increase from the coastal area to inland with three noticeable magnetic high zones (Fig. 6). The third magnetic zone which is part of the BMA split into narrow magnetic high zones as revealed by the vertical derivative map. The magnetic maps (Figs. 3, 4, 5, 6, 7, 8, 9, 10, 11) clearly show ring-like structures and lineaments which coincide with the mapped dolerite intrusions and some new ones are also revealed. The average magnetic depths of about 0.6 and 15 km were determined from the radially averaged power spectrum as the depths to the top of the

shallow and deep sources, respectively. The BMA appears in all the depth slices and becomes broader and stronger up to a depth of 15.2 km. This could be an indication that the source(s) of the BMA is deeper possibly within the basement. Lindeque et al. (2007) documented that the BMA could be due to a buried body in the basement (that is possibly intersected by a fault plane dipping in a southern direction) or/and a buried massive sulphide-magnetite body within the basement. The average magnetic depths of about 15 km that was determined as the depths to the top of the deep sources could be linked to the BMA since the BMA appeared in all the presented depth slices. The estimated depth is also similar to the depth envisaged by Pitts et al. (1992) that the buried magnetic body of about 30 km width dips southward from about 7 km below the ground surface to a depth of up to 30 km. The BMA is one of the largest crustal anomalies in the world and trends northeast to southwest of the study area. It is suggested to be due to deeper source(s) possibly within the basement which agrees with the findings of Pitts et al. (1992) and Weckmann et al. (2007a). The bean shape anomaly also behaves in a similar manner such as the BMA, and thus it could possibly have a source(s) at similar depths to the BMA based on the depth slicing results.

Fig. 18 Gravity modelling of profile D–D'. VE = 1.9; initial and final RMS errors are 48.3 and 1.5, respectively

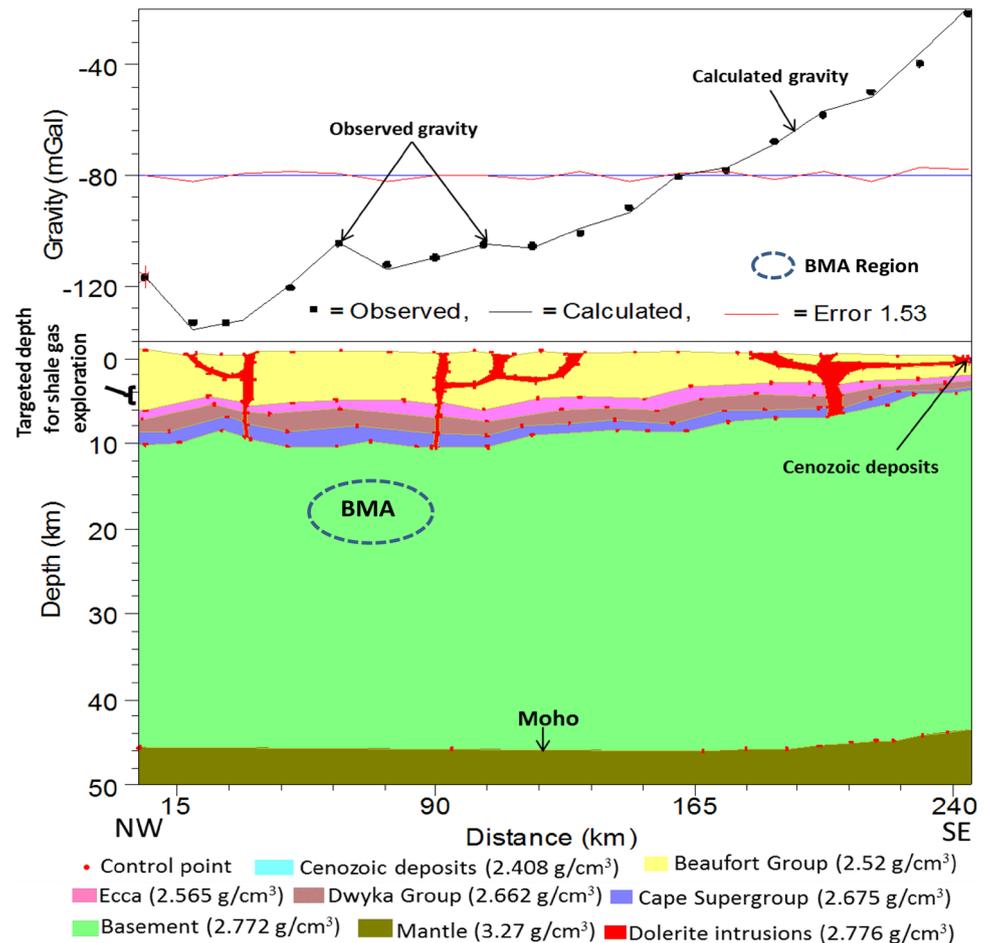


Table 3 Rock density values used for gravity profile modelling

| | Beaufort | Ecca | Dwyka | Cape | Dolerite intrusions | Basement or pre-cape | Below the Moho (mantle) |
|--------------------------------------|--------------------|----------|------------|------------------------|---------------------|----------------------|-------------------------|
| Lithology | Sandstone/mudstone | Mudstone | Diamictite | Sandstone/conglomerate | Dolerite | Chert/schist | – |
| No of samples | 85 | 66 | 20 | 32 | 36 | 19 | – |
| Highest density (g/cm ³) | 2.77 | 2.78 | 2.68 | 2.72 | 2.83 | 2.82 | 3.40 |
| Lowest density (g/cm ³) | 2.27 | 2.35 | 2.59 | 2.63 | 2.71 | 2.72 | 3.14 |
| Average density (g/cm ³) | 2.52 | 2.57 | 2.66 | 2.68 | 2.77 | 2.77 | 3.27 |

The Bouguer gravity anomaly map (Fig. 14) shows an increase in gravity values from about -100 mGal inland to around 62 mGal in coastal areas. The $2\frac{1}{2}$ D gravity profile modelling results show that the long wavelength variation is due to a deeper source/interface inland that shallows towards the coast, e.g., basement and/or Moho. Generally, the results of the modelling indicate a horizontal Moho at a depth of about 45 km inland and it shallows to about 42 km at the coastal area with a slope of about 3° . This shallowing of the Moho gives rise to the long wavelength gravity anomaly from inland to coastal areas. The Moho

depth obtained in this study are in agreement with the depth results from Tedla et al. (2011) as well as Stankiewicz and de Wit (2013), which are in the range of 42 – 45 km for inland sections. The $2\frac{1}{2}$ D gravity profile modelling results reveal a thickness of up to 8 km for the Karoo sedimentary sequence (i.e., depth from Beaufort to the Dwyka Group). The targeted area for shale gas exploration is the Ecca Group. All the four profiles cross the Ecca Group. Rowsell and De Swardt (1976) as well as Lindeque et al. (2011) have suggested a targeted exploration depth from about 3.5 – 5 km for shale gas and this depth range is within the

0–7.2 km determined from the gravity modelling. At the deepest section, the top of the Ecça Group is at 6.5 km depth and at the bottom is at 7.2 km. The determined laboratory dry densities range within about 2.59–2.68, 2.35–2.78, and 2.52–2.77 g/cm³ for the Dwyka, Ecça, and Beaufort Group, respectively. The calculated porosities range within 0.29–0.87, 1.29–5.52, and 0.48–3.34% for the Dwyka, Ecça, and Beaufort Group, respectively. The high porosity and low density values of the Ecça samples could be possibly due to weathering. Based on the gravity profile models, the variation in thickness of the geologic units could be inferred to be due to deformation. In addition, it could be inferred that dolerite intrusions which are prominent in the study area are interconnected at depth as depicted in the models. This could possibly affect the quality of the shale resources and probably pose threat by increasing the risk of fracking the Karoo for shale gas exploration.

Conclusions

Based on the results presented in this paper, the following can be inferred:

- The anomaly which is part of the BMA tends to divide into two anomalies in the study area and continues to trend in an NE–SW direction. The BMA is possibly due to a source(s) buried deep, as shown by the depth slices.
- From the radially averaged power spectrum, the average depths to the top of the shallow and deep magnetic sources were estimated to be approximately 0.6 and 15 km, respectively.
- Based on the 2½ D gravity profile modelling and depth slice results, it could be deduced that geologic structures such as dolerite intrusions are interconnected and extend from the surface to a depth of about 12 km.
- The bean-shape anomaly, based on depth slice results, behaves in a similar manner, such as the BMA, and thus it could possibly have a source(s) at similar depth to the BMA.
- The dominant gravity signature in the coastal area is of long wavelength, which is possibly due to a deep source/interface inland that shallows towards the coast, e.g., basement and/or Moho. The Moho is at about 45 km depth inland and shallows to about 42 km at the coast.
- The 2½ D gravity profile modelling results reveal the configuration of the basin, with the Beaufort Group being the thickest (6342 ± 295 m) of all the groups that make up the Karoo Supergroup in the study area, followed by the Ecça (3207 ± 263 m) and Dwyka Groups (727 ± 25 m), respectively.

Acknowledgements Thanks to Govan Mbeki Research and Development Centre (GMRDC) for supervisory link bursary and NRF bursary to Mr C. Baiyegunhi (UID: 101980). The authors are grateful to Fugro Airborne Surveys, Council for Geoscience for supplying the geophysical data. Geosoft is also appreciated for providing the “Extended GM-SYS” that was used for the modelling. We thank the editors and anonymous reviewers for their constructive comments, which helped us to improve the manuscript.

References

- Beattie JC (1909) Report of a magnetic survey of South Africa. R Soc Lond Publ 76:1–35
- Catuneanu O, Hancox PJ, Rubidge BS (1998) Reciprocal flexural behaviour and contrasting stratigraphies: a new basin development model for the Karoo retroarc foreland system, South Africa. *Basin Res* 10:417–439. doi:10.1046/j.1365-2117.1998.00078.x
- Catuneanu O, Hancox PJ, Cairncross B, Rubidge BS (2002) Foredeep submarine fans and forebulge deltas: orogenic off-loading in the underfilled Karoo Basin. *J Afr Earth Sci* 35:489–502. doi:10.1016/S0899-5362(02)00154-9
- Catuneanu O, Wopfner H, Eriksson PG, Cairncross B, Rubidge BS, Smith RMH, Hancox PJ (2005) The Karoo basins of south-central Africa. *J Afr Earth Sci* 43:211–253. doi:10.1016/j.jafrearsci.2005.07.007
- Chevallier L, Woodford AC (1999) Morpho-tectonics and mechanism of emplacement of the dolerite rings and sills of the western Karoo, South Africa. *S Afr J Geol* 102:43–54
- Chevallier L, Goedhart M, Woodford AC (2001) The influences of dolerite sill and ring complexes on the occurrence of groundwater in Karoo fractured aquifers: a morpho-tectonic approach. *Water Res Comm Rep* 37(1):4–146
- Cook FA, White DJ, Jones AG, Eaton DW, Hall J, Clowes RM (2010) How the crust meets the mantle: lithoprobe perspectives on the Mohorovičić discontinuity and crust–mantle transition. *Can J Earth Sci* 47:315–351. doi:10.1139/E09-076
- Council for Geoscience (1995) Geological map of South Africa 1:250000, Sheet 2922, Prieska. Geological Survey of South Africa, Pretoria
- De Beer JH, Gough DI (1980) Conductive structures in southernmost Africa: a magnetometer array study. *Geophys J R Astron Soc* 63:479–495. doi:10.1111/j.1365-246X.1980.tb02633.x
- De Beer JH, Meyer R (1983) Geo-electrical and gravitational characteristics of the Namaqua–Natal mobile belt and its boundaries. *Spec Publ Geol Soc S Afr* 10:91–100
- De Beer JH, Meyer R (1984) Geophysical characteristics of the Namaqua–Natal mobile belt and its boundaries, South Africa. *J Geodyn* 1:473–494. doi:10.1016/0264-3707(84)90020-6
- De Beer JH, Van Zijl JSV, Gough DI (1982) The southern cape conductive belt (South Africa): its composition, origin and tectonic significance. *Tectonophysics* 83:205–225. doi:10.1016/0040-1951(82)90019-1
- De Wit MJ, Horsfield B (2006) Inkaba yeAfrica Project surveys sector of Earth from core to space. *EOS Trans Am Geophys Union* 87(11):113–117. doi:10.1029/2006EO110002
- De Wit MJ, Ransome IG (1992) Regional inversion tectonics along the southern margin of Gondwana. In: De Wit MJ, Ransome IG (eds) *Inversion tectonics of the cape fold belt, Karoo and Cretaceous Basins of Southern Africa*. Balkema, Rotterdam, pp 15–22
- Du Plessis A, Simpson ESW (1974) Magnetic anomalies associated with the southeastern continental margin of South Africa. *Mar Geophys Res* 2:99–110. doi:10.1007/BF00340028

- Flint SS, Hodgson DM, Sprague AR, Brunt RL, Van Der Merwe WC, Figueiredo J, Prélat A, Box D, Di Celma C, Kavanagh JP (2011) Depositional architecture and sequence stratigraphy of the Karoo basin floor to shelf edge succession, Laingsburg depocentre, South Africa. *Mar Petrol Geol* 28:658–674. doi:[10.1016/j.marpetgeo.2010.06.008](https://doi.org/10.1016/j.marpetgeo.2010.06.008)
- Fuller BD (1967) Two-dimensional frequency analysis and design of grid operators. In: Hansen DA, Heinrichs WE, Holmer RC, MacDougall RE, Rogers GR, Sumner JS, Ward SH (eds) Society of exploration geophysicists' mining geophysics, vol 2. pp 658–708
- Geel C, Schulz HM, Booth P, De Wit M, Horsfield B (2013) Shale gas characteristics of Permian black shales in South Africa: results from recent drilling in the Ecca Group (Eastern Cape). *Energy Procedia* 40:256–265. doi:[10.1016/j.egypro.2013.08.030](https://doi.org/10.1016/j.egypro.2013.08.030)
- Gough D, De Beer JH, Van Zijl J (1973) A magnetometer array study in southern Africa. *Geophys J R Astron Soc* 34:421–433. doi:[10.1111/j.1365-246X.1973.tb02405.x](https://doi.org/10.1111/j.1365-246X.1973.tb02405.x)
- Gunn PJ (1972) Application of Wiener filters to transformations of gravity and magnetic data. *Geophys Prospect* 20:860–871. doi:[10.1111/j.1365-2478.1972.tb00671.x](https://doi.org/10.1111/j.1365-2478.1972.tb00671.x)
- Hälbich IW (1983) A tectogenesis of the Cape Fold Belt (CFB). In: Söhne APG, Hälbich IW (eds) *Geodynamics of the Cape Fold Belt*. Special publications of the geological society of South Africa, vol 12. pp 165–175
- Hälbich IW (1993) The Cape Fold Belt—Aguilhas Bank transect. In: Hälbich IW (ed) *Cape Fold Belt—Aguilhas Bank transect across Gondwana suture, Southern Africa*. American Geophysical Union, Washington, pp 3–18. doi:[10.1002/9781118668016.ch1](https://doi.org/10.1002/9781118668016.ch1)
- Harvey JD, De Wit MJ, Stankiewicz J, Doucouré CM (2001) Structural variations of the crust in the Southwest Cape, deduced from seismic receiver functions. *S Afr J Geol* 104:231–242. doi:[10.2113/1040231](https://doi.org/10.2113/1040231)
- Hynes A, Snyder DB (1995) Deep-crustal mineral assemblages and potential for crustal rocks below the Moho in the Scottish Caledonides. *Geophys J Int* 123(2):323–339. doi:[10.1111/j.1365-246X.1995.tb06857.x](https://doi.org/10.1111/j.1365-246X.1995.tb06857.x)
- Johnson MR, van Vuuren CJ, Hegenberger WF, Key R, Shoko U (1996) Stratigraphy of the Karoo Supergroup in southern Africa: an overview. *J Afr Earth Sci* 23:3–15. doi:[10.1016/S0899-5362\(96\)00048-6](https://doi.org/10.1016/S0899-5362(96)00048-6)
- Johnson MR, van Vuuren CJ, Visser JNJ, Cole DI, Wickens HV, Christie ADM, Roberts DL, Brandl G (2006) Sedimentary rocks of the Karoo Supergroup. In: Johnson MR, Anhaeusser CR, Thomas RL (eds) *The geology of South Africa*. Geological Society of South Africa, Johannesburg/Council for Geoscience, Pretoria, pp 461–500
- Li X (2006) Understanding 3D analytic signal amplitude. *Geophysics* 71(2):13–16. doi:[10.1190/1.2184367](https://doi.org/10.1190/1.2184367)
- Lindeque A, Ryberg T, Weber M, Stankiewicz J, De Wit MJ (2007) Deep crustal seismic reflection experiment across the Southern Karoo basin, South Africa. *S Afr J Geol* 110:419–438. doi:[10.2113/gssaj.110.2-3.419](https://doi.org/10.2113/gssaj.110.2-3.419)
- Lindeque A, De Wit MJ, Ryberg T, Weber M, Chevallier L (2011) Deep crustal profile across the Southern Karoo basin and Beattie magnetic anomaly, South Africa: an integrated interpretation with tectonic implications. *S Afr J Geol* 114:265–292. doi:[10.2113/gssaj.114.3-4.265](https://doi.org/10.2113/gssaj.114.3-4.265)
- Maus S, Dimri V (1996) Depth estimation from the scaling power spectrum of potential fields. *Geophys J Int* 124(1):113–120. doi:[10.1111/j.1365-246X.1996.tb06356.x](https://doi.org/10.1111/j.1365-246X.1996.tb06356.x)
- Mjelde R, Goncharov A, Müller RD (2013) The Moho: boundary above upper mantle peridotites or lower crustal eclogites, a global review and new interpretations for passive margins. *Tectonophysics* 609:636–650. doi:[10.1016/j.tecto.2012.03.001](https://doi.org/10.1016/j.tecto.2012.03.001)
- Pángaro F, Ramos VA (2012) Palaeozoic crustal blocks of onshore and offshore central Argentina: new pieces of the southwestern Gondwana collage and their role in the accretion of Patagonia and the evolution of Mesozoic south Atlantic sedimentary basins. *Mar Petrol Geol* 37:145–162. doi:[10.1016/j.marpetgeo.2012.05.010](https://doi.org/10.1016/j.marpetgeo.2012.05.010)
- Pitts B, Mahler M, de Beer J, Gough D (1992) Interpretation of magnetic, gravity and magnetotelluric data across the Cape Fold Belt and Karoo Basin. In: De Wit M, Ransome I (eds) *Inversion tectonics of the Cape Fold Belt, Karoo and Cretaceous basins of southern Africa*. Balkema, Rotterdam, Netherlands, pp 27–32
- Rabeh T, Abdallatif T, Mekkawi M, Khalil A, El-Emam A (2008) Magnetic data interpretation and depth estimation constraints: a correlative study on magnetometer and gradiometer data. *NRIAG J Geophys Spec Issue* 6:185–209
- Rowell DM, De Swardt AMJ (1976) Diagenesis in Cape and Karoo sediments, South Africa, and its bearing on their hydrocarbon potential. *Trans Geol Soc S Afr* 79:81–145
- Rust IC (1973) The evolution of the Paleozoic Cape Basin, southern margin of Africa. In: Nairn AEM, Stehli FG (eds) *The ocean basins and margins, vol 1, the South Atlantic*. Springer, New York, pp 247–276. doi:[10.1007/978-1-4684-3030-1_6](https://doi.org/10.1007/978-1-4684-3030-1_6)
- Scheiber-Enslin SE, Webb SJ, Ebbing J (2014) Geophysically plumbing the Main Karoo Basin, South Africa. *S Afr J Geol* 117(2):275–300. doi:[10.2113/gssaj.117.2.275](https://doi.org/10.2113/gssaj.117.2.275)
- Schreiber-Enslin S, Ebbing J, Eberle D, Webb S (2013) 3D Geophysical modelling of the Beattie magnetic anomaly and Karoo Basin, South Africa. *Geophys Res Abstr* 15:EGU2013-4407
- Smith RMH, Erickson PG, Botha WJ (1993) A review of the stratigraphy and sedimentary environments of the Karoo-aged basins of Southern Africa. *J Afr Earth Sci* 132:143–169. doi:[10.1016/0899-5362\(93\)90164-L](https://doi.org/10.1016/0899-5362(93)90164-L)
- Spector A, Grant FS (1970) Statistical models for interpreting aeromagnetic data. *Geophysics* 35(2):293–302. doi:[10.1190/1.1440092](https://doi.org/10.1190/1.1440092)
- Stankiewicz J, De Wit MJ (2013) 3.5 billion years of reshaped Moho, southern Africa. *Tectonophysics* 609:675–689. doi:[10.1016/j.tecto.2013.08.033](https://doi.org/10.1016/j.tecto.2013.08.033)
- Stankiewicz J, Ryberg T, Schulze A, Lindeque A, Weber MH, De Wit MJ (2007) Initial results from wide-angle seismic refraction lines in the Southern Cape. *S Afr J Geol* 110:407–418
- Stollhofen H, Stanistreet IG, Bangert B, Grill H (2000) Tuffs, tectonism and glacially related sea-level changes, Carboniferous–Permian, southern Namibia. *Palaeogeogr Palaeoclimatol Palaeoecol* 161(1–2):127–150. doi:[10.1016/S0031-0182\(00\)00120-6](https://doi.org/10.1016/S0031-0182(00)00120-6)
- Svensen H, Planke S, Chevallier L, Malthe-Sorensen A, Corfu F, Jamtveit B (2007) Hydrothermal venting of greenhouse gases triggering Early Jurassic global warming. *Earth Planet Sci Lett* 256(3/4):554–566
- Tankard AJ, Jackson MPA, Eriksson KA, Hobday DK, Hunter DR, Minter WEL (1982) *Crustal evolution of Southern Africa. 3.8 billion years of earth history*. Springer, New York, p 523
- Tankard A, Welsink H, Aukes P, Newton R, Stettler E (2009) Tectonic evolution of the Cape and Karoo basins of South Africa. *Mar Petrol Geol* 26:1379–1412. doi:[10.1016/j.marpetgeo.2009.01.022](https://doi.org/10.1016/j.marpetgeo.2009.01.022)
- Tankard A, Welsink H, Aukes P, Newton R, Stettler E (2012) Geodynamic interpretation of the Cape and the Karoo basins, South Africa. In: *Regional geology and tectonics: Phanerozoic passive margins, cratonic basins and global tectonics maps*. Elsevier, pp 868–945. doi:[10.1016/B978-0-444-56357-6.00022-6](https://doi.org/10.1016/B978-0-444-56357-6.00022-6)
- Tedla GE, Van Der Meijde M, Nyblade AA, Van Der Meer FD (2011) A crustal thickness map of Africa derived from a global

- gravity field model using Euler deconvolution. *Geophys J Int* 187:1–9. doi:[10.1111/j.1365-246X.2011.05140.x](https://doi.org/10.1111/j.1365-246X.2011.05140.x)
- Thybo H, Artemieva IM (2014) Moho and magmatic underplating in continental lithosphere. *Tectonophysics* 609:605–619. doi:[10.1016/j.tecto.2013.05.032](https://doi.org/10.1016/j.tecto.2013.05.032)
- Turner BR (1999) Tectono-stratigraphical development of the upper Karoo foreland basin: orogenic unloading versus thermally-induced Gondwana rifting. *J Afr Earth Sci* 28:215–238. doi:[10.1016/S0899-5362\(99\)00025-1](https://doi.org/10.1016/S0899-5362(99)00025-1)
- Van Der Voort I (2001), Risk based decision tool for managing and protecting groundwater resources. Ph.D. Thesis, University of the Free State, Bloemfontein
- Veevers JJ, Cole DI, Cowan EJ (1994) Southern Africa: Karoo Basin and Cape Fold Belt. *GSA Mem* 184:223–280. doi:[10.1130/MEM184-p223](https://doi.org/10.1130/MEM184-p223)
- Visser JNJ (1987) The palaeogeography of part of south-western Gondwana during the Permo-Carboniferous glaciation. *Palaeogeogr Palaeoclimatol Palaeocol* 61:205–219. doi:[10.1016/0031-0182\(87\)90050-2](https://doi.org/10.1016/0031-0182(87)90050-2)
- Visser JNJ (1995) Post-glacial Permian stratigraphy and geography of southern and central Africa: boundary conditions for climatic modelling. *Palaeogeogr Palaeoclimatol Palaeocol* 118:213–243. doi:[10.1016/0031-0182\(95\)00008-3](https://doi.org/10.1016/0031-0182(95)00008-3)
- Visser JNJ, Praekelt HE (1996) Subduction, mega-shear systems and Late Palaeozoic basin development in the African segment of Gondwana. *Geol Rundsch* 85:632–646. doi:[10.1007/BF02440101](https://doi.org/10.1007/BF02440101)
- Weckmann U, Ritter O, Jung A, Branch T, De Wit M (2007a) Magnetotelluric measurements across the Beattie magnetic anomaly and the Southern Cape Conductive Belt, South Africa. *J Geophys Res* 112:416. doi:[10.1029/2005JB003975](https://doi.org/10.1029/2005JB003975)
- Weckmann U, Jung A, Branch T, Ritter O (2007b) Comparison of electrical conductivity structures and 2D magnetic modelling along two profiles crossing the Beattie Magnetic Anomaly, South Africa. *S Afr J Geol* 110:449–464. doi:[10.2113/gssajg.110.2-3.449](https://doi.org/10.2113/gssajg.110.2-3.449)
- Woodford AC, Chevallier LP (2002) Regional characterization and mapping of Karoo fractured aquifer systems—an integrated approach using a geographical information system and digital processing. *Water Res Comm Rep* 653/1/02:40–55

Shear waves in elastic medium with void pores welded between vertically inhomogeneous and anisotropic magnetoelastic semi-infinite media

Shishir Gupta¹ · Mostaid Ahmed¹ · Abhijit Pramanik¹

Received: 20 January 2017 / Accepted: 25 January 2017 / Published online: 10 February 2017
© Institute of Geophysics, Polish Academy of Sciences & Polish Academy of Sciences 2017

Abstract The paper intends to study the propagation of horizontally polarized shear waves in an elastic medium with void pores constrained between a vertically inhomogeneous and an anisotropic magnetoelastic semi-infinite media. Elasto-dynamical equations of elastic medium with void pores and magnetoelastic solid have been employed to investigate the shear wave propagation in the proposed three-layered earth model. Method of separation of variables has been incorporated to deduce the dispersion relation. All possible special cases have been envisaged and they fairly comply with the corresponding results for classical cases. The role of inhomogeneity parameter, thickness of layer, angle with which the wave crosses the magnetic field and anisotropic magnetoelastic coupling parameter for three different materials has been elucidated and represented by graphs using MATHEMATICA.

Keywords Vertical inhomogeneity · Void volume fraction · Maxwell's equation · Magnetoelastic coupling parameter · Dispersion equation

Introduction

The crustal layer of earth is largely composed of several igneous, sedimentary and metamorphic rocks and is relatively more inhomogeneous than the mantle. The interface

waves in layered elastic media with different boundaries help us to gather information about the basic characteristics of various rocks and have great importance due to its relevance for other phenomena in seismology, geophysics and earth system sciences. It also provides information regarding the configuration of rock layers for oil exploration, earthquake disaster management, signal processing and civil engineering. A wide range of basic literatures can be found in Love (1927), Biot (1965), Achenbach (1973), Bullen and Bolt (1985) and Gubbins (1990).

Owing to the variations in elastic properties due to presence of typical rock formations and overburdened stress medium, elastic solids exhibit inhomogeneity to a large extent which forms the reasonable ground for assuming anisotropy in continents. It is well known that earth's crust (0–20 km) is neither isotropic nor homogeneous in nature. Inhomogeneity in the upper semi-infinite medium has been considered because inside the earth, material properties may vary linearly, quadratically or exponentially. The wave velocities are considerably influenced by inhomogeneities in the earth's crust. Exponential inhomogeneities in the material properties of different media have taken into account by Qian et al. (2008), Pal and Acharya (1998), Kumar and Pal (2014) and many others whereas Manolis and Shaw (1996), Dey et al. (2000), Manolis et al. (2002) considered quadratic inhomogeneities. Singh et al. (2016b) devised a model on Love-type waves in prestressed inhomogeneous fibre-reinforced layer with corrugated boundary surfaces lying over a viscoelastic semi-infinite medium under hydrostatic state of stress. Pal and Mandal (2014) derived the existence of SH-type waves due to shearing stress discontinuity in a sandy layer overlying an isotropic and inhomogeneous elastic half-space using Cagniard–De Hoop technique.

✉ Mostaid Ahmed
mostaidahmed@yahoo.in

¹ Department of Applied Mathematics, Indian Institute of Technology (Indian School of Mines), Dhanbad 826004, India

The theory of linear elastic materials with void pores remained one of the most important generalizations of classical theory of elasticity. The basis of this theory lies on the fact that the bulk density of a material can be written as product of two fields, the density field in the matrix material and the volume fraction field. The presence of void pores effectively impairs the several mechanical properties and moduli of the elastic media. The theory is based on the elastic materials composed by a distribution of small pores (voids) in which the void volume is regarded among the kinematic variables and in limiting case of volume tending to zero.

Inadequacy of broadly explaining various geological and biological materials leads Nunziato and Cowin (1979) to develop the nonlinear theory of elastic materials with voids. A mathematical model to study the mechanical behaviour of porous solids was also discussed by Cowin and Nunziato (1983). Numerous applications can be found while studying the response of materials to homogeneous deformations, pure bending of beams and acoustic waves of small amplitudes. Birsan (2000) accomplished the existence and weak solution in linear theory of elastic materials with voids. A similar extension was made by Cowin (1985) to show how linear elastic materials with voids behave like viscoelastic materials. Recently, Chirita and Ghiba (2010) demonstrated the inhomogeneous plane wave solutions within the ambit of linear theory for strongly elliptic homogeneous poroelastic materials with a center of symmetry. Ke et al. (2006) modelled the propagation of Love waves in inhomogeneous fluid saturated porous medium with linearly varying properties. Wang and Zhang (1998) impeccably demonstrated the effect of Love waves in fluid-saturated porous media. Magdalena et al. (2014) derived the wave reflection and transmission coefficients in emerged porous media using shallow water equations with a friction term of the linearized Dupuit–Forchheimer’s formula. Singh et al. (2016a) established the dispersion relation of horizontally polarized shear waves in corrugated poroelastic layer lying over a fibre-reinforced half-space under the effect of corrugated boundary surfaces. Chattaraj and Samal (2013) exemplified with a model on how the Love waves transmit in the fibre-reinforced layer over a gravitating porous half-space. Samal and Chattaraj (2011) had shown the dispersion equation of Rayleigh type surface wave propagation in fibre-reinforced anisotropic elastic layer welded between liquid saturated porous half-space and uniform liquid layer.

The main remarkable characteristic retained by magnetoelastic materials is the presence of magnetic domain in addition to more conventional microstructural features like twin boundaries, planar defects dislocation and so on.

Several interactions compete at a wide range of length scales rendering it a large variety of microstructural configurations and macroscopic properties.

The interplay of Maxwell electromagnetic field with the motion of deformable solids is largely being undertaken by several investigators owing to the possibility of its application to geophysical problems and in optics and acoustics. The earth is subject to its own magnetic field and the material of the earth is electrically conducting. As a consequence, the magnetic nature of the earth’s material affects the seismic waves.

Knopoff (1955) was the first to study the effect of earth’s magnetic field on seismic wave propagation in the conducting core of the earth. Dunkin and Eringen (1963) demonstrated the problem of plane waves transmitting through an infinite medium and an infinite plate in the presence of large magnetoelastic and electrostatic field. Chattopadhyay and Choudhury (1990) investigated few problems regarding magnetoelastic and thermoelastic medium. An excellent review of wave motion in magnetizable deformable media can also be seen in the research paper of Maugin (1998). Kumar et al. (2015) elucidated reflection and transmission phenomena of SH-waves through magnetoelastic medium constrained between two viscoelastic media. Chattopadhyay et al. (2012) discussed the effect of point source and heterogeneity on the transmission of SH-waves in a viscoelastic layer over a viscoelastic half-space. Acharya et al. (2009) well formulated the effect of magnetic field and initial stress on interface waves in transversely isotropic media.

It is noteworthy to mention that igneous or soft materials which exhibit magnetoelastic characteristics are the significant constituents of the earth’s upper mantle. In an endeavour to expose the converted characteristics of the high density mantle and to gather a vivid information, the study on seismic waves passing through this region is of great significance. Asthenosphere that forms the transition zone and acts as an intermediary between low-density crust and high-density mantle is utterly magnetoelastic in nature and can be held responsible for the vast elasto-dynamical processes causing earthquakes. Interplay of magnetic field with the motion of an elastic solid has proved its relevance due to its theoretical and practical significance in various branches of science as well.

Thus, horizontally polarized waves are always influenced by the magnetoelastic nature of the media through which they propagate. With this geophysical notion and motivated by certain practical situations, the authors have considered the proposed three layered model which has some realistic implications in non-destructive

measurements of mechanical and electromagnetic parameter of the materials found abundantly in asthenosphere.

The present investigation expounds the effect of various parameters (e.g. void parameter, magnetoelastic coupling parameter, inhomogeneity, thickness of layer, etc.) on the propagation of horizontally polarized shear waves in elastic medium of void pores clamped between vertically inhomogeneous and anisotropic magnetoelastic semi-infinite media. The dispersion relation as well as several numerical examples illuminating the effects of those parameters for different magnetoelastic materials have been established. The ocular interference of magnetoelastic coupling parameter and the angle with which wave crosses the magnetic field are depicted by means of surface plots.

Formulation of the problem

We consider an elastic medium with void pores of thickness H bonded by a welded contact of vertically inhomogeneous and anisotropic magnetoelastic half-space (Fig. 1). Referring to the rectangular Cartesian coordinate system, the lower semi-infinite media occupies the region $z \geq 0$ and the axis is chosen in the direction of increasing depth. $z = 0$ is taken as the interface of elastic layer with void pores and anisotropic magnetoelastic semi-infinite medium. Therefore, the porous layer lies in the region $-H \leq z < 0$ whereas the upper semi-infinite medium with exponential inhomogeneities occupies $z \leq -H$. The variations in rigidity and density are taken as $\mu = \mu_1 e^{az}$, $\rho = \rho_1 e^{az}$, where a has the dimension that is inverse of length.

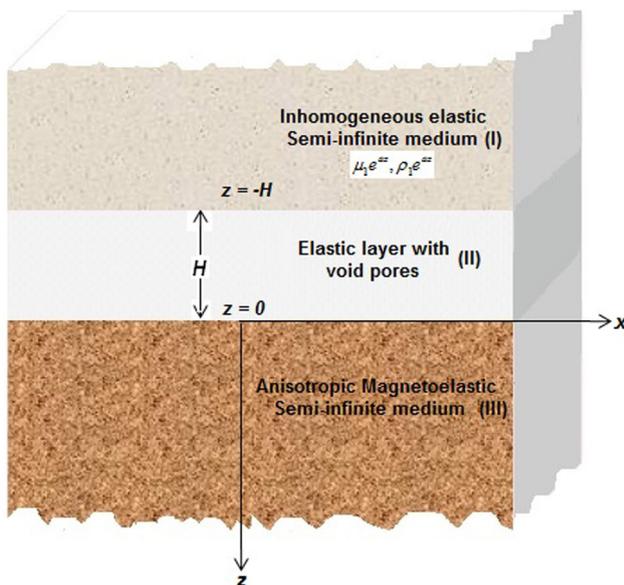


Fig. 1 Geometry of the three-layered earth model

Solution of the uppermost inhomogeneous half-space

Since the propagation of SH-waves is considered in x -direction and causing displacements only in the y -direction, the displacement components are as expressed as:

$$u_1 = 0, \quad v_1 = v_1(x, z, t), \quad w_1 = 0$$

We consider μ , ρ as the rigidity and density of the inhomogeneous medium, respectively. Neglecting the body forces in the equations of motion for SH-wave is given by

$$\frac{\partial}{\partial x} \left(\mu \frac{\partial v}{\partial x} \right) + \frac{\partial}{\partial z} \left(\mu \frac{\partial v}{\partial z} \right) = \rho \frac{\partial^2 v}{\partial t^2} \quad (1)$$

Now, we have $\mu = \mu_1 e^{az}$, $\rho = \rho_1 e^{az}$, where a has the dimension that is inverse of length.

Putting the values of μ and ρ in Eq. (1), we get

$$\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial z^2} + a \frac{\partial v}{\partial z} = \frac{1}{\beta_1^2} \frac{\partial^2 v}{\partial t^2} \quad (2)$$

where, $\beta_1 = \sqrt{\frac{\mu_1}{\rho_1}}$.

Now, using the separation of variables by putting the transformation $v = V_1(z) e^{ik(x-ct)}$ in (2), the above equation reduces into

$$\frac{d^2 V_1}{dz^2} + m_1^2 V_1^2 = 0 \quad (3)$$

where

$$m_1^2 = k^2 \left(\frac{c^2}{\beta_1^2} - 1 \right) - \frac{a^2}{4}.$$

Hence,

$$V_1(z) = A_1 e^{imz} + B_1 e^{-imz}. \quad (4)$$

Satisfying the condition that $\lim_{z \rightarrow -\infty} V_1(z)$, the above Eq. (4) reduces to

$$v_1 = \frac{1}{\sqrt{\mu_1 e^{az}}} \cdot A_1 e^{imz} \quad (5)$$

where A_1 is the arbitrary independent constant.

Solution of the elastic layer with void pores

SH-wave motion has only one displacement component since the particle moves along the perpendicular to the direction of wave propagation and within the plane. Therefore, the displacement components of SH-waves can be written as $u_2 = 0$, $v_2 = v_2(x, z, t)$, $w_2 = 0$. The equations of motion which are not identically satisfied under no body forces are given by Cowin and Nunziato (1983) as:

$$\rho_2 \frac{\partial^2 v_2}{\partial t^2} = \mu_2 \left(\frac{\partial^2 v_2}{\partial x^2} + \frac{\partial^2 v_2}{\partial z^2} \right) + \beta_2 \left(\frac{\partial \phi_2}{\partial x} + \frac{\partial \phi_2}{\partial z} \right) \tag{6}$$

$$\rho_2 \bar{k} \frac{\partial^2 \phi_2}{\partial t^2} = \alpha \left(\frac{\partial^2 \phi_2}{\partial x^2} + \frac{\partial^2 \phi_2}{\partial z^2} \right) - \omega \frac{\partial \phi_2}{\partial t} - \xi \phi_2 \tag{7}$$

where α , β_2 , ω and ξ are void parameters whereas μ_2 , ρ_2 are the rigidity and density of the bulk material respectively. \bar{k} is the equilibrated inertia, and ϕ_2 is the change of void volume fraction.

For the waves propagating in the positive direction of the x -axis with velocity c , the solution of the above equations may be written as:

$$v_2 = \bar{v}_2(z) e^{ik(x-ct)} \tag{8}$$

$$\phi_2 = \bar{\phi}_2(z) e^{ik(x-ct)} \tag{9}$$

where $\bar{v}_2(z)$ and $\bar{\phi}_2(z)$ satisfy the equations.

$$\bar{v}_2''(z) - N^2 \bar{v}_2(z) + B [ik \bar{\phi}_2(z) + \bar{\phi}_2'(z)] = 0 \tag{10}$$

With the values of N , B as

$$N = k \sqrt{\left(1 - \frac{c^2}{A^2}\right)}, \quad B = \frac{\beta_2}{\mu_2}, \quad A = \sqrt{\frac{\mu_2}{\rho_2}}$$

and

$$\bar{\phi}_2''(z) - M^2 \bar{\phi}_2(z) = 0 \tag{11}$$

where

$$M = \sqrt{[(\alpha k^2 - \rho_2 \bar{k} k^2 c^2 - i\omega k c t + \xi)/\alpha]}$$

α, \bar{k}, ξ being constants for a particular material. Ignoring the damping term ω which is negligibly small for sinusoidal wave and, hence, the value may be taken as:

$$M = k \sqrt{1 - \frac{c^2}{\alpha/\rho_2 \bar{k}} + \frac{1}{k^2 \left(\frac{\alpha}{\xi}\right)}} \tag{12}$$

With the help of (12), the solution of (11) can be written as:

$$\bar{\phi}_2 = R_3 e^{Mz} + R_4 e^{-Mz} \tag{13}$$

Again with the help of (13), the solution of (10) can be written as:

$$\bar{v}_2 = R_1 e^{Nz} + R_2 e^{-Nz} - R_3 \frac{B(ik + M)}{M^2 - N^2} e^{Mz} - R_4 \frac{B(ik - M)}{M^2 - N^2} e^{-Mz} \tag{14}$$

Hence, the solution of (6) and (7) can be written as:

$$v_2 = \left[R_1 e^{Nz} + R_2 e^{-Nz} - R_3 \frac{B(ik + M)}{M^2 - N^2} e^{Mz} - R_4 \frac{B(ik - M)}{M^2 - N^2} e^{-Mz} \right] e^{ik(x-ct)} \tag{15}$$

$$\phi_2 = [R_3 e^{Mz} + R_4 e^{-Mz}] e^{ik(x-ct)} \tag{16}$$

R_i 's ($i = 1, 2, 3, 4$) are the arbitrary constants.

Solution of the anisotropic magnetoelastic layer

The governing equations of motion in a perfectly conducting elastic medium in the presence of Lorentz forces can be written as:

$$\tau_{ij,j} + (\vec{J} + \vec{B})_i = \rho_3 \frac{\partial^2 u_i}{\partial t^2} \tag{17}$$

where τ_{ij} are the Cartesian components of the stress tensor, ($i = 1, 2, 3$).

$(\vec{J} + \vec{B})_i$ is the component of force $(\vec{J} + \vec{B})$ in the i th direction, \vec{J} is the electric current density and \vec{B} is the magnetic induction vector. u_i are the components of the displacement vector \vec{u} . ρ_3 is the mass density. Comma in the first term denotes the partial derivatives w.r.to spatial coordinates. The Maxwell's fundamental equations that govern the electro-magnetic field are given by

$$\begin{aligned} \nabla \times \vec{H} &= \vec{J}, \quad \nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \\ \nabla \cdot \vec{B} &= 0 \quad \text{where } \vec{B} = \mu_2 \vec{H} \\ \vec{J} &= \sigma \left(\vec{E} + \frac{\partial \vec{u}}{\partial t} \times \vec{B} \right) \end{aligned} \tag{18}$$

Following Chaudhary et al. (2004), the Maxwell's stress tensor is defined as:

$$(\tau_{ij})_{II} = \mu_e (H_i h_j + H_j h_i - H_k h_k \delta_{ij}) \tag{19}$$

where \vec{E} denotes the induced electric field.

$\vec{H} = (H_x, H_y, H_z)$ is the total applied and induced magnetic field, μ_e is the constant magnetic permeability, σ is the electric conductivity, $h_i = (h_1, h_2, h_3)$ is the change in basic magnetic field. Referring Chattopadhyay and Choudhury (1990) from Eq. (18), we can get

$$\nabla^2 \vec{H} = \sigma \mu_e \left[\frac{\partial \vec{H}}{\partial t} - \nabla \times \left(\frac{\partial \vec{u}}{\partial t} \times \vec{H} \right) \right]. \tag{20}$$

Since the motion of confines in the xz -plane, we take $\vec{u} = (0, u_3, 0)$ and $\frac{\partial}{\partial y} \equiv 0$.

Thus, the Eq. (20) gets into

$$\frac{\partial H_x}{\partial t} = \frac{1}{\sigma\mu_e} \nabla^2 H_x \quad (21)$$

$$\frac{\partial H_y}{\partial t} = \frac{1}{\sigma\mu_e} \nabla^2 H_y + \frac{\partial}{\partial x} \left(H_x \frac{\partial u_3}{\partial t} \right) + \frac{\partial}{\partial z} \left(H_z \frac{\partial u_3}{\partial t} \right) \quad (22)$$

$$\frac{\partial H_z}{\partial t} = \frac{1}{\sigma\mu_e} \nabla^2 H_z. \quad (23)$$

When $\sigma \rightarrow \infty$ it can be shown from (21) and (23) that there exists no perturbation in H_x and H_z . Hence, the Eq. (22) gets into the form

$$\frac{\partial H_y}{\partial t} = \frac{\partial}{\partial x} \left(H_x \frac{\partial u_3}{\partial t} \right) + \frac{\partial}{\partial z} \left(H_z \frac{\partial u_3}{\partial t} \right). \quad (24)$$

We take small perturbations h_3 in H_y so that

$$H_x = H_{01}, \quad H_y = H_{02} + h_3, \quad H_z = H_{03}.$$

We now consider

$$H_{01} = H_0 \cos \phi_3, \quad H_{02} = 0, \quad H_{03} = H_0 \sin \phi_3.$$

Hence

$$\vec{H} = (H_0 \cos \phi_3, h_3, H_0 \sin \phi_3) \quad (25)$$

where $H_0 = |\vec{H}_0|$ and ϕ_3 is the angle at which the wave crosses the magnetic field.

Using the relation (24) in (25), we get the equation

$$\frac{\partial h_3}{\partial t} = \frac{\partial}{\partial t} \left[H_0 \cos \phi_3 \frac{\partial u_3}{\partial x} + H_0 \sin \phi_3 \frac{\partial u_3}{\partial z} \right]. \quad (26)$$

Then, partially integrating the above equation w.r.to t we get

$$h_3 = H_0 \cos \phi_3 \frac{\partial u_3}{\partial x} + H_0 \sin \phi_3 \frac{\partial u_3}{\partial z}. \quad (27)$$

Initially, there is no perturbation in the magnetic field, i.e. $h_3 = 0$.

Hence from (18), we have

$$\vec{J} \times \vec{B} = \mu_e \left[(\vec{H} \cdot \nabla) \vec{H} - \frac{1}{2} \nabla \vec{H}^2 \right]. \quad (28)$$

Consequently with the aid of (25), the Eq. (28) gets into the form

$$\begin{aligned} (\vec{J} \times \vec{B})_I &= (\vec{J} \times \vec{B})_{III} = 0 \\ (\vec{J} \times \vec{B})_{II} &= \mu_e H_0 \left[\cos \phi_3 \frac{\partial h_3}{\partial x} + \sin \phi_3 \frac{\partial h_3}{\partial z} \right]. \end{aligned} \quad (29)$$

Now with the help of (29), the Eq. (26) produces

$$(\vec{J} \times \vec{B})_{II} = \mu_e H_0^2 \left[\cos^2 \phi_3 \frac{\partial^2 u_3}{\partial x^2} + \sin 2\phi_3 \frac{\partial^2 u_3}{\partial x \partial z} + \sin^2 \phi_3 \frac{\partial^2 u_3}{\partial z^2} \right]. \quad (30)$$

Thus, the only non-vanishing equation for the anisotropic magnetoelastic medium is

$$\frac{\partial \tau_{xy}}{\partial x} + \frac{\partial \tau_{yz}}{\partial z} + (\vec{J} \times \vec{B})_{II} = \rho_3 \frac{\partial^2 u_{33}}{\partial t^2} \quad (31)$$

where ρ_3 is the density of the medium, and τ_{xy} , τ_{yz} are the stress components. u_{33} is the displacement component in the y -direction.

The stress–strain relations of the anisotropic materials are taken according to Kumar and Pal (2014)

$$\begin{aligned} \tau_{xy} &= C_{46} \frac{\partial u_{33}}{\partial z} + C_{66} \frac{\partial u_{33}}{\partial x} \\ \tau_{yz} &= C_{44} \frac{\partial u_{33}}{\partial z} + C_{46} \frac{\partial u_{33}}{\partial x} \end{aligned} \quad (32)$$

where C_{44} , C_{46} and C_{66} are the elastic constants.

Using (30) and (32), the Eq. (31) converts into

$$R_{66} \frac{\partial^2 u_{33}}{\partial x^2} + 2R_{46} \frac{\partial^2 u_{33}}{\partial x \partial z} + R_{44} \frac{\partial^2 u_{33}}{\partial z^2} = \rho_3 \frac{\partial^2 u_{33}}{\partial t^2} \quad (33)$$

where

$$\begin{aligned} R_{44} &= C_{44}(1 + m_H \sin^2 \phi_3), \quad R_{66} = C_{44} \left(\frac{C_{66}}{C_{44}} + m_H \cos^2 \phi_3 \right) \\ R_{46} &= C_{44} \left(\frac{C_{46}}{C_{44}} + m_H \sin \phi_3 \cos \phi_3 \right) \end{aligned}$$

and $m_H = \frac{\mu_e H_0^2}{C_{44}}$ is the anisotropic magnetoelastic coupling parameter.

Assuming the solution $u_{33}(x, z, t) = U_{33}(z)e^{ik(x-ct)}$ and substituting in Eq. (33), we have

$$\frac{d^2 U_{33}}{dz^2} + 2ik\gamma_1 \frac{dU_{33}}{dz} + k^2\gamma_2 \left\{ \frac{c^2}{\beta_3^2} - 1 \right\} U_{33} = 0 \quad (34)$$

where

$$\gamma_1 = \frac{R_{46}}{R_{44}}, \quad \gamma_2 = \frac{R_{66}}{R_{44}} \quad \text{and} \quad \beta_3^2 = \frac{R_{66}}{\rho_3}$$

and β_3 denotes the shear wave velocity in the anisotropic magnetoelastic medium.

Thus, the solution of Eq. (33) is written as:

$$U_{33}(z) = A_3 e^{iks_2 z} + B_3 e^{-iks_3 z} \quad (35)$$

where

$$\begin{aligned} s_2 &= -\gamma_1 + \sqrt{\gamma_1^2 + \gamma_2 \left\{ \frac{c^2}{\beta_3^2} - 1 \right\}} \quad \text{and} \\ s_3 &= \gamma_1 + \sqrt{\gamma_1^2 + \gamma_2 \left\{ \frac{c^2}{\beta_3^2} - 1 \right\}}. \end{aligned}$$

Hence, the displacement and stress components for the layer are given by

$$u_{33}(x, z, t) = (A_3 e^{iks_2 z} + B_3 e^{-iks_3 z}) e^{ik(x-ct)} \tag{36}$$

$$(\tau_{yz})_{II} = R_{44} \frac{\partial u_{33}}{\partial z} + R_{46} \frac{\partial u_{33}}{\partial x}. \tag{37}$$

The displacement component would be bounded, as $z \rightarrow \infty$ and the first term of Eq. (36) makes the solution unbounded. Therefore, the approximate solution of Eq. (36) reduces to

$$u_{33}(x, z, t) = B_3 e^{-iks_3 z} e^{ik(x-ct)}. \tag{38}$$

Boundary conditions

We assume that the elastic medium with void pores is clamped with the semi-infinite media of inhomogeneous and anisotropic magnetoelastic material. Therefore, the boundary conditions are the continuity of stresses and displacement and change of void volume fraction at the interfaces. These boundary conditions are mathematically demonstrated as:

- (i) At $z = -H$, i.e. at the interface of upper semi-infinite medium and intermediary layer
 - (a) $(\tau_{yz})_I = (\tau_{yz})_{II}$
 - (b) $v_1 = v_2$
 - (c) $(\vec{n} \cdot \nabla \phi_2) = 0$.
- (ii) At $z = 0$, i.e. at the interface of intermediary layer and lower semi-infinite medium
 - (a) $(\tau_{yz})_{II} = (\tau_{yz})_{III}$
 - (b) $v_2 = v_3$
 - (c) $(\vec{n} \cdot \nabla \phi_2) = 0$.

From the boundary condition $ii(a)$ and using the Eqs. (5) and (15), we get

$$\begin{aligned} &\sqrt{\mu_1} \left(im_1 - \frac{a}{2} \right) e^{-(im_1 + \frac{a}{2})H} \cdot A_1 - \mu_2 N e^{-NH} \cdot R_1 + \mu_2 N e^{NH} \cdot \\ &R_2 + \frac{\mu_2 B M (ik + M) e^{-MH}}{M^2 - N^2} \cdot R_3 \\ &- \frac{\mu_2 B M (ik - M) e^{MH}}{M^2 - N^2} \cdot R_4 = 0. \end{aligned} \tag{39}$$

Similarly from $i(b)$ and with the aid of Eqs. (5), and (15), we get

$$\begin{aligned} &\frac{1}{\sqrt{\mu_1} e^{az}} e^{-im_1 H} \cdot A_1 - e^{-NH} \cdot R_1 - e^{NH} \cdot R_2 \\ &+ \frac{B(ik + M) e^{-MH}}{M^2 - N^2} \cdot R_3 + \frac{B(ik - M) e^{MH}}{M^2 - N^2} \cdot R_4 \\ &= 0 \end{aligned} \tag{40}$$

Similarly from $i(c)$ and with the aid of Eq. (16), we get

$$e^{-MH} R_3 - e^{MH} R_4 = 0 \tag{41}$$

Again from boundary condition $ii(a)$ and with the help of Eqs. (15), (37) and (38), we have

$$\begin{aligned} &\mu_2 \left[R_1 N - R_2 N - \frac{B M (ik + M)}{M^2 - N^2} R_3 + \frac{B M (ik - M)}{M^2 - N^2} R_4 \right] \\ &+ ik(s_3 M_{44} - M_{46}) B_3 \\ &= 0. \end{aligned} \tag{42}$$

Similarly from $ii(b)$ and with the aid of Eqs. (15) and (38), we get

$$R_1 + R_2 - \frac{B(ik + M)}{M^2 - N^2} R_3 - \frac{B(ik - M)}{M^2 - N^2} R_4 - B_3 = 0. \tag{43}$$

Similarly from $ii(c)$ and with the aid of Eq. (16), we get

$$R_3 - R_4 = 0. \tag{44}$$

Thus, for the non-zero solution, eliminating $A_1, R_1, R_2, R_3, R_4, B_3$ we have

$$|D_{ij}| = 0 \tag{45}$$

where $|D_{ij}|$ is the determinant of the coefficient matrix extracted from the Eqs. (39)–(44).

On simplification of (45), the equation refers to two different wave fronts. The first wave front which corresponds to the elasticity of the medium can be written as:

$$\tan \left[kH \sqrt{\frac{c^2}{\alpha/\rho_2 k} - 1 - \frac{1}{k^2 \left(\frac{z}{c} \right)}} \right] = 0. \tag{46}$$

The required dispersion relation deduced by equating the real parts of the second factor which refers to the attenuation of the wave can be expressed as:

$$\tan(N_1 H) = \frac{2e^{aH} N_1 \mu_1 \mu_2 a [k^2 R_{44}^2 s_3^2 - k^2 R_{46}^2 + 2k N_1 \mu_2 R_{46} - \mu_2^2 N_1^2]}{k^2 (a^2 + 4m_1^2) \mu_1^2 s_3^2 R_{44}^2 - 8e^{aH} k m_1 R_{44} N_1 s_3 \mu_1 \mu_2 (k R_{46} - N_1 \mu_2) + 4e^{2aH} N_1^2 \mu_2^2 (k R_{46} - N_1 \mu_2)^2}. \tag{47}$$

Similarly, equating the imaginary parts from the second factor which refers to the damping equation of the wave can be written as:

$$\frac{(a^2 + 4m_1^2)\mu_1^2 + 4e^{2aH}N_1^2\mu_2^2}{4\{k^2R_{44}^2s_3^2 + (kR_{46} - N_1\mu_2)^2\}} - \frac{e^{aH}m_1N_1\mu_1\mu_2}{kR_{44}s_3(kR_{46} - N_1\mu_2)} = 0 \tag{48}$$

where

$$N_1 = k\sqrt{\left(\frac{c^2}{A^2} - 1\right)}, \quad A = \sqrt{\frac{\mu_2}{\rho_2}}$$

The study laid down an emphasis on the fact that SH-waves in elastic media with void pores transmit in two wave fronts. First wave front (46) is related with the parameters of void pores that are involved with the change in void volume fraction and equilibrated inertia. The another wave front (47) relates the elastic medium without void pores. It is clear that both the fronts are dispersive in nature.

Special cases

Case I

When $m_H = 0$, which implies $R_{44} = C_{44}$, $R_{46} = C_{46}$ and $R_{66} = C_{66}$, i.e. the lower half-space is simply anisotropic, then the dispersion Eq. (47) can be written as:

$$\tan(N_1H) = \frac{2e^{aH}N_1\mu_1\mu_2a[k^2C_{44}^2s_3^2 - k^2C_{46}^2 + 2kN_1\mu_2C_{46} - \mu_2^2N_1^2]}{k^2(a^2 + 4m_1^2)\mu_1^2s_3^2C_{44}^2 - 8e^{aH}km_1C_{44}N_1s_3\mu_1\mu_2(kC_{46} - N_1\mu_2) + 4e^{2aH}N_1^2\mu_2^2(kC_{46} - N_1\mu_2)^2} \tag{49}$$

Case II

When $m_H = 0$ and $C_{44} = C_{66} = \mu_3$, $C_{46} = 0$ which implies $R_{44} = R_{66} = \mu_3$, $R_{46} = 0$, $s_3 = \sqrt{\frac{c^2}{\mu_3/\rho_3} - 1}$ i.e. the half-space becomes isotropic, then the dispersion Eq. (47) takes the form

$$\tan\left(kH\sqrt{\frac{c^2}{\mu_2/\rho_2} - 1}\right) = \frac{2e^{aH}\mu_1\mu_2ak\sqrt{\frac{c^2}{\mu_2/\rho_2} - 1}\left[\mu_3^2\left(\frac{c^2}{\mu_3/\rho_3} - 1\right) - \mu_2^2\left(\frac{c^2}{\mu_2/\rho_2} - 1\right)\right]}{(a^2 + 4m_1^2)\mu_1^2\mu_3^2\left(\frac{c^2}{\mu_3/\rho_3} - 1\right) - 8e^{aH}km_1\mu_1\mu_2^2\mu_3\left(\frac{c^2}{\mu_2/\rho_2} - 1\right)\sqrt{\frac{c^2}{\mu_3/\rho_3} - 1} - 4e^{2aH}\mu_2^3k\left(\frac{c^2}{\mu_2/\rho_2} - 1\right)^{\frac{3}{2}}} \tag{50}$$

where μ_3 being the rigidity of that elastic material.

Numerical calculations and discussion

To graphically illustrate the effect of inhomogeneity parameter, thickness of layer, angle at which the wave crosses the magnetic field the following values are taken. For the exponentially varying inhomogeneous upper semi-infinite medium, the values are taken from Gubbins (1990) which are as follows:

$$\mu_1 = 6.54 \times 10^{10} \text{ N/m}^2, \quad \rho_1 = 3409 \text{ kg/m}^3$$

In case elastic layer with void pores, the values for void parameters are taken from Dey et al. (2004) whereas values for rigidity and inhomogeneity are taken as:

$$\mu_2 = 7.5 \times 10^9 \text{ N/m}^2, \quad \rho_2 = 1973 \text{ kg/m}^3$$

The data given by Zakharenko (2005) have been considered for the anisotropic magnetoelastic materials which reflect a negligible piezoelectric and piezomagnetic effect on propagation of surface waves.

Figures 2, 3, and 4 depict the effect of the angle with which the wave crosses the magnetic field on phase velocity together with dimensionless wave number (kH). Figures 2, 3, and 4 represent the curves for three different magnetoelastic materials Coesite, Diopside and Jadeite (Table 1), respectively. In all three cases, for a fixed

dimensionless wave number the phase velocity increases profoundly with the gradual increase of angle with which the wave crosses the magnetic field. In all cases, the fixed values of inhomogeneity parameter (a), magnetoelastic coupling parameter (m_H) and the thickness of layer (H) are taken as 0.5, 0.8 and 1.0, respectively.

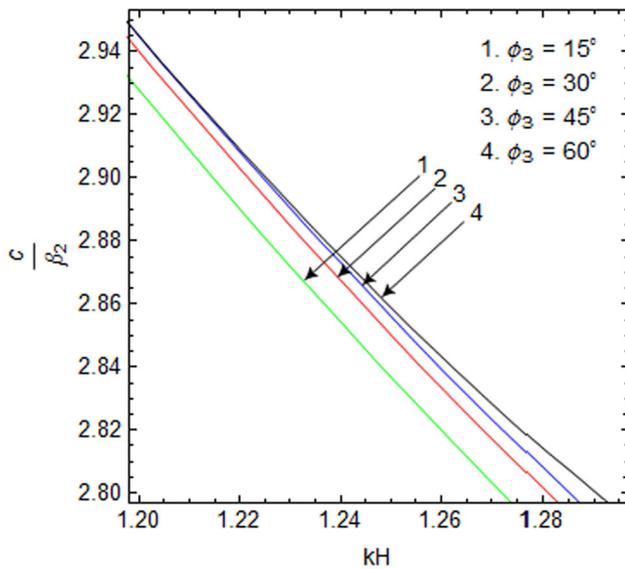


Fig. 2 Variation of dimensionless phase velocity and dimensionless wave number against the angle at which the wave crosses the magnetic field for Coesite (SiO_2)

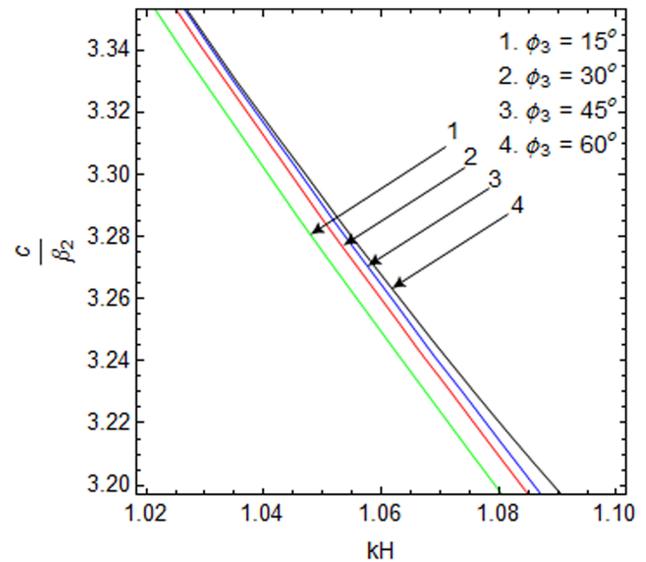


Fig. 4 Variation of dimensionless phase velocity and dimensionless wave number against the angle at which the wave crosses the magnetic field for Jadeite ($\text{NaAlSi}_2\text{O}_6$)

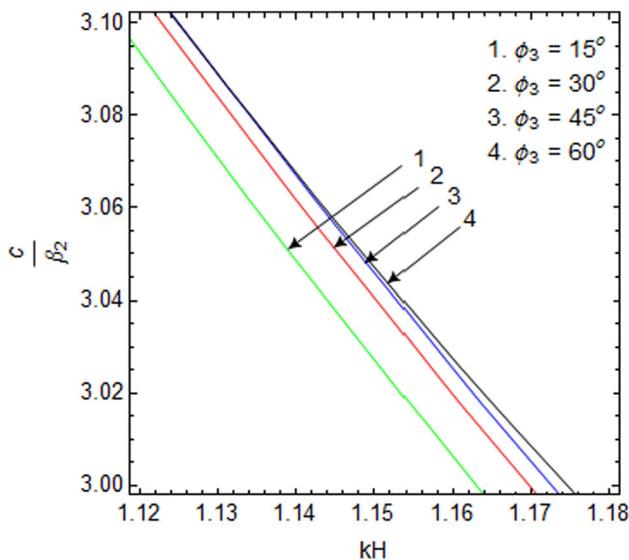


Fig. 3 Dimensionless phase velocity versus dimensionless wave number against the angle at which the wave crosses the magnetic field for Diopside ($\text{CaMgSi}_2\text{O}_6$)

In Figs. 5, 6, and 7, curves are drawn to show the variations of dimensionless phase velocity ($\frac{c}{\beta_2}$) with dimensionless wave number (kH) for the different values (0.6, 0.8, 1.0, 1.2) of magnetoelastic coupling parameter and for three different magnetoelastic materials, namely Coesite, Diopside and Jadeite, respectively. It is evident that the phase velocity shows a gradual increase with the increase of magnetoelastic coupling parameter (m_H). The fixed values of the inhomogeneity parameter (a), angle (ϕ_3) with which the wave crosses magnetic field and the

thickness of layer (H) are taken as 0.5, 15° and 1.0, respectively.

Variations in the inhomogeneity parameter (a) with dimensionless phase velocity and dimensionless wave number have been shown in Fig. 8 for Coesite material. It is quiet clear that as the inhomogeneity parameter increases the phase velocity and also increases accordingly. The fixed values of thickness of layer (H), angle (ϕ_3) with which the wave crosses magnetic field and magnetoelastic coupling parameter (m_H) are taken as 1.0, 15° and 0.8, respectively.

The curves in Fig. 9 trace out the effect of the thickness of intermediary layer (H) with dimensionless phase velocity and dimensionless wave number for Coesite material. It illustrates that there would be a gradual diminution on phase velocity with the decrease of the thickness of layer for a fixed wave number. The values of inhomogeneity parameter (a), magnetoelastic coupling parameter (m_H) and angle (ϕ_3) with which the wave crosses the magnetic field are considered as 0.5, 0.8 and 15°

The surface plot in Fig. 10 focuses on the combined effect of dimensionless phase velocity and dimensionless wave number and angle (ϕ_3) with which the wave crosses magnetic field in case of Coesite. The inhomogeneity parameter (a), magnetoelastic coupling parameter (m_H) and thickness of layer (H) are taken as 0.5, 0.8 and 1.0, respectively.

The surface plot in Fig. 11 demonstrates the joint effect of magnetoelastic coupling parameter (m_H) with dimensionless phase velocity and dimensionless wave number for Coesite material. The inhomogeneity parameter (a), angle (ϕ_3) with

Table 1 Material constants of monoclinic crystals (point group symmetry of 2/m)

| Magnetoelastic solids | C_{44} (N/m ²) | C_{46} (N/m ²) | C_{66} (N/m ²) | Density ρ_3 (kg/m ³) |
|--|------------------------------|------------------------------|------------------------------|---------------------------------------|
| Coesite (SiO ₂) | 6.78×10^{10} | 1.0×10^{10} | 5.88×10^{10} | 2920 |
| Diopside (CaMgSi ₂ O ₆) | 7.4×10^{10} | 0.73×10^{10} | 6.6×10^{10} | 3290 |
| Jadeite (NaAlSi ₂ O ₆) | 8.8×10^{10} | 1.3×10^{10} | 9.4×10^{10} | 3330 |

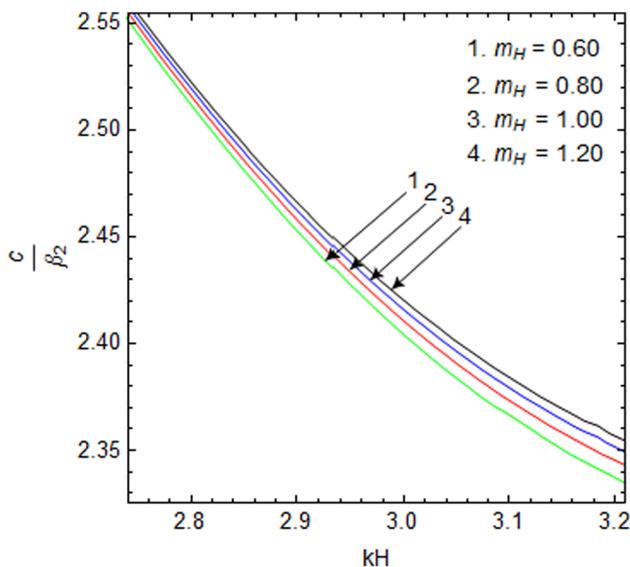


Fig. 5 Dimensionless phase velocity versus dimensionless wave number against magnetoelastic coupling parameter m_H for Coesite (SiO₂)

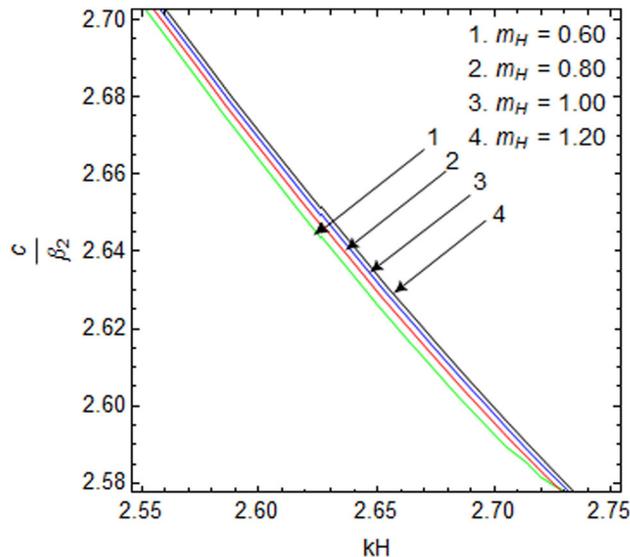


Fig. 7 Dimensionless phase velocity versus dimensionless wave number against magnetoelastic coupling parameter m_H for Jadeite (NaAlSi₂O₆)

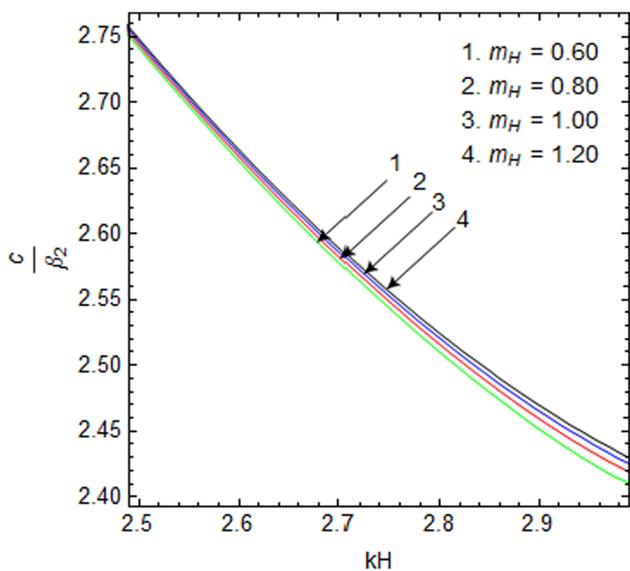


Fig. 6 Variation of dimensionless phase velocity and dimensionless wave number against magnetoelastic coupling parameter m_H for Diopside (CaMgSi₂O₆)

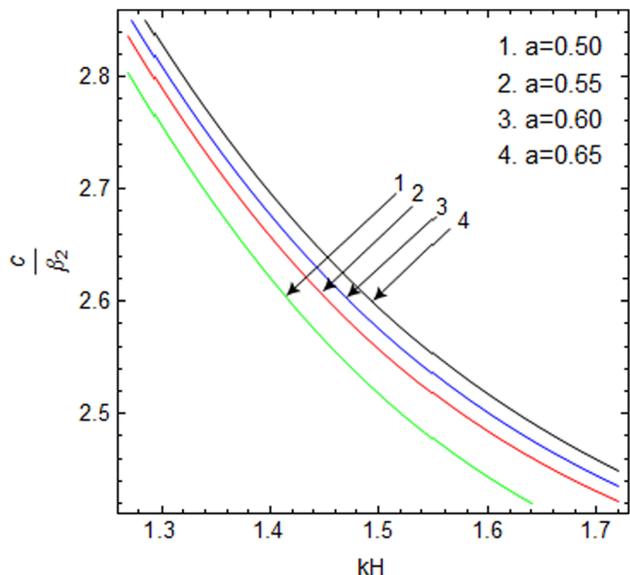


Fig. 8 Variation of dimensionless phase velocity and dimensionless wave number against inhomogeneity parameter a for Coesite (SiO₂)

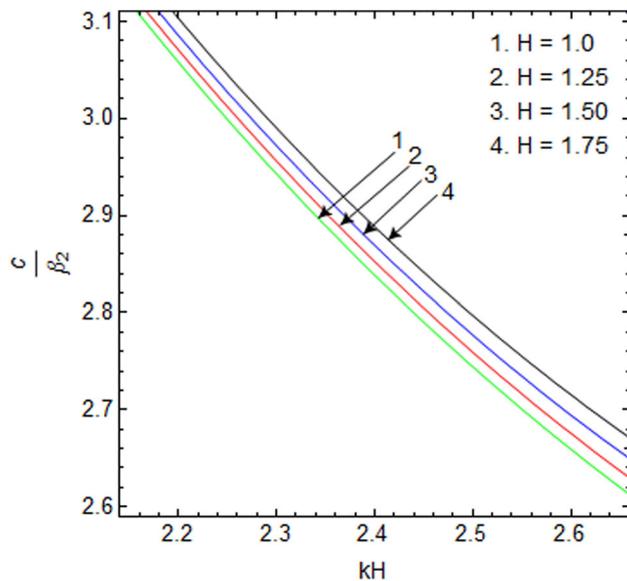


Fig. 9 Dimensionless phase velocity versus dimensionless wave number against thickness H of intermediary layer for Coesite (SiO_2)

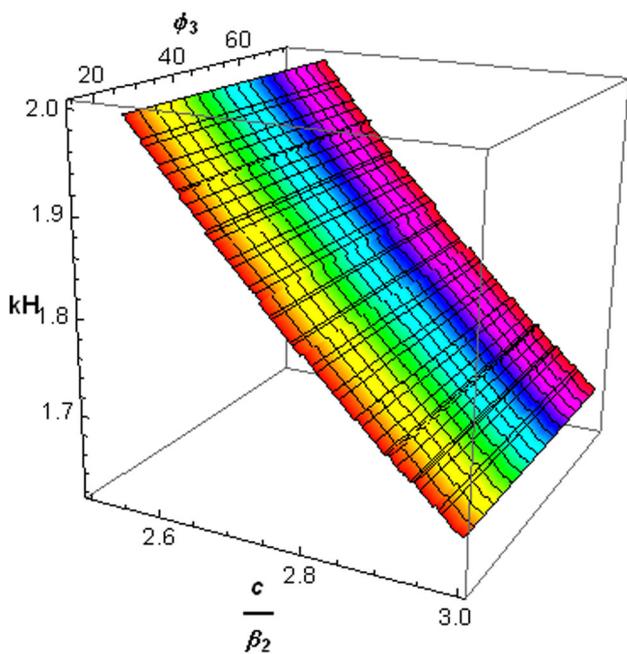


Fig. 10 Variation of dimensionless phase velocity ($\frac{c}{\beta_2}$) against dimensionless wave number and the angle ϕ_3 at which the wave crosses the field for Coesite (SiO_2)

which the wave crosses the magnetic field and thickness of layer (H) are taken as 0.5, 15° and 1.0, respectively.

Conclusions

An analytical and numerical approach has been made to investigate the transmission of horizontally polarized shear waves in elastic void layer welded between

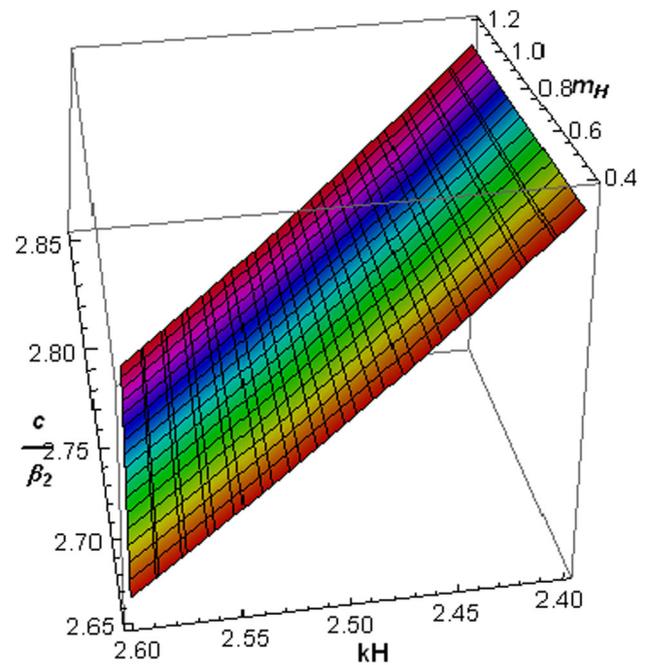


Fig. 11 Variation of dimensionless phase velocity ($\frac{c}{\beta_2}$) against dimensionless wave number and the magnetoelastic coupling parameter for Coesite (SiO_2)

vertically inhomogeneous and anisotropic magnetoelastic half-spaces. The frequency relations have obtained for both the wave fronts associated with elasticity of medium and change in void volume. There exists a damping part in the dispersion relation related to the change in void volume as well. Three different magnetoelastic materials which negligibly react to the piezoelectric and piezomagnetic fields have been taken into account to represent the effect of various parameters on the propagation of shear waves.

The observations in this study arrive with the conclusions that

- (i) Numerical results represent that the angle with which the wave crosses the magnetic field largely affects the phase velocity of the shear waves. For different set of anisotropic magnetoelastic materials, there are clear difference in phase velocities.
- (ii) For a fixed dimensionless wave number, the phase velocity increases with the gradual increase of magnetoelastic coupling parameter (m_H) and the same results have been found for different set of anisotropic magnetoelastic materials.
- (iii) The inhomogeneity parameter (a) has significant effect on the phase velocity. As the inhomogeneity decreases, the phase velocity also decreases accordingly.
- (iv) The thickness (H) of elastic layer with void pores also affects the shear wave transmission. The

diminution of thickness of elastic void layer results in a decrease in phase velocity.

Within the theoretical framework, the outcome of the investigated three layered earth model has certain practical applications to gather a clear insight to explore the natural resources lying abundantly in the form of mineral deposits beneath the earth's surface. Moreover, the results are found to be helpful in non-destructive measurements of mechanical and electromagnetic parameters of the materials found in asthenosphere.

Acknowledgements The fellowship received from University Grants Commission, New Delhi through Grant No. F.7-79/2007(BSR) is gratefully acknowledged by the authors.

References

- Acharya DP, Roy I, Sengupta S (2009) Effect of magnetic field and initial stress on the propagation of interface waves in transversely isotropic perfectly conducting media. *Acta Mech* 202(1–4):35–45
- Achenbach JD (1973) Wave propagation in elastic solids. North-holland series in applied mathematics and mechanics. North-Holland Publishing Company, New York
- Biot MA (1965) Mechanics of incremental deformations. Wiley, New York
- Birsan M (2000) Existence and uniqueness of weak solution in the linear theory of elastic shells with voids. *Lib Math* 20:95–106
- Bullen KE, Bolt BA (1985) An introduction to the theory of seismology. Cambridge University Press, Cambridge
- Chattaraj R, Samal SK (2013) Love waves in the fiber-reinforced layer over a gravitating porous half-space. *Acta Geophys* 61(5):1170–1183
- Chattopadhyay A, Choudhury S (1990) Propagation, reflection and transmission of magnetoelastic shear waves in a self-reinforced medium. *Int J Eng Sci* 28(6):485–495
- Chattopadhyay A, Gupta S, Kumari P, Sharma VK (2012) Effect of point source and heterogeneity on the propagation of SH-waves in a viscoelastic layer over a viscoelastic half space. *Acta Geophys* 60(1):119–139
- Chaudhary S, Kaushik VP, Tomar SK (2004) Reflection/transmission of plane SH-waves through a self-reinforced elastic layer between two half-spaces. *Acta Geophys Pol* 52(2):219–235
- Chirita S, Ghiba LD (2010) Inhomogeneous plane waves in elastic materials with voids. *Wave Motion* 47(6):333–342
- Cowin SC (1985) The viscoelastic behavior of linear elastic materials with voids. *J Elast* 15(2):185–191
- Cowin SC, Nunziato JW (1983) Linear elastic materials with voids. *J Elast* 13(2):125–147
- Dey S, Gupta AK, Gupta S (2000) Torsional surface waves in nonhomogeneous anisotropic medium under initial stress. *J Eng Mech* 126(11):1120–1123
- Dey S, Gupta S, Gupta AK (2004) Propagation of Love waves in an elastic layer with void pores. *Sādhanā* 29(4):355–363
- Dunkin JW, Eringen EC (1963) On the propagation of waves on electromagnetic elastic solids. *Int J Eng Sci* 1:461–495
- Gubbins D (1990) Seismology and plate tectonics. Cambridge University Press, London
- Ke LI, Wang YS, Zhang ZM (2006) Love waves in an inhomogeneous fluid saturated porous layered half-space with linearly varying properties. *Soil Dyn Earthq Eng* 26(6):574–581
- Knopoff L (1955) The interaction between elastic wave motion and a magnetic field in electrical conductors. *J Geophys Res* 60:441–456
- Kumar S, Pal PC (2014) Wave propagation in an inhomogeneous anisotropic generalized thermoelastic solid under the effect of gravity. *Comput Therm Sci* 6(3):241–250
- Kumar S, Pal PC, Majhi S (2015) Reflection and transmission of plane SH-waves through an anisotropic magnetoelastic layer sandwiched between two semi-infinite inhomogeneous viscoelastic half-spaces. *Pure Appl Geophys* 172(2):2621–2634
- Love AEH (1927) Mathematical theory of elasticity. Dover, New York
- Magdalena I, Pudjaprasetya SR, Wiryanto LH (2014) Wave interaction with an emerged porous media. *Adv Appl Math Mech* 6(5):680–692
- Manolis GD, Shaw RP (1996) Greens function for the vector wave equation in a mildly heterogeneous continuum. *Wave Motion* 24:59–83
- Manolis GD, Rangelov TV, Shaw RP (2002) Conformal mapping methods for variable parameter elastodynamics. *Wave Motion* 36:185–202
- Maugin GA (1998) Continuum mechanics of electromagnetic solids. Elsevier, Amsterdam
- Nunziato J, Cowin SC (1979) A nonlinear theory of elastic materials with voids. *Arch Ration Mech Anal* 72(2):175–201
- Pal PK, Acharya D (1998) Effects of inhomogeneity on surface waves in anisotropic media. *Sādhanā* 23(3):247–258
- Pal PC, Mandal D (2014) Generation of SH-type waves due to shearing stress discontinuity in a sandy layer overlying an isotropic and inhomogeneous elastic half-space. *Acta Geophys* 62(1):44–58
- Qian Z, Jin F, Lu TJ (2008) Transverse surface waves in functionally graded piezoelectric materials with exponential variation. *Smart Mater Struct* 17:065005
- Samal SK, Chattaraj R (2011) Surface wave propagation in fiber-reinforced anisotropic elastic layer between liquid saturated porous half space and uniform liquid layer. *Acta Geophys* 59(3):470–482
- Singh AK, Das A, Lakshman A, Chattopadhyay A (2016a) Effect of corrugation and reinforcement on the dispersion of SH-wave propagation in corrugated poroelastic layer lying over a fibre-reinforced half-space. *Acta Geophys* 64(5):1340–1369
- Singh AK, Mistri KC, Das A (2016b) Propagation of Love-type wave in a corrugated fibre-reinforced layer. *J Mech* 32(6):693–708
- Wang YS, Zhang ZM (1998) Propagation of Love waves in a transversely isotropic fluid-saturated porous layered half-space. *J Acoust Soc Am* 103(2):695–701
- Zakharenko AA (2005) Analytical studying the group velocity of three-partial Love (type) waves in both isotropic and anisotropic media. *Nondestruct Test Eval* 20(4):237–254

Volcanic ash cloud detection from MODIS image based on CPIWS method

Lan Liu¹ · Chengfan Li¹ · Yongmei Lei¹ · Jingyuan Yin² · Junjuan Zhao¹

Received: 20 January 2017 / Accepted: 25 January 2017 / Published online: 9 February 2017
© Institute of Geophysics, Polish Academy of Sciences & Polish Academy of Sciences 2017

Abstract Volcanic ash cloud detection has been a difficult problem in moderate-resolution imaging spectroradiometer (MODIS) multispectral remote sensing application. Principal component analysis (PCA) and independent component analysis (ICA) are effective feature extraction methods based on second-order and higher order statistical analysis, and the support vector machine (SVM) can realize the nonlinear classification in low-dimensional space. Based on the characteristics of MODIS multispectral remote sensing image, via presenting a new volcanic ash cloud detection method, named combined PCA-ICA-weighted and SVM (CPIWS), the current study tested the real volcanic ash cloud detection cases, i.e., Sangeang Api volcanic ash cloud of 30 May 2014. Our experiments suggest that the overall accuracy and Kappa coefficient of the proposed CPIWS method reach 87.20 and 0.7958%, respectively, under certain conditions with the suitable weighted values; this has certain feasibility and practical significance.

Keywords Remote sensing image · Principal component analysis (PCA) · Independent component analysis (ICA) · Support vector machine (SVM) · Volcanic ash cloud

Introduction

Volcanic ash cloud monitoring from moderate-resolution imaging spectroradiometer (MODIS) multispectral remote sensing images is an important research topic in the field of aviation safety and multispectral remote sensing applications (Chang et al. 2002; Duda and Canty 2002; Wang et al. 2006; Corradini et al. 2008; Mackie et al. 2014; Li et al. 2015), and it has been widely applied in recent years along with the rapid development of computer interpretation technology. Compared with the conventional digital images, the multispectral remote sensing image characteristics are much more complicated (Liang et al. 2010; Mountrakis et al. 2011; Wen et al. 2011; Prata and Prata 2012; Sahin et al. 2012; Spinetti et al. 2013; Western et al. 2015); not only do they have a lot of spectrum bands in which there is a huge data redundancy, but also there are many kinds of random-distributed earth objects in each spectrum band bearing extremely complex spectrum characteristics. In addition, it has also been influenced by the sunshine time, imaging angle, slope direction, and so on (Xu et al. 2010; Corradini et al. 2011; Heblinski et al. 2011; Sanchez-Azofeifa et al. 2011; Yi et al. 2011; Montopoli et al. 2013). Therefore, detecting the volcanic ash cloud from remote sensing images using remote sensing technique has been regarded as an effective way of monitoring the volcanic ash cloud.

Volcanic ash cloud detection from MODIS image consists in building a set of new features according to the existing combination in remote sensing image (Ward and Starks 1999; Segl et al. 2003). Generally speaking, the principal component analysis (PCA), independent component analysis (ICA), and support vector machine (SVM) method, with the huge ability of data compression and redundancy elimination, and being easy to calculate and analyze, are widely used at present (Lee et al. 1999;

✉ Chengfan Li
david-0904@163.com

Jingyuan Yin
lchf@shu.edu.cn

¹ School of Computer Engineering and Science, Shanghai University, Shanghai, China

² Earthquake Administration of Shanghai Municipality, Shanghai, China

Hyvarinen 1999; Hillger and Clark 2002a, b). Although these methods have their respective limitations in the actual application, they also have great application potential in the volcanic ash cloud detection from MODIS multispectral remote sensing image. On the basis of the previous work (Schumann et al. 2011; Christopher et al. 2012; Li and Yin 2013; Centre LVAA 2014; Li et al. 2014, 2015), this paper proposes a new volcanic ash cloud detection method combining PCA-ICA-weighted and SVM (CPIWS) for MODIS multispectral remote sensing image, which could better implement the data compression and elimination of redundancy from multispectral remote sensing image and obtain highly accurate volcanic ash cloud components. To further understand the CPIWS method, the current study investigates the two classic volcanic ash cloud cases using the proposed method. The experimental result shows that the proposed CPIWS method is correct and feasible, and improves the detection accuracy of multispectral remote sensing image to some extent. The main idea and structure of this paper are as follows: the first section is the introduction; the second section is the basic theory of PCA, ICA, and SVM and the proposed CPIWS method; the third section is the classification experiment of MODIS image; the fourth and fifth sections are the classic volcanic ash cloud cases, and we end our study with conclusions and discussions in the last section.

Basic theory

PCA method

As one of the most common data analysis methods in statistics, PCA consists in finding a set of variables (sometimes also called principal component information) with dimensional invariance, and unrelated with each other in terms of the linear transformation technology, mainly used in the field of data compression and feature extraction.

In essence, the remote sensing image can be seen as a linear superposition by a set of independent image bases (Hyvarinen and Oja 2000; Hillger and Clark 2002a, b), aiming at the remote sensing image X ,

$$X = (x_1, x_2, \dots, x_N) = (X_1, X_2, \dots, X_P)^T,$$

where N is the number of pixels, P is the number of bands, x_i ($i = 1, 2, \dots, N$) is the i -th pixel of remote sensing X , and x_k ($k = 1, 2, \dots, P$) is the k -th band of remote sensing X .

Step 1: Calculate the mean $m = \frac{1}{N} \sum_{k=1}^N x_k$ and covariance matrix $\sum_{\lambda} = E\{(X - m)(X - m)^T\}$ of the vector, where

the unbiased estimation of the covariance matrix can be expressed as $\sum_{\lambda} = \frac{1}{P-1} \sum_{j=1}^P (x_j - m)(x_j - m)^T$.

Step 2: Calculate the maximum eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_P \geq 0$ and corresponding unit feature vector t_1, t_2, \dots, t_P of the covariance matrix.

Step 3. Transform the remote sensing image X using transformation matrix $A = A^T$, and $Y = [Y_1, Y_2, \dots, Y_P] = AX$. Here, $Y_i = [Y_{i1}, Y_{i2}, \dots, Y_{iP}]^T$ is the i -th principal component of X .

ICA method

The PCA method is used to find some mutually orthogonal axes along with the maximum variance of data sets. ICA evolved from PCA method. Compared with the PCA method, ICA is an analysis method based on high-order statistical characteristics such as fourth-order kurtosis and better conforms to practical applications. Specifically, ICA can identify independent component information from non-Gaussian signals, whereas PCA only can get orthogonal vectors from mixed signals in terms of the variance information. Therefore, ICA method not only obtains the mutually independent components, but also realizes the PCA's decorrelation.

In actual application, ICA method is usually used for solving a set of mutually independent image bases from the multispectral remote sensing images, and these image bases are used to build a subspace, and then the feature information of remote sensing images is further extracted according to the projection coefficient in the subspace images. A brief comparison between PCA and ICA method is shown in Table 1.

Regarding the remote sensing image X , it can be a linear combination of statistically independent image basis s_1, s_2, \dots, s_n , and can be expressed in the form of matrix as $X = AS$, where A is the unknown mixed matrix, $x = \{x_1(t), x_2(t), \dots, x_n(t)\}^T$ is the observation signal, and $s = \{s_1, s_2, \dots, s_n\}^T$ ($m \geq n$) are the component variables.

Step 1: Assuming that s is a statistically independent and non-Gaussian distribution, the number of source signals $s_i(t)$ is equal to the number of the observation signals.

Step 2: Estimate the unknown mixed matrix A and the estimated value y .

Step 3: Calculate the inverse matrix A^{-1} .

Step 4: Calculate the estimate value y of s , $y = Wx$, $y = [y_1, y_2, \dots, y_p]^T$.

Table 1 Comparison between PCA and ICA method

| | PCA method | ICA method |
|---------------------------|---------------------|------------------------------|
| Signal type | Gaussian | Non-Gaussian |
| Statistical property | Second-order | Higher order |
| Relationship of component | Orthogonal | Orthogonal or non-orthogonal |
| Computational complexity | $O[\min(m, n)^3]$ | $O(m^{2.5}, n)$ |
| Results | Principal component | Independent component |

m is the number of observation signal, and n is the number of original signal

SVM method

SVM is a learning method whose core idea is that the hardly linear partition sample data was nonlinear mapping to high dimension space characteristics, and seeking the best compromise between model complexity and learning ability according to limited sample information and structural risk minimization, and the nonlinear classification of sample data were finally fulfilled (Li et al. 2014; Liang et al. 2010; Wang et al. 2007; Mountrakis et al. 2011).

Regarding the remote sensing image, SVM can realize the nonlinear classification and display by sort.

Step 1 Carry on the differentiations of different earth object types and select each earth object type's training samples. Assuming that $T = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\} \in (X, Y)^l$, where $x_i \in X \in R^P, y_i \in Y = \{-1, 1\}, i = 1, 2, \dots, l$.

Step 2 Calculate the optimal solution between the object samples and other feature samples: $a^* = (a_1^*, a_2^*, \dots, a_l^*)^T$. The optimal hyperplane for classification is $w^T x_i + b = 0$, and then the weight value w and bias b should satisfy the following conditions:

$$y_i(\omega^T x_i + b) \geq 1 - \zeta_i,$$

where ζ_i is a slack variable, $\zeta_i \geq 0, i = 1, 2, \dots, N$.

Step 3 Acquire the support vector sets and their corresponding credibility of each sample feature. Calculate $\omega^* = \sum_{i=1}^l y_i a_i^* x_i$ and the positive component $b^* = y_i - \sum_{i=1}^l y_i a_i^* (x_i \cdot x_j)$ Construct the optimal hyperplanes for classification and obtain the decision function

$$f(x) = \text{sgn}\left(\sum_{i=1}^l a_i^* y_i (x \cdot x_i) + b^*\right)$$

Step 4 Obtain SVM input vector and construct the SVM classifier; classify and display the feature information by sort.

CPIWS method

The detailed process of the CPIWS method proposed in this paper is as follows:

Step 1: Feature weighting of PCA and ICA. For the multispectral remote sensing image x, x_p and x_i represent, respectively, the extracted features by PCA and ICA; y represents the weighted feature and $y = wx_i + (1 - w)x_p$. Assuming that there is N -dimensional remote sensing image $x_1, x_2, x_3, \dots, x_N$, we have the corresponding weighted feature $(y_1, y_2, y_3, \dots, y_N), i = 1, 2, 3, \dots, N$.

Step 2: Distance similarity measure. For the given unclassified remote sensing image y , the distance similarity measure equation between y and different spectrum bands image y_i can be expressed as $d(u, v) = \|y - y_i\|^2, i = 1, 2, 3, \dots, N$. Hence, we have the similarity measure $S(y, y_i) = 1 - \frac{d(y, y_i) - \text{mind}(y, y_i)}{\max_i d(y, y_i) - \text{mind}(y, y_i)}, i = 1, 2, 3, \dots, N$.

Step 3: Recognition rule. For the unclassified multi-spectral remote sensing image y , the similarity between y and each spectrum band image y_i is calculated and summarized into the most similar training samples class j according to the similarity value. Based on the distance similarity measure equation, the class j has the following form: $j = \text{Arg max}_i \{S(y, y_i)\}$.

Step 4: Construct SVM classifier, classification, and feature information display. In our study, all 16 components obtained by weighted PCA and ICA are treated as the input feature statistics because of the low spatial resolution of MODIS multispectral remote sensing image and few texture features. The radial basis function (RBF) is the optimal kernel function and the corresponding suitable parameters are penalty coefficient and interval coefficient.

Classification experiment of modis image

Experimental data and environment

As the updated sensor of advanced very high-resolution radiometer (AVHRR), the MODIS remote sensor is carried in the TERRA/AQUA, and the average height of MODIS sensor orbit which belongs to sun-synchronous orbits is 705 km. Table 2 shows the detailed band distribution and main applications. As can be seen from Table 2, it has a total of 36 spectrum band images (image sets) whose spectrum ranges between 0.62 and 14.385 μm . Different bands have different application fields. Therefore, the MODIS multispectral remote sensing image contains visible and near-infrared bands (i.e., bands 1–19 and 20) and thermal infrared bands (i.e., bands 20–36). For the volcanic ash cloud, in essence the detection of volcanic ash cloud from remote sensing image lies in separating the volcanic ash cloud from the whole cloud information; the wavelength of MODIS sensor thermal infrared bands is between 3.660 and 14.385 μm and the main applications are surface temperature of sea, land, and cloud layer. Therefore, in the actual applications the thermal infrared bands are often adopted in the volcanic ash cloud detection based on satellite remote sensing data.

All experiments are complemented on Inter (R) Core (TM) i5-2400 CPU @ 3.10 GHz, 2 G memory, Windows platform, and the program code is written and run in ENVI 4.6 and MATLAB 7.0.

Experimental process

Suppose that ω is the weighted value of PCA, and $0 \leq \omega \leq 1$; thus, the weighted value of ICA is given by $1 - \omega$. In this experiment,

1. The weighted value $\omega = 0$, which means that only PCA method was used.
2. The weighted value $\omega = 1$, which means that only ICA method is used.
3. The weighted value $\omega \in (0, 1)$, which means the PCA-ICA-weighted is used.

Table 2 The MODIS sensor band distribution and main applications

| Bands | Resolution [m] | Band width [μm] | Main applications |
|------------------------|----------------|------------------------------|---|
| 1–2 | 250 | 0.620–0.876 | Chlorophyll determination and landuse change |
| 3–7 | 500 | 0.459–2.155 | Difference of cloud, land, soil and cloud |
| 8–19 | 1000 | 0.405–0.965 | Chlorophyll determination and atmosphere monitoring |
| 20–36 (except band 26) | 1000 | 3.660–14.385 | Surface temperature of sea, land and cloud layer |
| 26 | 1000 | 1.360–1.390 | Detection of infrared cloud |

According to the characteristics of MODIS multispectral remote sensing image, the penalty coefficient C and interval coefficient γ are selected as the SVM kernel function parameters in the experiment, and $C = 900$ and $\gamma = 0.05$, respectively.

In this paper, the proposed CPIWS method is tested and evaluated on a real optical MODIS multispectral remote sensing image, 1000×2000 pixels, representing a researching area which is used in our empirical study. Each remote sensing spectrum band image is composed of four kinds of earth objects, i.e., volcanic ash cloud, meteorological cloud, seawater, and land. According to the previous works (Li et al. 2014, 2015), the visible bands of MODIS sensor are usually used to the chlorophyll determination and landuse change detection, and it is almost impossible to detect the diffusion change of volcanic ash cloud. Therefore, in this experiment the thermal infrared bands 20–36 are used to train in the classification, and the B20 is water absorption and should also be ignored in the actual application. Then, the top 5 remote sensing spectrum band images (ex.: B20–B24) serve as the training samples, so there are 20 earth object samples in the training set due to the fact that every band image has four kinds of earth objects. Similarly, the remaining 11 remote sensing spectrum band images (ex.: B25, B27–B36) serve as the test samples, and there are 44 earth object samples in the test samples.

Experimental results

Based on the test optical MODIS multispectral remote sensing image, because there are four kinds of earth objects (ex.: seawater, meteorological cloud, volcanic ash cloud and land) in every spectrum band image, the 500 sample points in the test image were randomly generated by program. And then the classification accuracy of volcanic ash cloud with CPIWS method could be got under different weighted value conditions by comparing with the real distribution information of earth object by visual interpretation method. The classification accuracy of CPIWS method from MODIS image is illustrated in Fig. 1. Here, the classification accuracy and the weighted value ω is between 0 and 1.

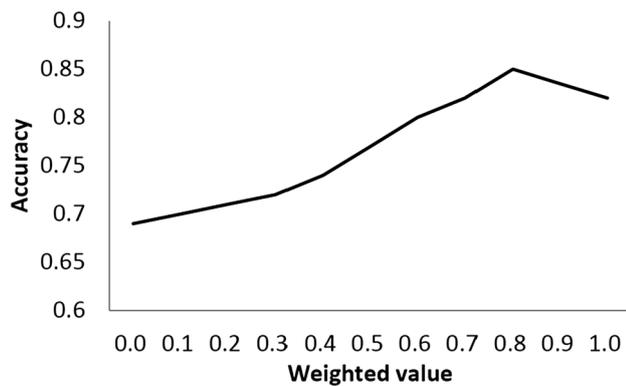


Fig. 1 Relationship between classification accuracy and weighted value

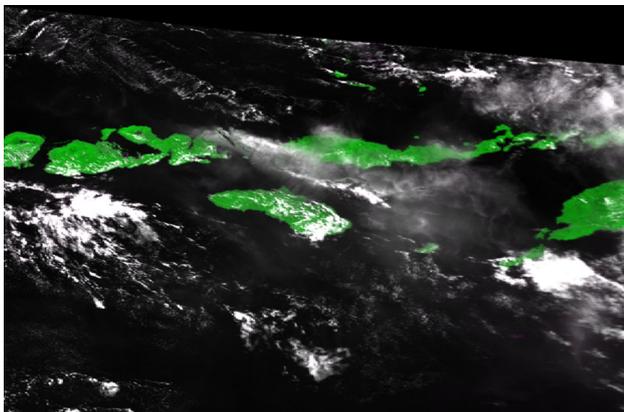


Fig. 2 MODIS image of Sangeang Api volcanic ash cloud on 30 May 2014

As can be seen from Fig. 1, we have concluded as follows:

1. When only PCA-SVM method is used for classification of MODIS multispectral remote sensing images, that is to say, $\omega = 0$, the classification accuracy of CPIWS reaches 69%.
2. When only ICA-SVM method is used for classification of MODIS multispectral remote sensing images, that is to say, $\omega = 1$, the classification accuracy of CPIWS reaches 82%.
3. From the perspective of using single PCA or ICA method, the classification accuracy of MODIS image achieved using ICA-SVM method is better than that achieved using PCA-SVM method.
4. When the weighted value $\omega = 0.8$, the classification accuracy of proposed CPIWS method reaches 85% and has good classification performance.

Sangeang api volcanic ash cloud case

General situation of study area

Sangeang Api, part of the Lesser Sunda Islands, is located in Indonesia and immediately south of the Equator, and it is also one of the most active islands in Lesser Sunda Islands. It is known as a volcanic island and began to erupt on 30 May 2014 (see Fig. 2). The enormous cloud of volcanic ash rose to a height of 15–20 km and spread to the continent of Australia by wind, subsequently causing the closure of most local airports and the cancellation of all the flights passing through were canceled.

Figure 2 shows the MODIS false color image combination of bands 2, 1, and 3 on 30 May 2014. The center wavelengths of three corresponding bands are 0.859, 0.645, and 0.469 μm , respectively. Figure 3 clearly represents the distribution of Sangeang Api volcanic ash clouds during the eruption in May 2014, and the contrast between volcanic ash cloud and other background factors, including sea water and island land, is very clear. However, the color of volcanic ash cloud in the image is close to that of the meteorological cloud, so it is very difficult to distinguish them just visually. In addition, based on the literature and the meteorological conditions, the MODIS images with better imaging conditions on 30 May 2014 is eventually used in this study.

Volcanic ash cloud classification

First, the weighted value $\omega = 0.8$ based on PCA and ICA is used to process the MODIS images. The obtained component images (CIs) of Sangeang Api on 30 May 2014 are shown in Fig. 3. It can be seen from Fig. 3 that CIs cover most of the volcanic ash cloud information.

Second, in order to verify the effect of the PCA/ICA weighted method, the histograms of the MODIS CIs and MODIS original spectrum images B20–B36 (except of B26) (see Fig. 4) of Sangeang Api volcanic ash cloud are analyzed statistically and calculated, respectively. Figures 5 and 6 show MODIS CIs' histogram with weighted PCA and ICA, and original false color images (B20–B36) of Sangeang Api. For such, the spectral feature in original spectrum images mainly reflects the calibrated radiance for each band, the spectral feature of MODIS CIs usually reflects the histogram distribution, and shows better spectral feature of different earth objects in remote sensing image.

As can be seen from Figs. 5 and 6, we conclude as follows:

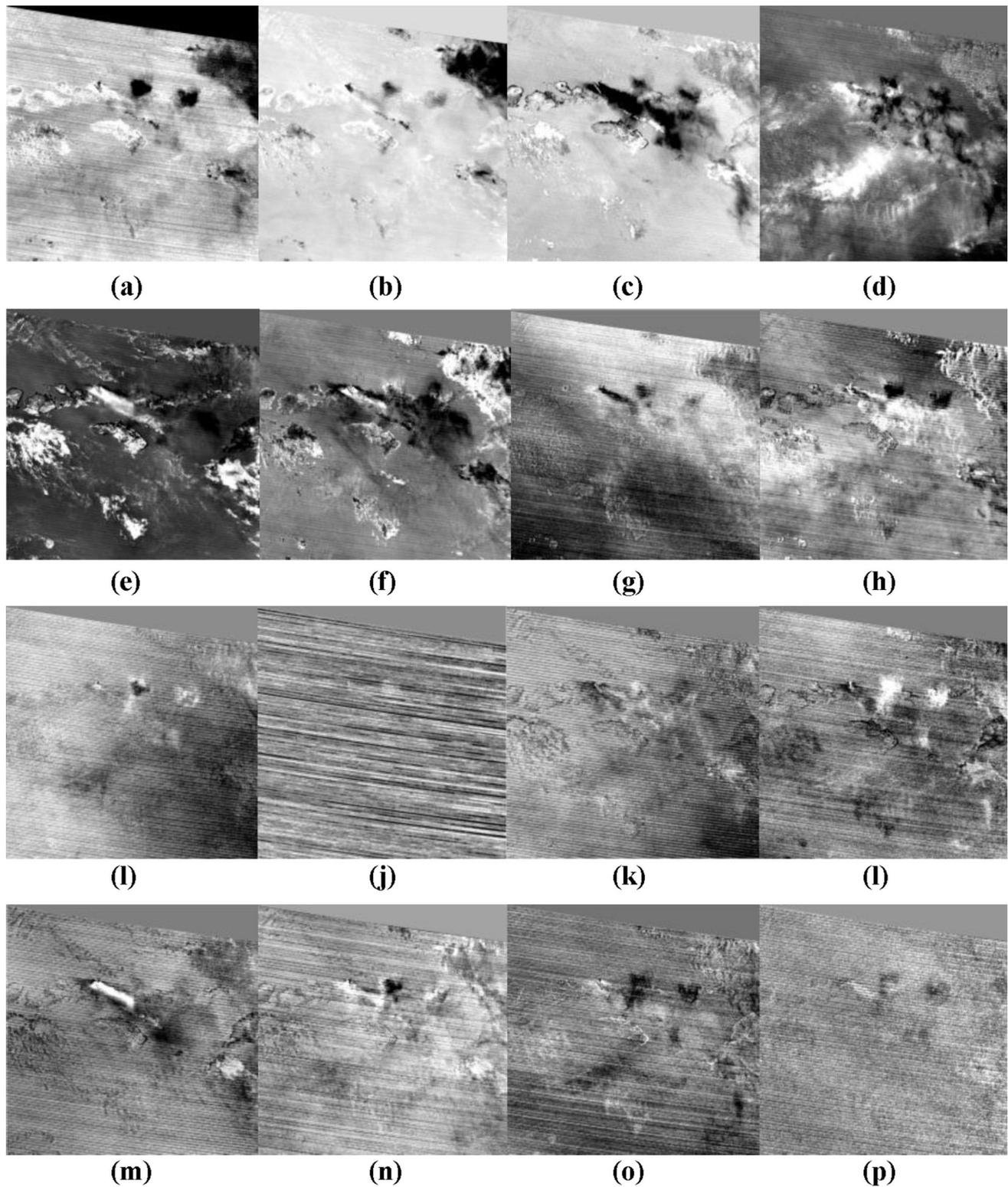


Fig. 3 MODIS CIs of Sangeang Api based on weighted PCA and ICA, **a–p** are the MODIS CI1–CI16, respectively

1. From the histogram shape point of view, the color histogram shape of MODIS original false color images of Sangeang Api volcanic ash cloud is very similar, and the distribution is very sparse, which means that these spectrum bands are highly correlated. In contrast, the color histogram shape of

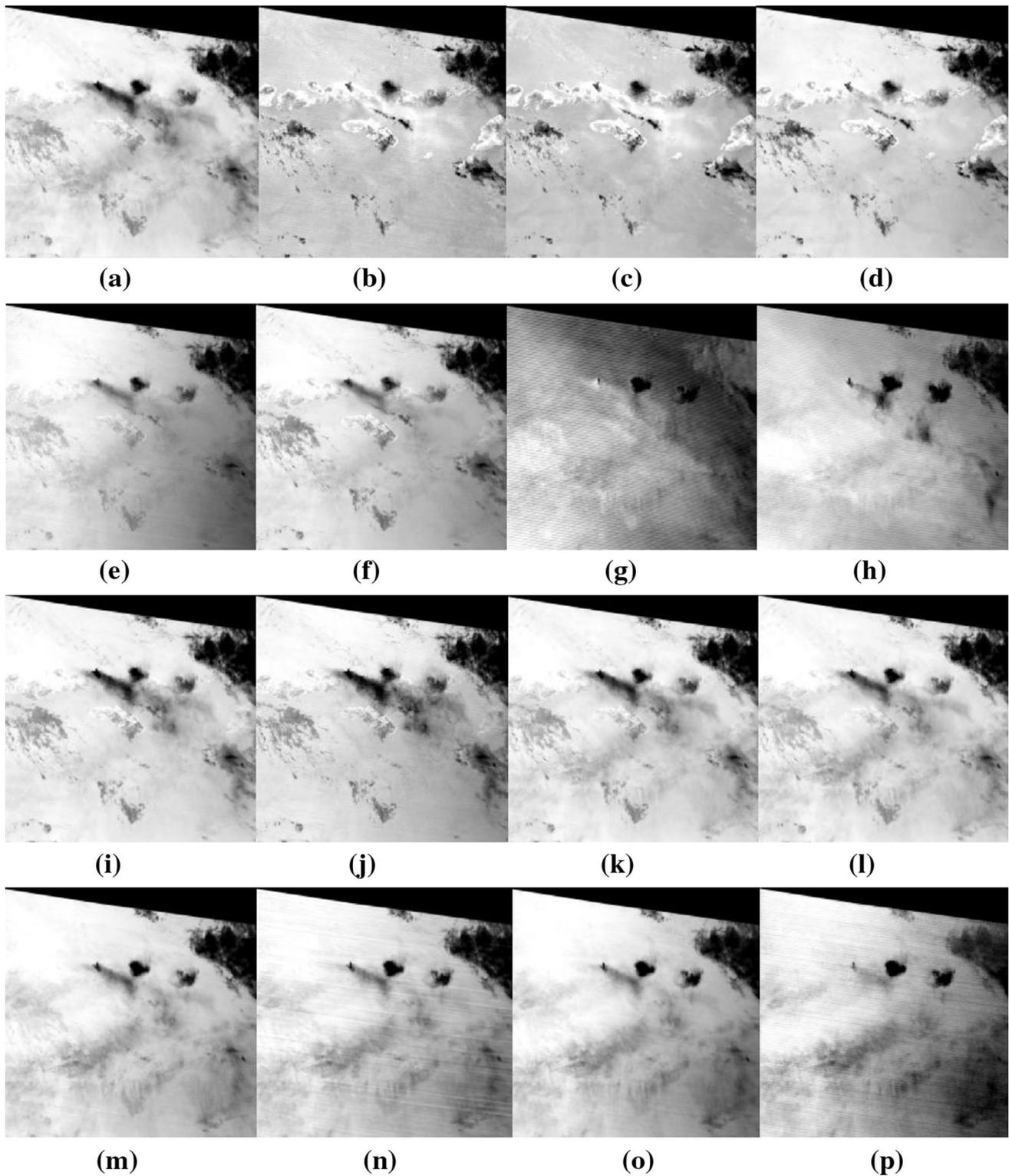


Fig. 4 MODIS original spectrum images of Sangeang Api; **a–p** are the MODIS B20–B36 (except of B26), respectively

MODIS CIs of Sangeang Api volcanic ash cloud is almost exactly the same and the color histogram distribution of MODIS CIs is very dense, which

means that the amount of information of different earth objects in MODIS CIs is relatively concentrated and has a more independent distribution.

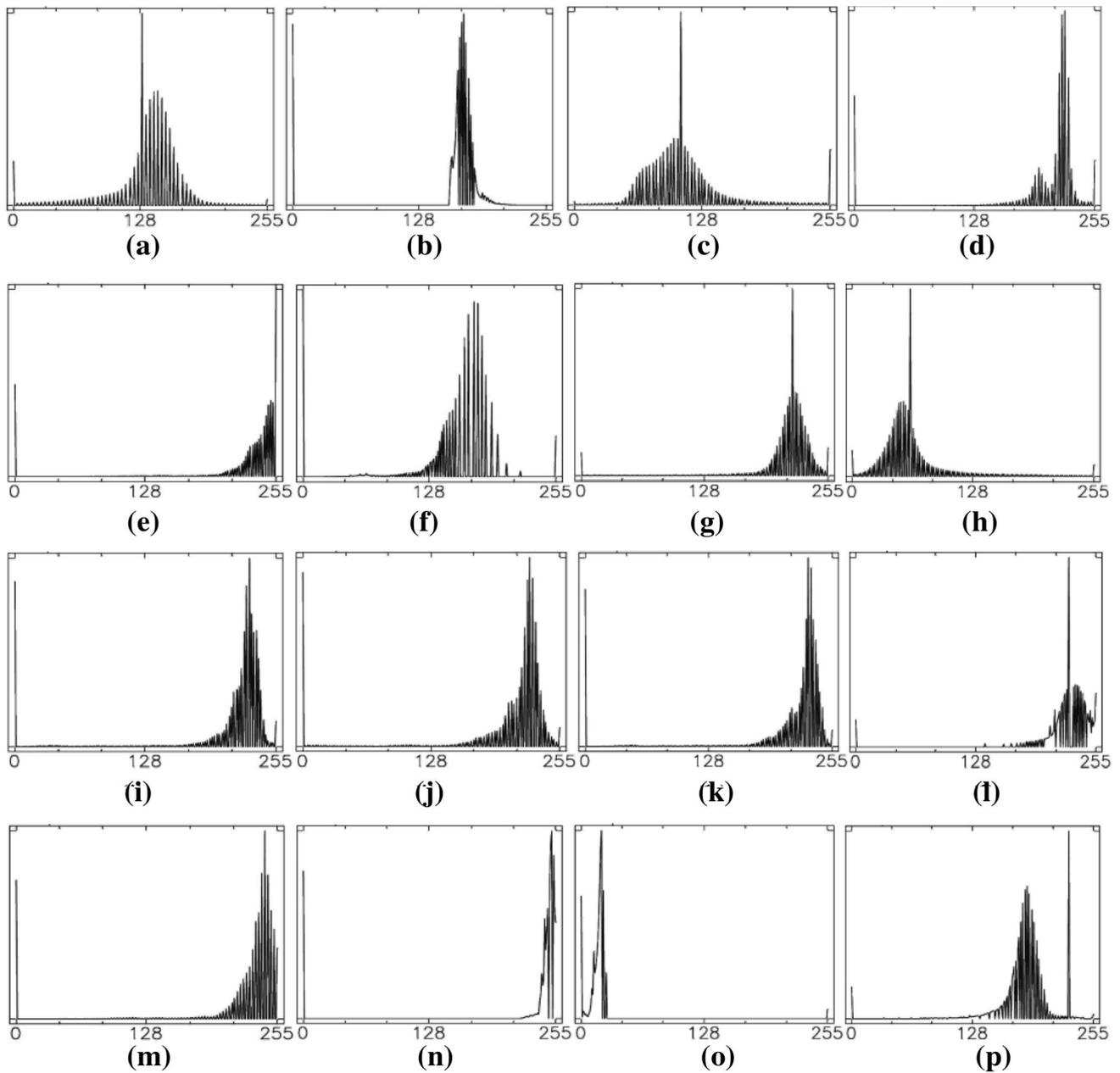


Fig. 5 MODIS CIs' histogram of Sangeang Api based on weighted PCA and ICA; **a–p** are the histograms of MODIS C11–C16, respectively

2. Compared with the original false color images' histogram, the color histogram of MODIS CIs has obvious peak values, which shows that the earth object distributions are more homogeneous and can generate more distinct pixel values. In the volcanic ash cloud detection from MODIS image, this means that the contrasts between volcanic ash cloud feature information and other features could be enlarged and that the spectrum images' complexity of MODIS original false color image can be reduced considerably.

Then it is necessary to select and ascertain the preferences of the types and parameters of kernel function. In our study, the RBF is selected as the kernel function of SVM classifier, and the corresponding penalty coefficient (C) and interval coefficient (γ) are selected as the optimum parameters of RBF kernel function. Finally, $C = 800$ and $\gamma = 3$ are used to detect the volcanic ash cloud information using the massive experiment research. The volcanic ash cloud of Sangeang Api from MODIS image by CPIWS method is illustrated in Fig. 7.

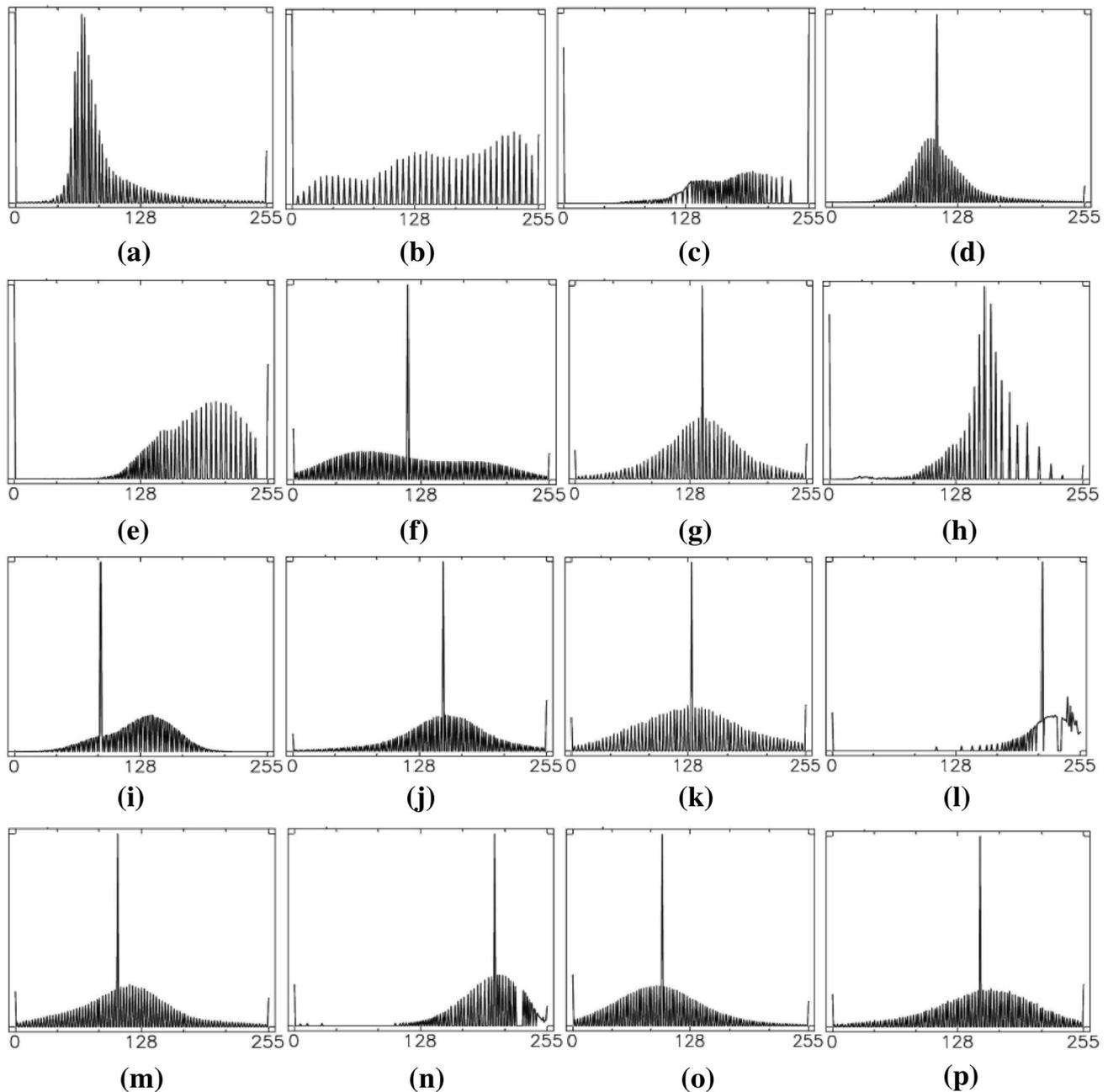


Fig. 6 MODIS original spectrum image histogram of Sangeang Api; **a–p** are the histograms of MODIS B20–B36 (except of B26), respectively

From Fig. 7, it can be seen that the CPIWS method could accurately detect the Sangeang Api volcanic ash cloud on 30 May 2014 from MODIS image. The distribution of volcanic ash cloud was uniform in horizontal direction; there are only a few broken spots in the detected volcanic ash cloud, and the detected results are very consistent with Fig. 2. It has good image quality and visual effect.

Experimental results and accuracy comparison

In order to assess the volcanic ash cloud detection with CPIWS method, in the experiment, 500 data points were randomly generated first by program and classified into the volcanic ash cloud and non-volcanic ash cloud in terms on the visual interpretation. Next, the confusion matrix, Kappa coefficient and overall accuracy were introduced and used

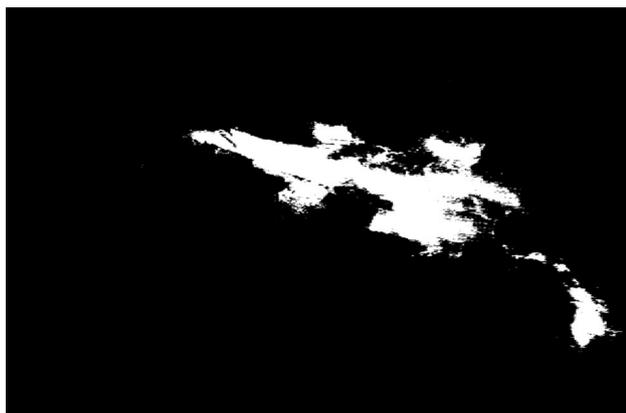


Fig. 7 Sangeang Api volcanic ash cloud on 30 May 2014 detected using CPIWS method; the white part is the volcanic ash cloud information

for accuracy assessment. And then the accuracy assessment of the detected volcanic ash cloud was performed. The confusion matrix and detection accuracy are shown in Tables 3 and 4.

The Kappa coefficient represents the relationship between the membership degree and the detection results. The larger the value of Kappa coefficient is, the higher the membership degree is, and the better the detection effect of thematic information from remote sensing images is.

As it is, the volcanic ash cloud detection is one of the most important applications of volcanic ash cloud monitoring. In order to assess the effectiveness of proposed CPIWS method, the split-window temperature difference (SWTD) method and RGB pseudo-color (RGB PC) method in traditional volcanic ash cloud monitoring were introduced to detect the Sangeang Api volcanic ash cloud on 30 May 2014 from the same MODIS image in this study. The processes of detection and assessment are same as in Sect. “Volcanic ash cloud classification” and “Experimental results and accuracy comparison”. The detected

volcanic ash cloud and accuracy comparison are illustrated in Fig. 8 and Table 5. In our study, all of the software packages used for simulation, like SWTD and RGB PC, include ENVI 4.6 and MATLAB 7.0.

From Fig. 8a and Table 5, it can also be seen that the CPIWS method could detect accurately the Sangeang Api volcanic ash cloud on 30 May 2014 from MODIS multi-spectral remote sensing image; the overall accuracy and Kappa coefficient reach 87.20 and 0.7958%, respectively. There is an obvious error detection of volcanic ash cloud. For example, a large high meteorological cloud is misclassified as the volcanic ash cloud by SWTD method. To some extent, this also further corroborates the previous research (Ellrod 2004; Ellrod and Schreiner 2004; Pavolonis et al. 2013; Zhao et al. 2014) showing that there are obvious error detections when SWTD method is used to monitor the high and cold clouds and the earth surface covered by snow and ice.

From Fig. 8b and Table 5, it can be seen that there are also obvious error detections of volcanic ash cloud by RGB PC method, and the overall accuracy and Kappa coefficient reaches 69.83 and 0.6258%, respectively. The meteorological clouds around the volcanic ash cloud were error detected while detecting the main section of volcanic ash cloud. What is more, the detected volcanic ash cloud usually exists in the form of broken plaques and noise information. And a large part of the errors of RGB PC method is caused by the low genting temperature of meteorological cloud.

Compared with the classic method in the field of volcanic ash cloud traditional monitoring, i.e., SWTD and RGB PC method, the overall accuracy and Kappa coefficient of CPIWS method reaches 87.20 and 0.7958%, respectively. From Table 4, the accuracy comparison of different method shows that the CPIWS yields better results than the traditional SWTD and RGB PC method. Hence, compared with the SWTD and RGB PC methods,

Table 3 Confusion matrix

| | Volcanic ash cloud | Non-volcanic ash cloud | Total |
|------------------------|--------------------|------------------------|-------|
| Volcanic ash cloud | 121 | 35 | 156 |
| Non-volcanic ash cloud | 29 | 315 | 344 |
| Total | 150 | 350 | 500 |

Table 4 Detection accuracy

| | Reference pixels number | Detection pixels number | Right detection pixels number | Producer accuracy, % | User accuracy, % |
|---------------------------|-------------------------|-------------------------|-------------------------------|----------------------|----------------------------|
| Volcanic ash cloud | 150 | 156 | 121 | 80.67 | 77.56 |
| Non-volcanic ash cloud | 350 | 344 | 315 | 90.00 | 91.57 |
| Total | 500 | 500 | 436 | – | – |
| Overall accuracy = 87.20% | | | | | |
| | | | | | Kappa coefficient = 0.7958 |

Fig. 8 Sangeang Api volcanic ash cloud on 30 May 2014: **a** SWT method, **b** RGB PC method; the *black part* is the volcanic ash cloud information, and the *white part* indicates seawater, meteorological cloud, and land surface

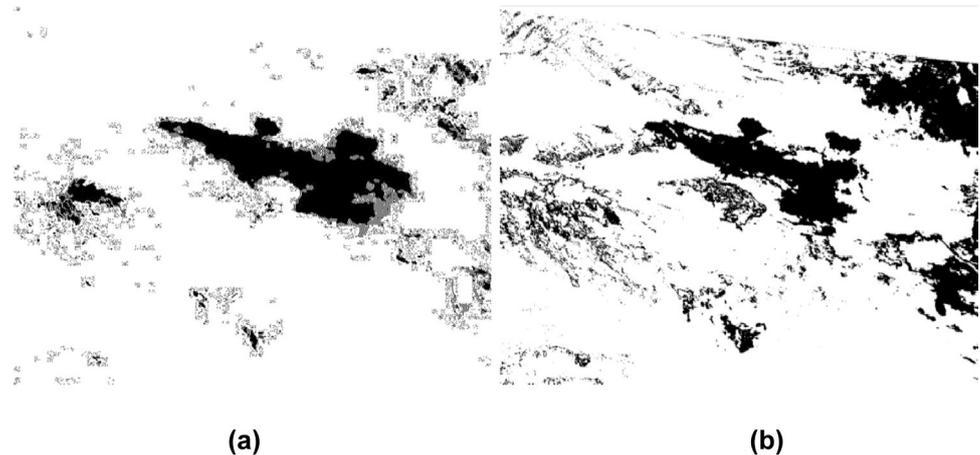


Table 5 Accuracy comparison of different methods

| | Kappa coefficient | Overall accuracy, % |
|--------|-------------------|---------------------|
| CPIWS | 0.7958 | 87.20 |
| SWTD | 0.6648 | 72.53 |
| RGB PC | 0.6258 | 69.83 |

CPIWS method has both the better detection accuracy and the better image quality.

Conclusions and discussions

This work presents a new method for detecting the volcanic ash cloud from remote sensing image based on CPIWS. In this method, the PCA, ICA and SVM techniques were applied to detect the volcanic ash cloud from MODIS image. Future experiments will consider the weighted PCA and ICA and the real case of Sangeang Api volcanic ash cloud on 30 May 2014. Compared to some previous works, for instance, those using only either PCA or ICA method (ex.: Hyvarinen 1999; Hillger and Clark 2002a; Li et al. 2014, 2015), and the traditional remote sensing monitoring methods of volcanic ash cloud (ex.: Duda and Canty 2002; Ellrod 2004; Ellrod and Schreiner 2004; Zhao et al. 2014), the proposed method is much simpler and easier to calculate in a practical application.

From our study, we have concluded that

- From the point of view of using single PCA or ICA method, ICA-SVM method is better than PCA method. From the point of view of using single ICA or PCA-ICA-weighted method, the latter is better than the former. The classification accuracy of MODIS image by CPIWS method can achieve the best effect in suitable conditions.

- From the point of view of the detection accuracy, although the traditional SWT and RGB PC volcanic ash cloud monitoring methods have fine detection results, the volcanic ash cloud detected using MODIS image was worse, with the overall accuracy of 72.53 and 69.83% in Sangeang Api volcanic ash cloud case. Compared to the SWT and RGB PC methods, the total detection accuracy of CPIWS method and Kappa coefficient in this case reaches 87.20 and 0.7958, respectively, which is obviously higher than that derived using the other two methods. According to the above-mentioned analysis, this is mainly caused by the various stages of volcanic ash cloud diffusion. For example, the SWT method is usually used in the early stage of volcanic ash cloud diffusion, while the RGB PC method is usually used in the late stage of volcanic ash cloud diffusion.
- From the point of view of the visual effect, the CPIWS method has the best visual effect and image quality, while the second-best one is the SWT method, and the one with the worst visual effect is the RGB PC method. As the CPIWS method is capable of making full use of the spectrum characteristics of volcanic ash cloud in MODIS image, it can especially effectively distinguish volcanic ash cloud from meteorological cloud.

In summary, the CPIWS method can not only improve the PCA or ICA methods in volcanic ash cloud detection in the condition of selecting suitable weighted value, but also overcome the limitation of traditional remote sensing monitoring method for different volcanic ash cloud diffusion processes. In the field of volcanic ash cloud monitoring, the detection and the tracing techniques of volcanic ash cloud are both bottlenecks limiting various methods to attain the practical request. The key to volcanic ash cloud tracing lies in effectively detecting volcanic ash cloud from

remote sensing data. As a whole, there is a lot of work to do in the future. For instance,

1. From the point of view of the algorithm itself, the proposed method also has disadvantages, i.e., reasonable SVM kernel function and weighted value.
2. From the point of view of the applicability, this study is just aiming at the MODIS multispectral remote sensing image; whether it is suitable for other types of remote sensing images (i.e., hyperspectral remote sensing image and high-spatial resolution remote sensing image) and volcanic ash cloud cases (i.e., volcanos near the equator, mid-latitudes and high-latitudes) is still unknown.
3. This work mainly focuses on the detection of volcanic ash cloud from the single remote sensing image, and not tracing of volcanic ash cloud. Current volcanic ash cloud tracing, i.e., diffusion and prediction, is usually based on the mathematical models, and it does not involve the use of remote sensing technology to accurately detect volcanic ash cloud and further verify the diffusion and prediction. Furthermore, the tracing of volcanic ash cloud is one of the follow-up research concerns and needs more investigations to be made.

Acknowledgements We thank the the anonymous reviewers and editorial team for their vital comments and suggestions made. This work was supported by the Projects of National Science Foundation of China (41404024), Shanghai Science and Technology Development Foundation (16142203000) and Young Teachers Training and Supporting Plan in Shanghai Universities (2014-2016).

References

- Centre LVAA (2014), London volcanic ash advisory centre (VAAC), <http://www.metoffice.gov.uk/aviation/vaac/>. Accessed 14 Aug 2014
- Chang CI, Chiang SS, Smith JA, Ginsberg IW (2002) Linear spectral random mixture analysis for hyperspectral imagery. *IEEE Trans Geosci Remote* 40(2):375–392. doi:10.1109/WHISPERS.2009.5289096
- Christopher SA, Feng N, Naeger A, Johnson B, Marengo F (2012) Satellite remote sensing analysis of the 2010 Eyjafjallajökull volcanic ash cloud over the North Sea during 4–18 May 2010. *J Biol Chem* 117(D20):401–409. doi:10.1029/2011JD016850
- Corradini S, Spinetti C, Carboni E (2008) Mt. Etna tropospheric ash retrieval and sensitivity analysis using moderate resolution imaging spectroradiometer measurements. *J Appl Remote Sens* 2(1):23550–23570. doi:10.1117/1.3046674
- Corradini S, Merucci L, Folch A (2011) Volcanic ash cloud properties: comparison between MODIS satellite retrievals and FALL3D transport model. *IEEE Geosci Remote Sens Lett* 8(2):248–252. doi:10.1109/LGRS.2010.2064156
- Duda T, Canty M (2002) Unsupervised classification of satellite imagery: choosing a good algorithm. *Int J Remote Sens* 23(11):2193–2212. doi:10.1080/01430060110078467
- Ellrod GP (2004) Impact on volcanic ash detection caused by the loss of the 12.0 μm “Split Window” band on GOES imagers. *J Volcanol Geotherm Res* 135(1–2):91–103. doi:10.1016/j.jvolgeores.2003.12.009
- Ellrod GP, Schreiner AJ (2004) Volcanic ash detection and cloud top height estimates from the GOES-12 imager: coping without a 12 μm infrared band. *Geophys Res Lett* 31(15):1–4. doi:10.1029/2004GL020395
- Heblinski J, Schmieder K, Heege T, Agyemang TK, Sayadyan H, Vardanyan L (2011) High-resolution satellite remote sensing of littoral vegetation of Lake Sevan (Armenia) as a basis for monitoring and assessment. *Hydrobiologia* 61(1):97–111. doi:10.1007/s10750-010-0466-6
- Hillger DW, Clark J (2002a) Principal component image analysis of MODIS for volcanic ash. Part I: most important bands and implications for future GOES imagers. *J Appl Meteorol* 41(1):985–1001. doi:10.1175/1520-0450(2002)041<0985:PCIAOM>2.0.CO;2
- Hillger DW, Clark J (2002b) Principal component image analysis of MODIS for volcanic ash. Part II: simulation of current GOES and GOES-M imagers. *J Appl Meteorol* 41(10):1003–1010. doi:10.1175/1520-0450(2002)041<1003:PCIAOM>2.0.CO;2
- Hyvarinen A (1999) Survey on independent component analysis. *Neural Comput Surv* 2(2):94–128
- Hyvarinen A, Oja E (2000) Independent component analysis: algorithms and applications. *Neural Netw* 13(4–5):411–430. doi:10.1016/S0893-6080(00)00026-5
- Lee TW, Girolami M, Sejnowski TJ (1999) Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural Comput* 11(2):417–441. doi:10.1162/089976699300016719
- Li CF, Yin JY (2013) Variational Bayesian independent component analysis-support vector machine for remote sensing classification. *Comput Electr Eng* 39(3):717–726. doi:10.1016/j.compeleceng.2012.10.004
- Li CF, Dai YY, Zhao JY, Yin Y, Zhou SQ (2014) Remote sensing detection of volcanic ash cloud using independent component analysis. *Seismol Geol* 36(1):137–142. doi:10.3969/j.issn.0253-4967.2014.01.011
- Li CF, Dai YY, Zhao JJ, Zhou SQ, Yin JY, Xue D (2015) Remote sensing monitoring of volcanic ash clouds based on PCA method. *Acta Geophys* 63(2):432–450. doi:10.2478/s11600-014-0257-y
- Liang L, Yang MH, Li YF (2010) Hyperspectral remote sensing image classification based on ICA and SVM algorithm. *Spectrosc Spectr Anal* 30(10):2724–2728. doi:10.3964/j.issn.1000-0593(2010)10-2724-05
- Mackie S, Millington S, Watson IM (2014) How assumed composition affects the interpretation of satellite observations of volcanic ash. *Meteorol Appl* 21(1):20–29. doi:10.1002/met.1445
- Montopoli M, Cimini D, Lamantea M, Herzog M, Graf HF, Marzano FS (2013) Microwave radiometric remote sensing of volcanic ash clouds from space: model and data analysis. *IEEE Trans Geosci Remote* 51(9):4678–4691. doi:10.1109/TGRS.2013.2260343
- Mountrakis G, Irn J, Ogole C (2011) Support vector machines in remote sensing: a review. *ISPRS J Photogramm* 66(3):247–249. doi:10.1016/j.isprsjprs.2010.11.001
- Pavlonis MJ, Heidinger AK, Sieglaff J (2013) Automated retrievals of volcanic ash and dust cloud properties from upwelling infrared measurements. *J Geophys Res* 118(3):1436–1458. doi:10.1002/JGRD.50173
- Prata AJ, Prata AT (2012) Eyjafjallajökull volcanic ash concentrations determined using spin enhanced visible and infrared imager measurements. *J Geophys Res* 117(D20):2156–2202. doi:10.1029/2011JD016800
- Sahin M, Yildiz BY, Senkal O, Pestemalci V (2012) Modeling and remote sensing of land surface temperature in Turkey. *J Indian Soc Remote* 40(3):399–409. doi:10.1007/s12524-011-0158-3

- Sanchez-Azofeifa A, Rivard B, Wright J, Feng JL, Li PJ, Chong MM, Bohlman SA (2011) Estimation of the distribution of *Tabebuia guayacana* (Bignoniaceae) using high-resolution remote sensing imagery. *Sensors* 11(4):3831–3851. doi:[10.3390/s110403831](https://doi.org/10.3390/s110403831)
- Schumann U, Weinzierl B, Reitebuch O, Schlager H, Minikin A, Forster C, Baumann R, Sailer T, Graf K, Mannstein H, Voigt C, Rahm S, Simmet R, Scheibe M, Lichtenstern M, Stock P, Rüba H, Schäuble D, Tafferner A, Rautenhaus M, Gerz T, Ziereis H, Krautstrunk M, Mallaun C, Gayet J-F, Lieke K, Kandler K, Ebert M, Weinbruch S, Stohl A, Gasteiger J, Groß S, Freudenthaler V, Wiegner M, Ansmann A, Tesche M, Olafsson H, Sturm K (2011) Airborne observations of the Eyjafjalla volcano ash cloud over Europe during air space closure in April and May 2010. *Atmos Chem Phys* 11(5):2245–2279. doi:[10.5194/acp-11-2245-2011](https://doi.org/10.5194/acp-11-2245-2011)
- Segl K, Roessner S, Heiden U (2003) Fusion of spectral and shape features for identification of urban surface cover types using reflective and thermal hyperspectral data. *ISPRS J Photogramm* 58(1–2):99–112. doi:[10.1016/S0924-2716\(03\)00020-0](https://doi.org/10.1016/S0924-2716(03)00020-0)
- Spinetti C, Barsotti S, Neri A, Buongiorno MF, Doumaz F, Nannipieri L (2013) Investigation of the complex dynamics and structure of the 2010 Eyjafjallajökull volcanic ash cloud using multispectral images and numerical simulations. *J Geophys Res* 118(10):4729–4747. doi:[10.1002/jgrd.50328](https://doi.org/10.1002/jgrd.50328)
- Wang XM, Zeng SG, Xia DS (2006) Remote sensing image classification based on a loose modified fast ICA algorithm. *J. Comput Res Dev* 43(4):708–715 (in Chinese)
- Wang ZQ, Liu XQ, Zhang GL, Wang ZJ (2007) Face recognition based on PCA and ICA. *J Huazhong Nor Univ* 41(3):373–376 (in Chinese)
- Ward CA, Starks SA (1999) An approach to predict Africanized honey bee migration using remote sensing. *Comput Electr Eng* 26(1):33–45. doi:[10.1016/S0045-7906\(99\)00028-2](https://doi.org/10.1016/S0045-7906(99)00028-2)
- Wen XP, Yang XF, Hu GD (2011) Relationship between land cover ration and urban heat island from remote sensing and automatic weather stations data. *J Ind Soc Remote* 39(2):193–201. doi:[10.1007/s12524-011-0076-4](https://doi.org/10.1007/s12524-011-0076-4)
- Western LM, Watson MI, Francis PN (2015) Uncertainty in two-channel infrared remote sensing retrievals of a well-characterised volcanic ash cloud. *Bull Volcanol* 77(8):1–12. doi:[10.1007/s00445-015-0950-y](https://doi.org/10.1007/s00445-015-0950-y)
- Xu YM, Qin ZH, Wan HX (2010) Spatial and temporal dynamics of urban heat island and their relationship with land cover changes in urbanization process: a case study in Suzhou, China. *J Ind Soc Remote* 38(4):654–663. doi:[10.1007/s12524-011-0073-7](https://doi.org/10.1007/s12524-011-0073-7)
- Yi WB, Tang H, Chen YH (2011) An object-oriented semantic clustering algorithm for high-resolution remote sensing images using the aspect model. *IEEE Geosci Remote Sens Lett* 8(3):522–526. doi:[10.1109/LGRS.2010.2090034](https://doi.org/10.1109/LGRS.2010.2090034)
- Zhao Y, Liang Y, Ma BJ, Li YS, Wu XJ (2014) Identification of Icelandic volcanic ash cloud based on FY-3A remote sensing data. *Acta Petrol Sin* 30(12):3693–3700 (in Chinese)

Forced oscillation measurements of seismic attenuation in fluid saturated sandstone

Shankar Subramaniyan¹ · Beatriz Quintal² · Erik H. Saenger³

Received: 23 January 2017 / Accepted: 27 January 2017 / Published online: 10 February 2017
© Institute of Geophysics, Polish Academy of Sciences & Polish Academy of Sciences 2017

Abstract Adopting the method of forced oscillation, attenuation was studied in Fontainebleau sandstone (porosity 10%, permeability 10 mD) at seismic frequencies (1–100 Hz). Confining pressures of 5, 10, and 15 MPa were chosen to simulate reservoir conditions. First, the strain effect on attenuation was investigated in the dry sample for 11 different strains across the range 1×10^{-6} – 8×10^{-6} , at the confining pressure of 5 MPa. The comparison showed that a strain of at least 5×10^{-6} is necessary to obtain a good signal to noise ratio. These results also indicate that nonlinear effects are absent for strains up to 8×10^{-6} . For all the confining pressures, attenuation in the dry rock was low, while partial (90%) and full (100%) saturation with water yielded a higher magnitude and frequency dependence of attenuation. The observed high and frequency dependent attenuation was interpreted as being caused by squirt flow.

Keywords Forced oscillation · Sandstone · Strain · Attenuation

Introduction

As it passes through the subsurface, the amplitude of a seismic wave is reduced in magnitude on account of elastic losses, such as scattering and mode conversion at interfaces, as well as anelastic absorption, where energy is lost due to its conversion into heat. Anelastic absorption, which is the focus of this study, is here simply referred to as attenuation. It represents the fraction of energy dissipated per cycle in a seismic wave ($1/Q$, O’Connell and Budiansky 1978). In contrast to dry rocks that exhibit low and frequency independent attenuation, fluid saturated rocks have been observed to exhibit high and frequency dependent attenuation (Mikhailsevitch et al. 2014; Pimienta et al. 2015; Subramaniyan et al. 2015). Hence, attenuation could be an indicator of pore fluid content in a reservoir if its magnitude as well as frequency dependence on rock physical properties is understood well (Chapman et al. 2006; Lambert et al. 2013). As attenuation is sensitive to the presence of fluids, it can potentially be employed in techniques such as time lapse monitoring of reservoirs.

In the laboratory, attenuation has been studied mainly using three techniques, i.e., ultrasonic (Toksöz et al. 1979; Winkler 1983), resonant bar (Gardner et al. 1964; Winkler and Nur 1979) and forced oscillation (Batzle et al. 2006; Behura et al. 2007). Frequencies in ultrasonic and resonant bar methods lie in the kHz and MHz range and correspond approximately to those used in well logging techniques. The technique of forced oscillation involves the simple principle of calculating attenuation from the phase shift of sinusoidal stress and strain signals, but the lack of sophisticated equipment to achieve that end has kept this study from moving

✉ Shankar Subramaniyan
shankar.subramaniyan@erdw.ethz.ch

Beatriz Quintal
beatriz.quintal@unil.ch

Erik H. Saenger
erik.saenger@hs-bochum.de

¹ Department of Earth Sciences, ETH Zurich, Zurich, Switzerland

² Institute of Earth Sciences, University of Lausanne, Lausanne, Switzerland

³ International Geothermal Centre and Ruhr Universität Bochum, Bochum, Germany

forward rapidly (Subramaniyan et al. 2014). The forced oscillation technique has the advantage that the frequencies measured (1–100 Hz) correspond to those of seismic datasets acquired on the earth's surface for oil and gas exploration.

The general aim of studying attenuation in the laboratory concerns with understanding the cause of frequency dependent attenuation in fluid saturated rocks (Adam et al. 2009; Tisato and Quintal 2013; Subramaniyan et al. 2015; Chapman et al. 2016). Once the cause is identified and understood, physical laws relating attenuation to properties of the rock and fluids can be employed to convert the attenuation attribute obtained from field data into information regarding the rock-fluid system present in subsurface. Wave induced fluid flow (Pride et al. 2004) is considered to be the dominant factor causing frequency dependent attenuation in fluid saturated rocks. Fluid flow is driven by pressure gradients developed on account of varying local compliances in the rock. These compliances can occur at various scales, corresponding to the different scales of heterogeneities in the rock frame such as fractures, cracks, voids or due to the combination of immiscible pore fluids in liquid and gaseous states. Müller et al. (2010) provided a summary of the mechanisms that exist at all scales. More laboratory data is required to narrow down mechanisms responsible for the predominant attenuation observed under conditions that are characteristic of deep reservoir rocks.

Sandstones have been extensively used in laboratory studies as their pore structure is simpler, when compared to carbonates (Zimmerman 1991). Attenuation related experiments have shown that dry sandstones exhibit low and frequency independent attenuation while fluid saturated sandstones can exhibit higher magnitude and frequency dependent attenuation (Usher 1962; Winkler 1985; Tisato and Quintal 2013, 2014; Subramaniyan et al. 2015; Pimienta et al. 2015; Chapman et al. 2016). We aim to deepen this understanding by studying attenuation and the corresponding modulus dispersion in Fontainebleau sandstone, which is mainly composed of quartz grains and its porosity increases linearly as the amount of cementation is reduced (Bourbié and Zinszner 1985). Once the controlling mechanisms causing attenuation in sandstones are understood, similar experiments can be carried out in more complicated rocks such as carbonates.

A series of laboratory experiments were conducted in Fontainebleau sandstone, using the Seismic Wave Attenuation Module (SWAM, Madonna and Tisato 2013) at seismic frequencies (1–100 Hz). The SWAM works on the principle of forced oscillation. To simulate reservoir conditions, confining pressures of 5, 10, and 15 MPa were adopted. A strain study was performed by measuring

attenuation in the dry rock across the strain range 1×10^{-6} – 8×10^{-6} , at the confining pressure of 5 MPa. For studying the effect of pore fluids, the sample was saturated, first partially, then fully, with water.

Methods

The SWAM is comprised of a piezo actuator to generate a vertical sinusoidal stress in the sample as well as in an aluminum standard. The voltage fed to the piezo actuator corresponds to the desired strain in the sample. The displacements of both the standard and the sample are measured by linear variable differential transformers (LVDTs), recorded on an oscilloscope after amplification. The module was calibrated using plexiglass and aluminum (Madonna and Tisato 2013). Attenuation ($1/Q_E$) and Young's modulus (E) of the sample were calculated from phase and amplitude of displacement signals acquired for the sample and the standard (Subramaniyan et al. 2015).

The sandstone sample used for the experiments was cylindrical and measured 60 mm in length and 25.4 mm in diameter. The properties of the sandstone are as follows: permeability of 10 mD, porosity of 10%, and dry rock bulk density of 2386 kg/m³. The SWAM is placed in a Paterson pressure rig (Paterson and Olgaard 2000) that uses argon gas as the confining medium. Prior to enclosing the sample in a shrink tube, an aluminum foil was glued around the curved surface of the sample to avoid fluid flow out of the sample during experiments and prevent argon from seeping into the sample. The aluminum standard was also enclosed in a shrink tube. The SWAM was connected to a syringe pump using pipes of 1 mm diameter to saturate the sample. As the experiment progressed, the shrink tube allowed for a small leak of argon gas into the sample (0.01–0.05 MPa/h, as per the pump reading). To allow argon gas to escape through the fluid pipes, valves on either side of the sample were kept open during the measurements and hence in all measurements, the pore pressure was equal to atmospheric pressure. This resulted in a possibility of fluid flow out of the sample where it was connected to the pore fluid pipes.

After the completion of measurements on dry sample, it was partially saturated with water, by allowing the corresponding volume of water into the sample. When partial saturation attenuation measurements were complete, the sample was then fully saturated with water. Under a constant pore pressure (~ 2 MPa), the sample was flushed up to ~ 10 – 15 times its pore volume to ensure full saturation with water. The sample was oven dried before the experiments. All experiments were

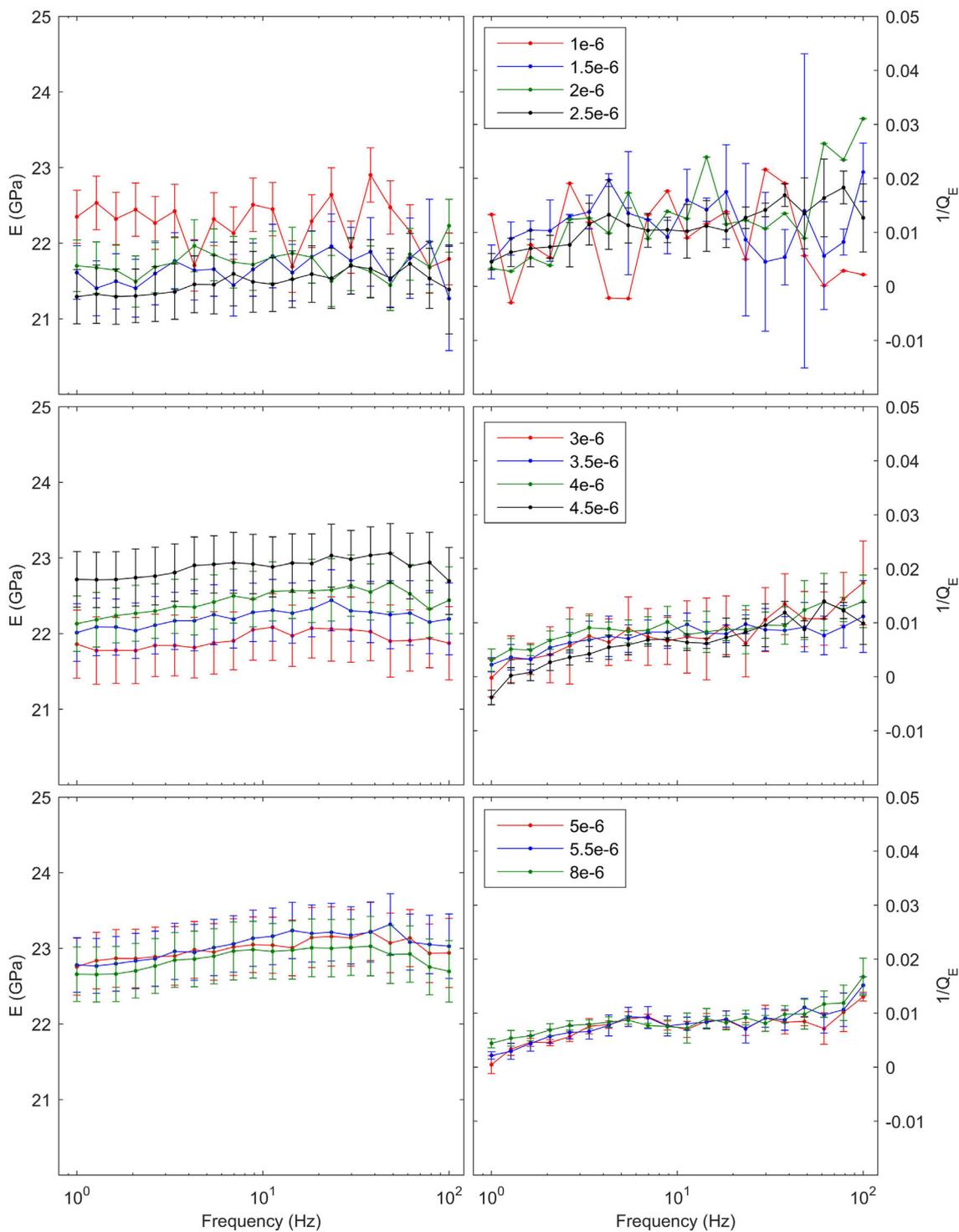


Fig. 1 Laboratory measurements of Young’s modulus (E) and attenuation ($1/Q_E$) for dry Fontainebleau sandstone at increasing sample strains from 1×10^{-6} to 8×10^{-6} , at the confining pressure of 5 MPa

performed at 20 °C. After performing the investigation of the effect of sample strain on attenuation, all subsequent measurements were conducted at a sample strain of 5×10^{-6} . The datasets represent attenuation/

dispersion acquired at 20 discrete frequencies, plotted on a log scale, while the error bars include repetition of measurements done by looping five times over the frequency range.

Results and discussion

Effect of strain

Figure 1 displays attenuation and dispersion data obtained for the dry sandstone, at a confining pressure of 5 MPa, for a series of sample strains from 1×10^{-6} – 8×10^{-6} . In the first strain range 1×10^{-6} – 2.5×10^{-6} (top plot), the attenuation signals are barely discernible with enormous error bars, while the Young's modulus values show an erratic trend. When the strain is increased to the range of 3×10^{-6} – 4.5×10^{-6} (middle plot), the attenuation signals are better, quite similar in magnitude, but still with large error bars. Additionally, the Young's modulus varies significantly for this strain range. However, when the strain is increased to the range of 5×10^{-6} – 8×10^{-6} (bottom plot), attenuation data overlaps well, exhibiting considerably smaller error bars and the Young's modulus values converge. These results indicate that a strain of at least 5×10^{-6} should be applied to the sample to provide a good quality signal. The results also indicate that a strain as high as 8×10^{-6} can be imposed on the sample without resulting in nonlinear effects, i.e., an increase in attenuation magnitude with an increase in strain.

Winkler et al. (1979) observed that attenuation remains constant with strain amplitude up to a certain cut-off value, after which it begins to increase with an increase in strain, resulting in nonlinearity. They also observed that this cut-off value shifts to higher strains as the confining pressure is increased, i.e., nonlinear effects begin to operate at higher strains when the confining pressure (effective stress) is increased. This implies that, for our study, it is sufficient to obtain an optimal strain value at 5 MPa and the measurements will lie in the linear range also for 10 and 15 MPa.

It is important to avoid strain related nonlinear effects so that attenuation measurements can be compared to intrinsic attenuation observed in seismic field data. In surface reflection seismic, strains are typically expected to be of the order of 10^{-6} (Karato and Spetzler 1990). So, the strain magnitude can be high enough to ensure high signal quality (small error bars), but it should be low enough to exclude nonlinear effects.

Figure 2 displays the displacement signals of the dry sample for selected strain cases for the frequencies of 1, 11, and 100 Hz, at the confining pressure of 5 MPa. Even though 18–20 periods were recorded for each frequency, only ~ 4 periods are shown in Fig. 2. For all displayed frequencies, it is visible that the quality of data improves as the strain is increased, correlating well with the observations of attenuation/dispersion in Fig. 1. Meanwhile, the signals are much smoother at 1 Hz than at 100 Hz due to the constant sampling rate, resulting in a worsening sampling with increasing frequency.

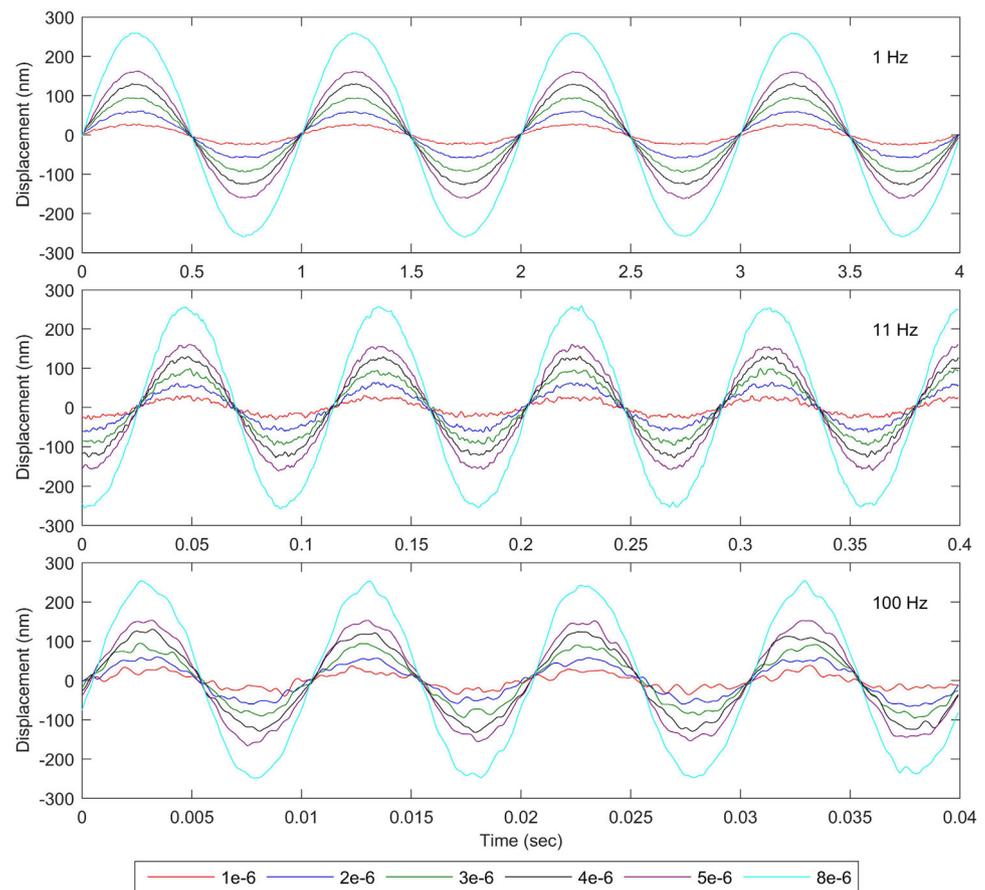
Effect of fluid saturation

Figure 3 displays attenuation/dispersion data acquired for the Fontainebleau sandstone, for the cases of dry rock and partial (50, 90%) as well as full (100%) saturation with water. At all the three confining pressures, 5, 10, and 15 MPa, the dry Fontainebleau exhibits low attenuation. For 50% water saturation, attenuation increases at 5 and 10 MPa, with no change in magnitude at 15 MPa. For 90% water saturation, attenuation increases further at all confining pressures. When the sandstone is fully (100%) saturated, results are similar to 90% water saturation at 5 and 15 MPa, but we observe remarkable attenuation increase and frequency dependence at 10 MPa. In general, attenuation is approximately frequency independent at 5 MPa while it exhibits frequency dependence at higher confining pressures for both dry and saturated cases. For Young's modulus, we observe an increase in its value with increasing saturation, leading to values in the range ~ 26 – 30 GPa, compared to ~ 22 – 23 GPa for the dry case.

Overall, a gradual increase in the magnitude of attenuation can be observed as the amount of water is increased in the rock sample (Fig. 3). This could be explained as an increase in local fluid flow within the sample as air is substituted with water. Replacing air with water also leads to stiffening of the rock, observed as the increasing values of Young's modulus.

Concerning the fluid flow mechanism responsible for the observed frequency dependent attenuation, the microscopic squirt flow (Mavko and Nur 1979) is most likely the predominant mechanism. We should not, however, ignore the fact that the attenuation for the water saturated cases are probably also influenced by some fluid flow occurring at the (mesoscopic) sample scale, from the sample into the fluid pipes (Gardner et al. 1964; Dunn 1987). To some extent, our results are in conformity with those obtained by Subramaniyan et al. (2015), where a similar low porosity Fontainebleau sandstone was saturated with water (1 cP) and with high viscosity fluids (8–1414 cP), resulting in high and frequency dependent bell-shaped attenuation curves in the seismic frequency range (1–100 Hz) for viscosities above 8 cP. The data for full saturation with glycerin (1414 cP) was clearly suggestive of squirt flow as the predominant mechanism, as indicated by the reasonable agreement with a squirt flow analytical solution (Gurevich et al. 2010). The attenuation measurements of Subramaniyan et al. (2015) for full water saturation, however, exhibited a very modest increase in magnitude and frequency dependence, compared to attenuation in the dry sample. This behaviour is somewhat different from that of our data for confining pressures of 10 and 15 MPa (Fig. 3), where the increase in attenuation is considerably more pronounced. This difference could be explained by a lower

Fig. 2 Displacement signals of dry Fontainebleau sandstone for varying sample strains in the range 1×10^{-6} – 8×10^{-6} , for the frequencies of 1, 11, and 100 Hz, at the confining pressure of 5 MPa



aspect ratio of micro-cracks in the sample used in this study, which would cause the attenuation to peak at a lower frequency (e.g., Mavko et al. 2009; Gurevich et al. 2010) than that for the sample used by Subramaniyan et al. (2015).

In light of an interpretation of our data (Fig. 3) based on the squirt flow mechanism, the fact that the attenuation behaviour for full water saturation significantly changes with confining pressure (5–10 MPa) can probably be explained by a change in the aspect ratio of some of the randomly distributed micro-cracks with changing confining pressure. In addition, closing completely a portion of the micro-cracks with an increase in confining pressure can cause a reduction of the attenuation magnitude and frequency dependence, as observed when comparing the data for 10 and 15 MPa.

Data consistency

As discussed and suggested by several authors (O'Donnell et al. 1981; Spencer 1981; Mikhaltsevitch et al. 2014; Chapman et al. 2016), an important technique to validate consistency of attenuation and dispersion datasets is to verify if they satisfy Kramers–Kronig relations, thereby confirming

their inherent dependence on each other as they arise from the same physical cause. Equation 1 describes the relationship between the spatial attenuation co-efficient (α) and the longitudinal phase velocity (C_E) (O'Donnell et al. 1981):

$$\alpha = \frac{\pi\omega^2}{2C_E^2} \frac{dC_E}{d\omega} \quad (1)$$

where ω is the angular frequency, which relates to the linear frequency f through $\omega = 2\pi f$. For low loss materials ($Q_E \gg 1$), the attenuation measured in the laboratory ($1/Q_E$) is related to the attenuation coefficient (α) as (Carcione 2007):

$$Q_E^{-1} = \frac{2\alpha C_E}{\omega}. \quad (2)$$

Equation 3 relates the longitudinal phase velocity (C_E) to Young's modulus (E) and density of the medium (ρ) (Mikhaltsevitch et al. 2014):

$$C_E = \sqrt{\frac{E}{\rho}}. \quad (3)$$

Combining Eqs. 1–3, the Kramers–Kronig relations between attenuation ($1/Q_E$) and Young's modulus (E) can be obtained as (Mikhalesevitch et al. 2014):

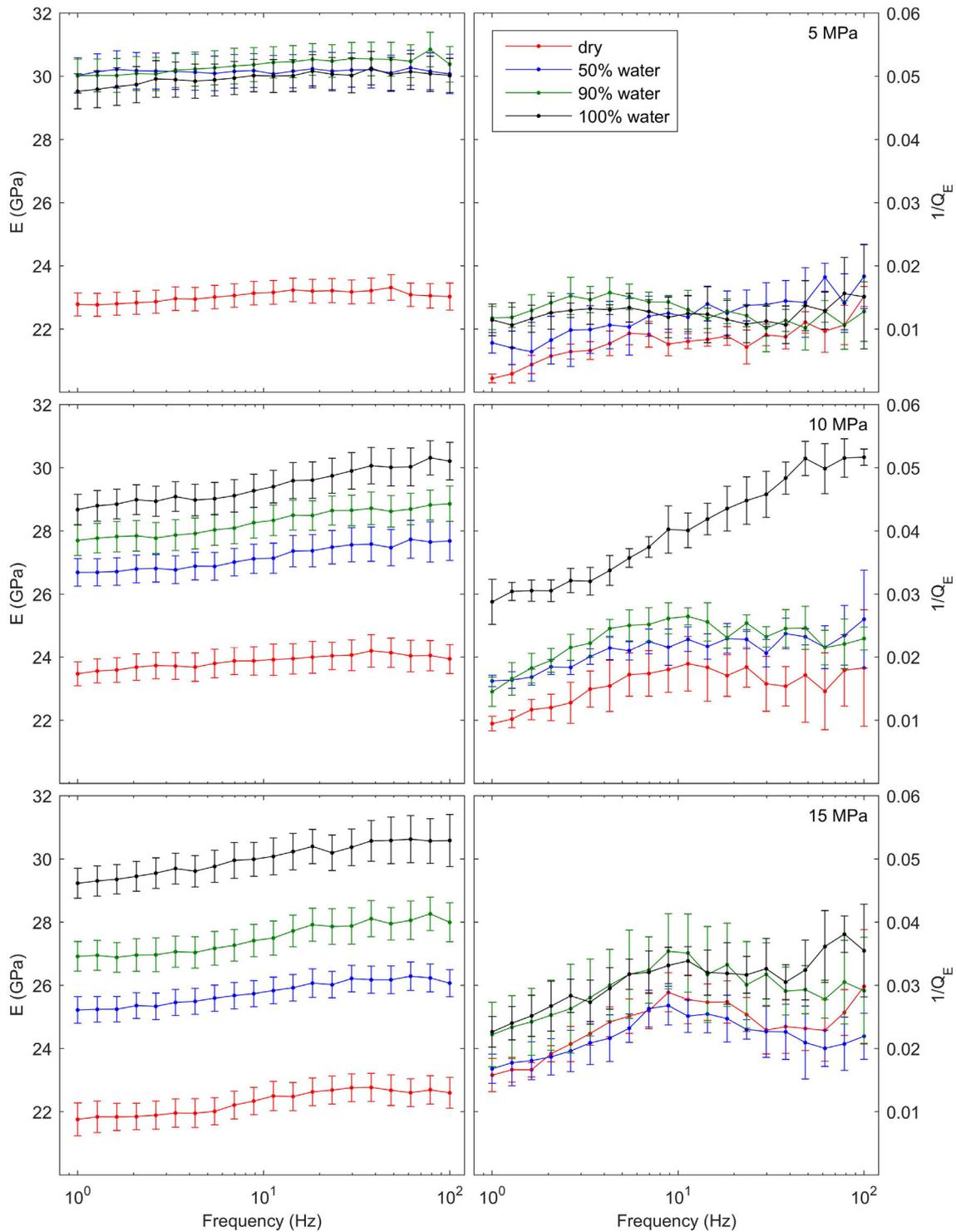


Fig. 3 Laboratory measurements of Young’s modulus (E) and attenuation ($1/Q_E$) for dry Fontainebleau sandstone as well as the cases of partial (50, 90%) and full (100%) saturation with water, at the confining pressures of 5, 10, and 15 MPa

$$E_f = E_0 e^b, \tag{4}$$

$$b = \frac{2}{\pi} \int_{f_0}^f \frac{Q_E^{-1}}{f'} df',$$

Equation 4 shows that the theoretical Young’s modulus (E_f) at frequency (f) can be derived by numerically integrating attenuation from the reference frequency f_0 (in our case, the lowest one, i.e., 1 Hz) up to the frequency (f) at

which the Young's modulus is being calculated. For the measurements shown in Fig. 3, Young's modulus values were derived from values of attenuation using the Kramers–Kronig relations (Eq. 4).

Figure 4 thus shows the resulting Young's modulus values at the confining pressures of 5, 10, and 15 MPa. We can observe that these predictions fit relatively well with the Young's modulus values measured in the laboratory for the cases of dry and partial (50, 90%) water saturation. At the confining pressures of 5 and 15 MPa, the case of full (100%) water saturation also shows a reasonable fit. However, for the case of full (100%) saturation with water at 10 MPa, Kramers–Kronig relations predict a higher dispersion of Young's modulus, which does not fit well the laboratory data acquired at frequencies close to 100 Hz, when attenuation is quite high. This, together with the fact that the laboratory measurements for Young's modulus tend to show a slightly negative dispersion for the dry

sample at frequencies close to 100 Hz, indicate that the Young's modulus measurements at those frequencies are not very accurate or reliable. We expect the attenuation data at those frequencies to be more reliable.

Conclusions

At the confining pressure of 5 MPa, a strain study performed on dry Fontainebleau sandstone revealed that a minimum sample strain of 5×10^{-6} is needed to obtain good signal quality. It is essential that while performing laboratory based attenuation experiments, an optimum strain value is chosen, or the quality of data may be too poor for the assessment of underlying mechanisms. Our strain study also showed that we operate within the linear regime. At confining pressures of 5, 10, and 15 MPa, the dry sandstone exhibited low attenuation, while its magnitude increased upon partial saturation with water. Full

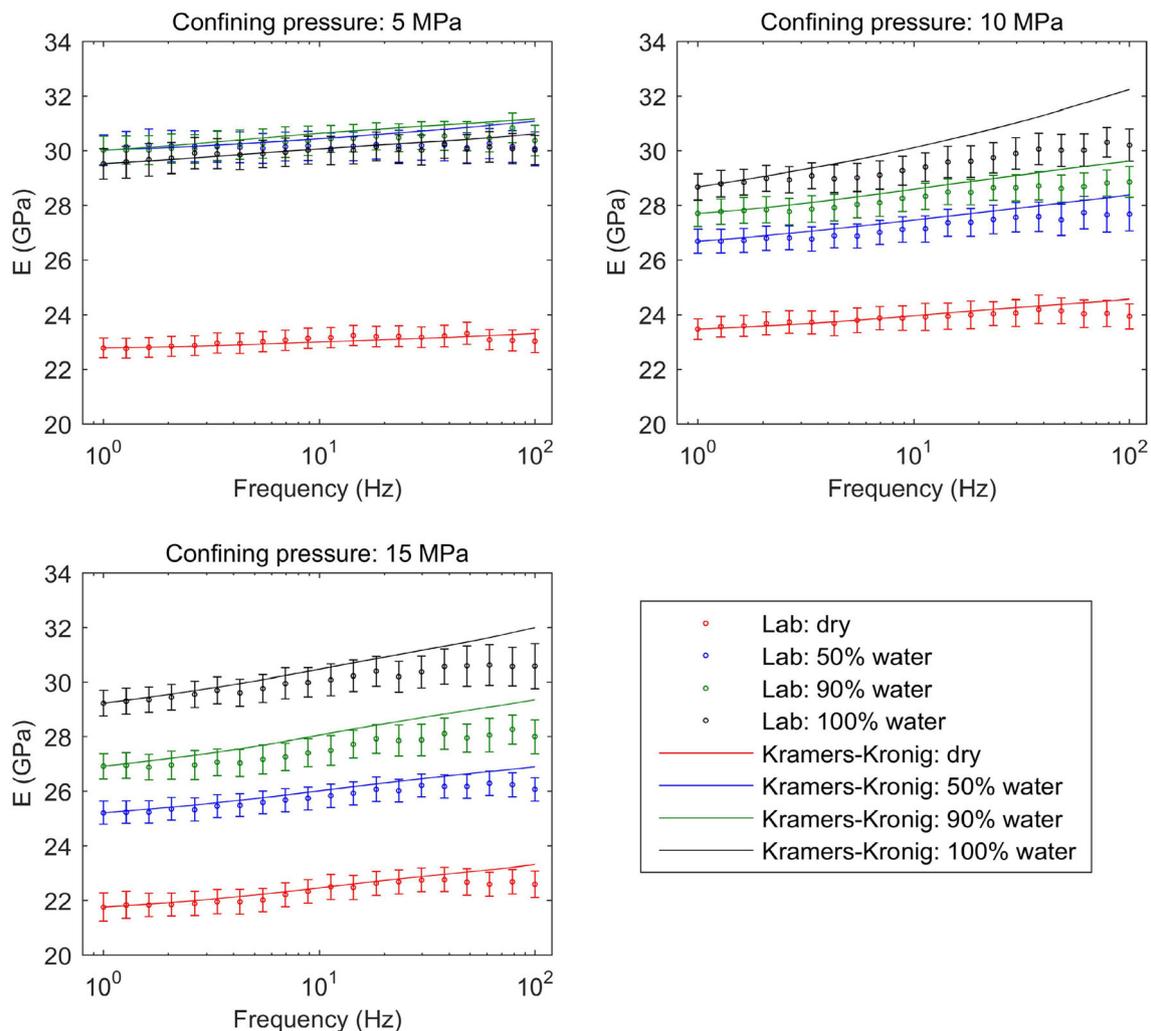


Fig. 4 Young's modulus derived from Kramers–Kronig relations using $1/Q_E$ as input (solid lines), compared to Young's modulus obtained from laboratory measurements (circles and error bars), at the confining pressures of 5, 10, and 15 MPa

saturation with water increased attenuation further, particularly at the confining pressures of 10 and 15 MPa. The observed high and frequency dependent attenuation was interpreted as being caused by squirt flow. The obtained results highlight the importance of the role played by fluid flow mechanisms and the necessity of further laboratory and theoretical investigations to identify the mechanisms controlling fluid related attenuation effects.

Acknowledgements Petrobras co-funded this project. We would like to thank Claudio Madonna for his support in the laboratory and Prof. Jean-Pierre Burg for providing a conducive research environment.

References

- Adam L, Batzle M, Lewallen KT, van Wijk K (2009) Seismic wave attenuation in carbonates. *J Geophys Res* 114(B6):B06208. doi:[10.1029/2008JB005890](https://doi.org/10.1029/2008JB005890)
- Batzle ML, Han D-H, Hofmann R (2006) Fluid mobility and frequency-dependent seismic velocity—direct measurements. *Geophysics* 71(1):N1–N9. doi:[10.1190/1.2159053](https://doi.org/10.1190/1.2159053)
- Behura J, Batzle M, Hofmann R, Dorgan J (2007) Heavy oils: their shear story. *Geophysics* 72(5):E175–E183. doi:[10.1190/1.2756600](https://doi.org/10.1190/1.2756600)
- Bourbie T, Zinszner B (1985) Hydraulic and acoustic properties as a function of porosity in Fontainebleau sandstone. *J Geophys Res* 90(B13):11524–11532. doi:[10.1029/JB090iB13p11524](https://doi.org/10.1029/JB090iB13p11524)
- Carcione JM (2007) Wave fields in real media: theory and numerical simulation of wave propagation in anisotropic. Anelastic porous and electromagnetic media. Elsevier, Amsterdam
- Chapman M, Liu E, Li X (2006) The influence of fluid-sensitive dispersion and attenuation on AVO analysis. *Geophys J Int* 167(1):89–105. doi:[10.1111/j.1365-246X.2006.02919.x](https://doi.org/10.1111/j.1365-246X.2006.02919.x)
- Chapman S, Tisato N, Quintal B, Holliger K (2016) Seismic attenuation in partially saturated Berea sandstone submitted to a range of confining pressures. *J Geophys Res Solid Earth* 121(3):1664–1676. doi:[10.1002/2015JB012575](https://doi.org/10.1002/2015JB012575)
- Dunn K-J (1987) Sample boundary effect in acoustic attenuation of fluid-saturated porous cylinders. *J Acoust Soc Am* 81(5):1259–1266. doi:[10.1121/1.394529](https://doi.org/10.1121/1.394529)
- Gardner GHF, Wyllie MRJ, Droschak DM (1964) Effects of pressure and fluid saturation on the attenuation of elastic waves in sands. *J Petrol Technol* 16(2):189–198. doi:[10.2118/721-PA](https://doi.org/10.2118/721-PA)
- Gurevich B, Makarynska D, de Paula OB, Pervukhina M (2010) A simple model for squirt-flow dispersion and attenuation in fluid-saturated granular rocks. *Geophysics* 75(6):N109–N120. doi:[10.1190/1.3509782](https://doi.org/10.1190/1.3509782)
- Karato S, Spetzler HA (1990) Defect microdynamics in minerals and solid-state mechanisms of seismic wave attenuation and velocity dispersion in the mantle. *Rev Geophys* 28(4):399–421. doi:[10.1029/RG028i004p00399](https://doi.org/10.1029/RG028i004p00399)
- Lambert M-A, Saenger EH, Quintal B, Schmalholz SM (2013) Numerical simulation of ambient seismic wavefield modification caused by pore-fluid effects in an oil reservoir. *Geophysics* 78(1):T41–T52. doi:[10.1190/GEO2011-0513.1](https://doi.org/10.1190/GEO2011-0513.1)
- Madonna C, Tisato N (2013) A new seismic wave attenuation module to experimentally measure low-frequency attenuation in extensional mode. *Geophys Prospect* 61(2):302–314. doi:[10.1111/1365-2478.12015](https://doi.org/10.1111/1365-2478.12015)
- Mavko GM, Nur A (1979) Wave attenuation in partially saturated rocks. *Geophysics* 44(2):161–178. doi:[10.1190/1.1440958](https://doi.org/10.1190/1.1440958)
- Mavko G, Mukerji T, Dvorkin J (2009) The rock physics handbook. Tool for seismic analysis of porous media. Cambridge University Press, Cambridge
- Mikhailsevitch V, Lebedev M, Gurevich B (2014) A laboratory study of low-frequency wave dispersion and attenuation in water-saturated sandstones. *Leading Edge* 33(6):616–618, 620–622. doi:[10.1190/le33060616.1](https://doi.org/10.1190/le33060616.1)
- Müller TM, Gurevich B, Lebedev M (2010) Seismic wave attenuation and dispersion resulting from wave-induced flow in porous rocks—A review. *Geophysics* 75(5):75A147–75A164. doi:[10.1190/1.3463417](https://doi.org/10.1190/1.3463417)
- O’Connell RJ, Budiansky B (1978) Measures of dissipation in viscoelastic media. *Geophys Res Lett* 5(1):5–8. doi:[10.1029/GL005i001p00005](https://doi.org/10.1029/GL005i001p00005)
- O’Donnell M, Jaynes ET, Miller JG (1981) Kramers–Kronig relationship between ultrasonic attenuation and phase velocity. *J Acoust Soc Am* 69(3):696–701. doi:[10.1121/1.385566](https://doi.org/10.1121/1.385566)
- Paterson MS, Olgaard DL (2000) Rock deformation tests to large shear strains in torsion. *J Struct Geol* 22(9):1341–1358. doi:[10.1016/S0191-8141\(00\)00042-0](https://doi.org/10.1016/S0191-8141(00)00042-0)
- Pimienta L, Fortin J, Guéguen Y (2015) Bulk modulus dispersion and attenuation in sandstones. *Geophysics* 80(2):D111–D127. doi:[10.1190/geo2014-0335.1](https://doi.org/10.1190/geo2014-0335.1)
- Pride SR, Berryman JG, Harris JM (2004) Seismic attenuation due to wave-induced flow. *J Geophys Res* 109(B1):B01201. doi:[10.1029/2003JB002639](https://doi.org/10.1029/2003JB002639)
- Spencer JW Jr (1981) Stress relaxations at low frequencies in fluid-saturated rocks: Attenuation and modulus dispersion. *J Geophys Res* 86(B3):1803–1812. doi:[10.1029/JB086iB03p01803](https://doi.org/10.1029/JB086iB03p01803)
- Subramaniyan S, Quintal B, Tisato N, Saenger EH, Madonna C (2014) An overview of laboratory apparatuses to measure seismic attenuation in reservoir rocks. *Geophys Prospect* 62(6):1211–1223. doi:[10.1111/1365-2478.12171](https://doi.org/10.1111/1365-2478.12171)
- Subramaniyan S, Quintal B, Madonna C, Saenger EH (2015) Laboratory-based seismic attenuation in Fontainebleau sandstone: Evidence of squirt flow. *J Geophys Res Solid Earth* 120(11):7526–7535. doi:[10.1002/2015JB012290](https://doi.org/10.1002/2015JB012290)
- Tisato N, Quintal B (2013) Measurements of seismic attenuation and transient fluid pressure in partially saturated Berea sandstone: evidence of fluid flow on the mesoscopic scale. *Geophys J Int* 195(1):342–351. doi:[10.1093/gji/ggt259](https://doi.org/10.1093/gji/ggt259)
- Tisato N, Quintal B (2014) Laboratory measurements of seismic attenuation in sandstone: strain versus fluid saturation effects. *Geophysics* 79(5):WB9–WB14. doi:[10.1190/geo2013-0419.1](https://doi.org/10.1190/geo2013-0419.1)
- Toksöz MN, Johnston DH, Timur A (1979) Attenuation of seismic waves in dry and saturated rocks: I. Laboratory measurements. *Geophysics* 44(4):681–690. doi:[10.1190/1.1440969](https://doi.org/10.1190/1.1440969)
- Usher MJ (1962) Elastic behavior of rocks at lower frequencies. *Geophys Prospect* 10(2):119–127. doi:[10.1111/j.1365-2478.1962.tb02002.x](https://doi.org/10.1111/j.1365-2478.1962.tb02002.x)
- Winkler KW (1983) Frequency dependent ultrasonic properties of high-porosity sandstones. *J Geophys Res* 88(B11):9493–9499. doi:[10.1029/JB088iB11p09493](https://doi.org/10.1029/JB088iB11p09493)
- Winkler KW (1985) Dispersion analysis of velocity and attenuation in Berea sandstone. *J Geophys Res* 90(B8):6793–6800. doi:[10.1029/JB090iB08p06793](https://doi.org/10.1029/JB090iB08p06793)
- Winkler KW, Nur A (1979) Pore fluids and seismic attenuation in rocks. *Geophys Res Lett* 6(1):1–4. doi:[10.1029/GL006i001p00001](https://doi.org/10.1029/GL006i001p00001)
- Winkler KW, Nur A, Gladwin M (1979) Friction and seismic attenuation in rocks. *Nature* 277(5697):528–531. doi:[10.1038/277528a0](https://doi.org/10.1038/277528a0)
- Zimmerman RW (1991) Compressibility of sandstones, developments in petroleum science, vol 29. Elsevier Science Publishing Co., Inc., New York

Automatic scaling of critical frequency f_oF2 from ionograms recorded at São José dos Campos, Brazil: a comparison between Autoscala and UDIDA tools

Michael Pezzopane¹  · Valdir Gil Pillat² · Paulo Roberto Fagundes²

Received: 23 January 2017 / Accepted: 27 January 2017 / Published online: 8 February 2017
© Institute of Geophysics, Polish Academy of Sciences & Polish Academy of Sciences 2017

Abstract This paper considers a dataset of ionograms recorded by the CADI ionosonde installed at São José dos Campos (SJC; 23.2°S, 45.9°W, magnetic latitude 14.1°S), Brazil, to compare two autoscaling systems: Autoscala, developed by the Istituto Nazionale di Geofisica e Vulcanologia, and the UDIDA-scaling, developed by the Universidade do Vale do Paraíba. The analysis, focused on the critical frequency of the F2 layer, f_oF2 , shows that the two systems work differently. The UDIDA-scaling gives always a value of f_oF2 as output, regardless of the presence of the ionogram trace and its definition, while Autoscala tends to reject ionograms for which the digital information is considered insufficient. As a consequence, the UDIDA-scaling can autoscale more f_oF2 values than Autoscala, but Autoscala can discard a larger number of ionograms for which the trace is unidentifiable. Discussions are made on the accuracy of the f_oF2 values given as output, as well as on the main shortcomings characterizing the two systems.

Keywords Ionogram · Ionosonde · Low-latitude ionosphere · Critical frequency f_oF2 · Automatic scaling

Introduction

The importance of recording real-time vertical sounding data, especially for space weather purposes, has greatly increased over the past years. In particular, the assimilation of this kind of data into ionospheric models has become more and more important (Pezzopane et al. 2011, 2013; Galkin 2012; Lee et al. 2012; McNamara et al. 2013; Gardner et al. 2014; Chartier et al. 2016; Sabbagh et al. 2016). For this reason, in the last decades, several groups of researchers worked hard to implement systems able to automatically scale the experimental trace obtained after performing an ionospheric vertical sounding, that is, an ionogram (Reinisch and Huang 1983; Fox and Blundell 1989; Igi et al. 1993; Tsai and Berkey 2000; Zabortin et al. 2006; Ding et al. 2007; Galkin and Reinisch 2008; Su et al. 2012; Chen et al. 2013; Zheng et al. 2013; Jiang et al. 2013, 2014, 2015a, b). In particular, the ARTIST system developed by the University of Lowell, Center for Atmospheric Research, has been widely used and tested (Gilbert and Smith 1988; Conkright and McNamara 1997; Jacobs et al. 2004; Reinisch et al. 2005; McNamara 2006; Bamford et al. 2008; Stankov et al. 2012). Although much progress has been made, there are many issues that have still to be faced, and a significant work is being continuously made to upgrade existing programs. The aim of these improvements is, on the one hand, to advance the reliability of the automatically scaled data, and on the other hand to adapt some systems for being installed in ionosondes that are not equipped with a tool to automatically scale the recorded trace.

One type of ionosonde that is built without being equipped with an automatic scaling system is the Canadian Advanced Digital Ionosonde (CADI) (MacDougall 1993). The present paper considers a dataset of ionograms

✉ Michael Pezzopane
michael.pezzopane@ingv.it

Valdir Gil Pillat
valdirgp@univap.br

Paulo Roberto Fagundes
fagundes@univap.br

¹ Istituto Nazionale di Geofisica e Vulcanologia, Rome, Italy
² Laboratório de Física e Astronomia, Universidade do Vale do Paraíba (UNIVAP), São José dos Campos, São Paulo, Brazil

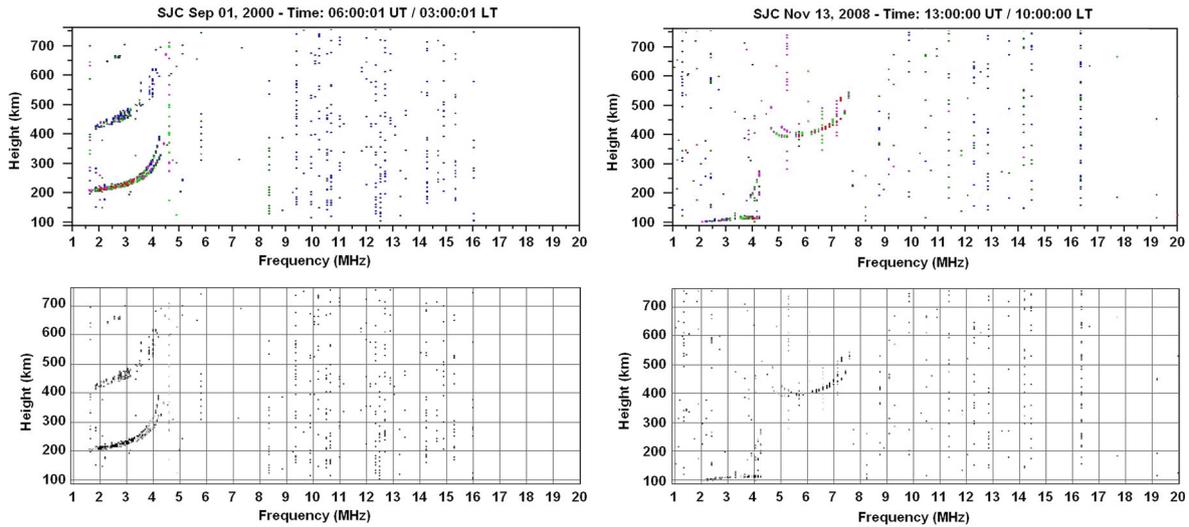


Fig. 1 Examples of conversion from (top) the MD4 format to (bottom) the RDF format for ionograms recorded at SJC on (left) 1 September 2000 at 06:00 UT and (right) 13 November 2008 at 13:00 UT

Table 1 The autoscaling statistics (see the text for the description of the four different subsets), and the corresponding autoscaling accuracy for subset 1, are shown for both Autoscala and UDIDA for the hourly ionograms recorded in September 2000 ($R_{12} = 116$)

September 2000 ($R_{12} = 116$) (603 ionograms)

Autoscaling statistics

| | Autoscala | UDIDA |
|----------|------------|-------|
| Subset 1 | 137 (~37%) | 369 |
| Subset 2 | 229 (~98%) | – |
| Subset 3 | 5 (~2%) | 234 |
| Subset 4 | 232 (~63%) | – |

Autoscaling accuracy subset 1

| $x_a =$ lautoscaled value – validated value (MHz) | Autoscala (137 ionograms) | | | UDIDA (369 ionograms) | | |
|---|---------------------------|------------------------|--------------------------|-----------------------|-------------------------|---------------------------|
| | Total ionograms (137) | Daytime ionograms (59) | Nighttime ionograms (78) | Total ionograms (369) | Daytime ionograms (242) | Nighttime ionograms (127) |
| $x_a \leq 0.5$ | 127 (~93%) | 53 (~90%) | 74 (~95%) | 265 (~72%) | 165 (~68%) | 100 (~79%) |
| $0.5 < x_a \leq 1.0$ | 5 (~3%) | 2 (~3%) | 3 (~4%) | 52 (~14%) | 37 (~15%) | 15 (~12%) |
| $1.0 < x_a \leq 1.5$ | – | – | – | 35 (~10%) | 28 (~12%) | 7 (~5%) |
| $x_a > 1.5$ | 5 (~3%) | 4 (~7%) | 1 (~3%) | 17 (~4%) | 12 (~5%) | 5 (~4%) |

| $x_r =$ lautoscaled value – validated value/validated value | Autoscala (137 ionograms) | | | UDIDA (369 ionograms) | | |
|---|---------------------------|------------------------|--------------------------|-----------------------|-------------------------|---------------------------|
| | Total ionograms (137) | Daytime ionograms (59) | Nighttime ionograms (78) | Total ionograms (369) | Daytime ionograms (242) | Nighttime ionograms (127) |
| $x_r \leq 0.1$ | 130 (~95%) | 55 (~93%) | 75 (~96%) | 328 (~89%) | 221 (~91%) | 107 (~84%) |
| $0.1 < x_r \leq 0.3$ | 3 (~2%) | 1 (~2%) | 2 (~3%) | 35 (~9%) | 17 (~7%) | 18 (~14%) |
| $0.3 < x_r \leq 0.5$ | 3 (~2%) | 2 (~3%) | 1 (~1%) | 3 (~1%) | 1 (~1%) | 2 (~2%) |
| $x_r > 0.5$ | 1 (~1%) | 1 (~2%) | – | 3 (~1%) | 3 (~1%) | – |

Concerning the “Autoscaling statistics”: the percentages between brackets of Autoscala for subset 1 and subset 4 refer to the total number of ionograms of (subset 1 + subset 4), that is, the ionograms for which the operator gave a value of $foF2$ as output; the percentages between brackets of Autoscala for subset 2 and subset 3 refer to the total number of ionograms of (subset 2 + subset 3), that is, the ionograms for which the operator did not give any value of $foF2$ as output

Table 2 Same as Table 1 but for March 2004 ($R_{12} = 47$)

| March 2004 ($R_{12} = 47$) (627 ionograms) | | | | | | |
|---|---------------------------|-------------------------|---------------------------|-----------------------|-------------------------|---------------------------|
| Autoscaling statistics | | | | | | |
| | Autoscala | | | UDIDA | | |
| Subset 1 | 265 (~51%) | | | 522 | | |
| Subset 2 | 85 (~81%) | | | – | | |
| Subset 3 | 20 (~19%) | | | 105 | | |
| Subset 4 | 257 (~49%) | | | – | | |
| Autoscaling accuracy subset 1 | | | | | | |
| $x_a =$ lautoscaled value – validated value (MHz) | Autoscala (265 ionograms) | | | UDIDA (522 ionograms) | | |
| | Total ionograms (265) | Daytime ionograms (125) | Nighttime ionograms (140) | Total ionograms (522) | Daytime ionograms (316) | Nighttime ionograms (206) |
| $x_a \leq 0.5$ | 241 (~91%) | 110 (~88%) | 131 (~94%) | 278 (~53%) | 183 (~58%) | 95 (~46%) |
| $0.5 < x_a \leq 1.0$ | 20 (~7%) | 11 (~8%) | 9 (~6%) | 109 (~21%) | 62 (~20%) | 47 (~22%) |
| $1.0 < x_a \leq 1.5$ | 2 (~1%) | 2 (~2%) | – | 87 (~17%) | 55 (~17%) | 32 (~16%) |
| $x_a > 1.5$ | 2 (~1%) | 2 (~2%) | – | 48 (~8%) | 16 (~5%) | 32 (~16%) |
| $x_r =$ lautoscaled value – validated value/validated value | Autoscala (265 ionograms) | | | UDIDA (522 ionograms) | | |
| | Total ionograms (265) | Daytime ionograms (125) | Nighttime ionograms (140) | Total ionograms (522) | Daytime ionograms (316) | Nighttime ionograms (206) |
| $x_r \leq 0.1$ | 247 (~93%) | 117 (~94%) | 130 (~93%) | 379 (~73%) | 261 (~83%) | 118 (~57%) |
| $0.1 < x_r \leq 0.3$ | 18 (~7%) | 8 (~6%) | 10 (~7%) | 94 (~18%) | 43 (~13%) | 51 (~25%) |
| $0.3 < x_r \leq 0.5$ | – | – | – | 30 (~6%) | 8 (~3%) | 22 (~11%) |
| $x_r > 0.5$ | – | – | – | 19 (~3%) | 4 (~1%) | 15 (~7%) |

recorded by the CADI ionosonde installed at São José dos Campos (SJC) (23.2°S, 45.9°W, magnetic latitude 14.1°S), Brazil, and discusses the results obtained after applying two different autoscaling systems, Autoscala, developed at Istituto Nazionale di Geofisica e Vulcanologia, Rome, Italy (Pezzopane and Scotto 2005, 2007, 2010), and the UNIVAP Digital Ionosonde Data Analysis-scaling (UDIDA-scaling, hereafter UDIDA), developed at the Laboratório de Física e Astronomia of the Universidade do Vale do Paraíba (UNIVAP), São José dos Campos, Brazil (Pillat et al. 2013).

The main characteristic of Autoscala is that it is based on an image recognition technique, which makes the algorithm unrelated to the hardware features of the ionosonde, and consequently able to scale ionograms independently of the information on the polarization tagging. The algorithm gives as output the main characteristics of the F2, F1, E, and Es ionospheric layers, as well as an estimation of the vertical electron density profile (Scotto 2009), and it has been applied to different types of ionosonde (Pezzopane and Scotto 2004; Scotto and Pezzopane

2008; Bullett et al. 2010; Krasheninnikov et al. 2010; Pezzopane et al. 2010; Scotto and MacDougall 2012; Enell et al. 2016).

UDIDA uses a fuzzy relation to fit the trace and gives as output the following F region characteristics: the critical frequency of the F region (f_oF2), the virtual height of the base of the F region ($h'F$), and the F region peak height ($h_pF2 = h(f_oF2 \cdot 0.834)$) (Piggott and Rawer 1972).

In this paper, ionograms recorded at SJC for quiet and disturbed periods are autoscaled by both Autoscala and UDIDA. The dataset considered, how the data analysis has been carried out, and the corresponding results are described in the next section followed by which discussion is given. Conclusions are given in the final section.

Datasets, analyses and results

The ionograms considered to compare the two autoscaling systems, Autoscala and UDIDA, are divided into two groups. The first group includes the hourly ionograms recorded at SJC

Table 3 Same as Table 1 but for November 2008 ($R_{12} = 2$)

| November 2008 ($R_{12} = 2$) (510 ionograms) | | | | | | |
|---|---------------------------|------------------------|---------------------------|-----------------------|-------------------------|---------------------------|
| Autoscaling statistics | | | | | | |
| | Autoscala | | | UDIDA | | |
| Subset 1 | 228 (~53%) | | | 428 | | |
| Subset 2 | 61 (~74%) | | | – | | |
| Subset 3 | 21 (~26%) | | | 82 | | |
| Subset 4 | 200 (~47%) | | | – | | |
| Autoscaling accuracy subset 1 | | | | | | |
| $x_a =$ lautoscaled value – validated value (MHz) | Autoscala (228 ionograms) | | | UDIDA (428 ionograms) | | |
| | Total ionograms (228) | Daytime ionograms (93) | Nighttime ionograms (135) | Total ionograms (428) | Daytime ionograms (234) | Nighttime ionograms (194) |
| $x_a \leq 0.5$ | 224 (~98%) | 93 (100%) | 131 (~97%) | 270 (~63%) | 153 (~65%) | 117 (~60%) |
| $0.5 < x_a \leq 1.0$ | 4 (~2%) | – | 4 (~3%) | 47 (~11%) | 18 (~8%) | 29 (~15%) |
| $1.0 < x_a \leq 1.5$ | – | – | – | 31 (~7%) | 14 (~6%) | 17 (~9%) |
| $x_a > 1.5$ | – | – | – | 80 (~19%) | 49 (~21%) | 31 (~16%) |
| $x_r =$ lautoscaled value – validated value/validated value | Autoscala (228 ionograms) | | | UDIDA (428 ionograms) | | |
| | Total ionograms (228) | Daytime ionograms (93) | Nighttime ionograms (135) | Total ionograms (428) | Daytime ionograms (234) | Nighttime ionograms (194) |
| $x_r \leq 0.1$ | 219 (~96%) | 92 (~99%) | 127 (~94%) | 282 (~66%) | 168 (~72%) | 114 (~59%) |
| $0.1 < x_r \leq 0.3$ | 9 (~4%) | 1 (~1%) | 8 (~6%) | 73 (~17%) | 28 (~12%) | 45 (~23%) |
| $0.3 < x_r \leq 0.5$ | – | – | – | 26 (~6%) | 14 (~6%) | 12 (~6%) |
| $x_r > 0.5$ | – | – | – | 47 (~11%) | 24 (~10%) | 23 (~12%) |

for high solar activity ($R_{12} = 116$) in September 2000, for medium solar activity ($R_{12} = 47$) in March 2004, and for low solar activity ($R_{12} = 2$) in November 2008, and it represents the “quiet” dataset, even though a geomagnetic storm occurred during 17–20 September 2000 ($D_{st} = -201$ nT). The second group, which represents the “disturbed” dataset, includes the 5-min ionograms recorded at SJC during three different geomagnetic storms that occurred during 24–26 August 2005 ($D_{st} = -184$ nT), during 14–16 December 2006 ($D_{st} = -162$ nT), and during 24–26 October 2011 ($D_{st} = -147$ nT), and during the well-known and well-studied St. Patrick storm that occurred during 17–19 March 2015 ($D_{st} = -223$ nT) (Astafyeva et al. 2015; Carter et al. 2016; De Michelis et al. 2016; Nava et al. 2016; Nayak et al. 2016; Pignalberi et al. 2016; Spogli et al. 2016; Tulasi Ram et al. 2016; Zhong et al. 2016).

It is worth highlighting that to process CADI ionograms with Autoscala it was necessary to perform an ionogram file format conversion from the MD4 one, which is the native CADI format, to the RDF format (Pezzopane 2004), which is the format with which the Autoscala system was run since the early stage of its development. This format conversion did

not cause any loss of information, as it is shown in Fig. 1 for typical nighttime and daytime ionograms.

To evaluate the two autoscaling systems, automatically scaled values of f_oF2 are compared with the ones hand-scaled by an experienced operator. We chose to carry out the test by considering the ionospheric characteristic f_oF2 because for this one both systems obtained the most reliable results. Moreover, with reference to the processing dataset of ionograms, the following subsets were considered: subset 1, including ionograms for which both the autoscaling system and the operator give a value as output; subset 2, including ionograms for which neither the autoscaling system nor the operator give a value as output; subset 3, including ionograms for which the autoscaling system gives a value as output, while the operator does not; subset 4, including ionograms for which the operator gives a value as output, while the autoscaling system does not.

Tables 1, 2, 3, 4, 5, 6 and 7 show the results obtained after applying both Autoscala and UDIDA on each of the seven single periods (the three quiet months and the four disturbed periods) forming the whole considered dataset of ionograms recorded at SJC. For each table, the top part

Table 4 Same as Table 1 but for the 5-min ionograms recorded during 24–26 August 2005, a disturbed period characterized by a D_{st} of -184 nT

| Storm 24–26 August 2005 ($D_{st} = -184$ nT) (826 ionograms) | | | | | | |
|---|---------------------------|-------------------------|---------------------------|-----------------------|-------------------------|---------------------------|
| Autoscaling statistics | | | | | | |
| | Autoscala | | | UDIDA | | |
| Subset 1 | 507 (~80%) | | | 635 | | |
| Subset 2 | 122 (~64%) | | | – | | |
| Subset 3 | 69 (~36%) | | | 191 | | |
| Subset 4 | 128 (~20%) | | | – | | |
| Autoscaling accuracy subset 1 | | | | | | |
| $x_a =$ lautoscaled value – validated value (MHz) | Autoscala (507 ionograms) | | | UDIDA (635 ionograms) | | |
| | Total ionograms (507) | Daytime ionograms (254) | Nighttime ionograms (253) | Total ionograms (635) | Daytime ionograms (362) | Nighttime ionograms (273) |
| $x_a \leq 0.5$ | 493 (~97%) | 248 (~97%) | 245 (~97%) | 325 (~51%) | 212 (~59%) | 113 (~41%) |
| $0.5 < x_a \leq 1.0$ | 11 (~2%) | 4 (~2%) | 7 (~2%) | 124 (~20%) | 62 (~17%) | 62 (~23%) |
| $1.0 < x_a \leq 1.5$ | 1 (~0.5%) | – | 1 (~1%) | 78 (~12%) | 19 (~5%) | 59 (~22%) |
| $x_a > 1.5$ | 2 (~0.5%) | 2 (~1%) | – | 108 (~17%) | 69 (~19%) | 39 (~14%) |
| $x_r =$ lautoscaled value – validated value/validated value | Autoscala (507 ionograms) | | | UDIDA (635 ionograms) | | |
| | Total ionograms (507) | Daytime ionograms (254) | Nighttime ionograms (253) | Total ionograms (635) | Daytime ionograms (362) | Nighttime ionograms (273) |
| $x_r \leq 0.1$ | 488 (~96%) | 249 (~98%) | 239 (~94%) | 338 (~53%) | 239 (~66%) | 99 (~36%) |
| $0.1 < x_r \leq 0.3$ | 19 (~4%) | 5 (~2%) | 14 (~6%) | 150 (~24%) | 63 (~17%) | 87 (~32%) |
| $0.3 < x_r \leq 0.5$ | – | – | – | 72 (~11%) | 24 (~7%) | 48 (~18%) |
| $x_r > 0.5$ | – | – | – | 75 (~12%) | 36 (~10%) | 39 (~14%) |

(named as “Autoscaling statistics”) shows the number of ionograms falling in each subset (1, 2, 3 or 4), while the bottom part (named as “Autoscaling accuracy subset 1”) shows the autoscaling accuracy related to subset 1, in terms of four different ranges ($x_a \leq 0.5$, $0.5 < x_a \leq 1.0$, $1.0 < x_a \leq 1.5$, and $x_a > 1.5$) of the absolute error $x_a =$ |autoscaled value – validated value|, and in terms of 4 different ranges ($x_r \leq 0.1$, $0.1 < x_r \leq 0.3$, $0.3 < x_r \leq 0.5$, and $x_r > 0.5$) of the relative error $x_r =$ (autoscaled value – validated value)/validated value). In both cases, a distinction between daytime ionograms (recorded between 6:00 and 18:00 LT) and nighttime ionograms (recorded between 18:00 and 6:00 LT) was also made.

Moreover, for each of the seven considered periods, corresponding linear fits between autoscaled and validated values of subset 1 are shown for Autoscala and UDIDA in Figs. 2 and 3, for quiet and disturbed periods, respectively; the number visible at the bottom right corner of each plot is the Pearson correlation coefficient, according to the following formula

$$\rho_{X,Y} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{E((X - E(X))(Y - E(Y)))}{\sigma_X \sigma_Y}, \quad (1)$$

where $\text{cov}()$ is the covariance between the variables X and Y , σ_X and σ_Y are the corresponding standard deviations, and $E()$ represents the expected value. Specifically, to compare the performances of both autoscaling systems at the same level, besides showing the linear fits for the whole subset 1 of Autoscala (left columns of plots) and for the whole subset 1 of UDIDA (right column of plots), Figs. 2 and 3 show also the linear fits for the part of subset 1 of UDIDA, which is coincident with the subset 1 of Autoscala (middle column of plots).

Discussion

Looking at Tables 1, 2, 3, 4, 5, 6 and 7, the different philosophy characterizing the two autoscaling systems is evident. On the one hand, UDIDA gives always a value for

Table 5 Same as Table 1 but for the 5-min ionograms recorded during 14–16 December 2006, a disturbed period characterized by a D_{st} of -162 nT

| Storm 14–16 December 2006 ($D_{st} = -162$ nT) (860 ionograms) | | | | | | |
|---|---------------------------|-------------------------|---------------------------|-----------------------|-------------------------|---------------------------|
| Autoscaling statistics | | | | | | |
| | Autoscala | | | UDIDA | | |
| Subset 1 | 316 (~48%) | | | 658 | | |
| Subset 2 | 133 (~66%) | | | – | | |
| Subset 3 | 69 (~34%) | | | 202 | | |
| Subset 4 | 342 (~52%) | | | – | | |
| Autoscaling accuracy subset 1 | | | | | | |
| $x_a =$ lautoscaled value – validated value (MHz) | Autoscala (316 ionograms) | | | UDIDA (658 ionograms) | | |
| | Total ionograms (316) | Daytime ionograms (157) | Nighttime ionograms (159) | Total ionograms (658) | Daytime ionograms (423) | Nighttime ionograms (235) |
| $x_a \leq 0.5$ | 292 (~92%) | 140 (~89%) | 152 (~96%) | 300 (~46%) | 195 (~46%) | 105 (~45%) |
| $0.5 < x_a \leq 1.0$ | 22 (~7%) | 15 (~9%) | 7 (~4%) | 101 (~15%) | 62 (~15%) | 39 (~17%) |
| $1.0 < x_a \leq 1.5$ | 1 (~0.5%) | 1 (~1%) | – | 73 (~11%) | 46 (~11%) | 27 (~11%) |
| $x_a > 1.5$ | 1 (~0.5%) | 1 (~1%) | – | 184 (~28%) | 120 (~28%) | 64 (~27%) |
| $x_r =$ lautoscaled value – validated value/validated value | Autoscala (316 ionograms) | | | UDIDA (658 ionograms) | | |
| | Total ionograms (316) | Daytime ionograms (157) | Nighttime ionograms (159) | Total ionograms (658) | Daytime ionograms (423) | Nighttime ionograms (235) |
| $x_r \leq 0.1$ | 296 (~93%) | 144 (~92%) | 152 (~96%) | 367 (~56%) | 256 (~61%) | 111 (~47%) |
| $0.1 < x_r \leq 0.3$ | 19 (~6%) | 12 (~7%) | 7 (~4%) | 168 (~25%) | 106 (~25%) | 62 (~27%) |
| $0.3 < x_r \leq 0.5$ | 1 (~1%) | 1 (~1%) | – | 56 (~9%) | 27 (~6%) | 29 (~12%) |
| $x_r > 0.5$ | – | – | – | 67 (~10%) | 34 (~8%) | 33 (~14%) |

$foF2$, independently of the presence of the ionogram trace and, if present, independently of its definition; on the other hand, Autoscala tends to discard ionograms for which the digital information is considered insufficient, giving no value of $foF2$ as output. This different kind of processing explains why the number of ionograms falling in the subset 1 is greater for UDIDA than for Autoscala, while the number of ionograms falling in the subset 4 is greater for Autoscala than for UDIDA. In fact, for both quiet and disturbed conditions, and independent of the solar activity, the trace referring to the F2 layer is weak and scarcely defined for most of the ionograms, and consequently Autoscala does not give any $foF2$ value as output. As an example, Fig. 4 shows an ionogram for which the operator scaled a value of 11.0 MHz for $foF2$, while Autoscala considered the digital information insufficient to identify the ionogram trace. This Autoscala behavior might be smoothed by varying the thresholds used by the algorithm

to identify the F2 trace, but this is a long and hard task that could be done as a further step.

At the same time, however, this feature allows Autoscala to identify correctly the ionogram cases for which the operator did not validate any $foF2$ value, and in fact the number of ionograms falling in the subset 2 for Autoscala is high. On the contrary, UDIDA cannot identify this kind of ionograms, and correspondingly the number of ionograms falling in the subset 3 is large. This is a feature of UDIDA that has to be certainly improved in the near future, namely this system should be equipped with a subroutine able to reliably identify whether the ionogram trace is absent or not.

Figures 5 and 6 show two examples of ionograms falling in the subsets 1 and 3, respectively, as processed by Autoscala and UDIDA.

Concerning the number of ionograms falling in subset 3, this shows a tendency to increase for both Autoscala and

Table 6 Same as Table 1 but for the 5-min ionograms recorded during 24–26 October 2011, a disturbed period characterized by a D_{st} of -147 nT

| Storm 24–26 October 2011 ($D_{st} = -147$ nT) (761 ionograms) | | | | | | |
|--|---------------------------|------------------------|--------------------------|-----------------------|-------------------------|---------------------------|
| Autoscaling statistics | | | | | | |
| | Autoscala | | | UDIDA | | |
| Subset 1 | 103 (~19%) | | | 539 | | |
| Subset 2 | 203 (~91%) | | | – | | |
| Subset 3 | 19 (~9%) | | | 222 | | |
| Subset 4 | 436 (~81%) | | | – | | |
| Autoscaling accuracy subset 1 | | | | | | |
| $x_a =$ lautoscaled value – validated value (MHz) | Autoscala (103 ionograms) | | | UDIDA (539 ionograms) | | |
| | Total ionograms (103) | Daytime ionograms (55) | Nighttime ionograms (48) | Total ionograms (539) | Daytime ionograms (391) | Nighttime ionograms (148) |
| $x_a \leq 0.5$ | 84 (~82%) | 37 (~67%) | 47 (~98%) | 253 (~47%) | 192 (~49%) | 61 (~41%) |
| $0.5 < x_a \leq 1.0$ | 10 (~10%) | 9 (~16%) | 1 (~2%) | 137 (~25%) | 92 (~24%) | 45 (~30%) |
| $1.0 < x_a \leq 1.5$ | 2 (~1%) | 2 (~4) | – | 100 (~19%) | 71 (~18%) | 29 (~20%) |
| $x_a > 1.5$ | 7 (~7%) | 7 (~13%) | – | 49 (~9%) | 36 (~9%) | 13 (~9%) |
| $x_r =$ lautoscaled value – validated value/validated value | Autoscala (103 ionograms) | | | UDIDA (539 ionograms) | | |
| | Total ionograms (103) | Daytime ionograms (55) | Nighttime ionograms (48) | Total ionograms (539) | Daytime ionograms (391) | Nighttime ionograms (148) |
| $x_r \leq 0.1$ | 92 (~89%) | 44 (~80%) | 48 (100%) | 424 (~79%) | 314 (~80%) | 110 (~74%) |
| $0.1 < x_r \leq 0.3$ | 9 (~9%) | 9 (~16%) | – | 102 (~18%) | 70 (~18%) | 32 (~22%) |
| $0.3 < x_r \leq 0.5$ | 1 (~1%) | 1 (~2%) | – | 9 (~2%) | 4 (~1%) | 5 (~3%) |
| $x_r > 0.5$ | 1 (~1%) | 1 (~2%) | – | 4 (~1%) | 3 (~1%) | 1 (~1%) |

UDIDA for disturbed periods. For Autoscala, also the number of ionograms falling in the subset 4 seems to increase for disturbed periods.

Concerning the number of ionograms falling in the subset 1, for both UDIDA and Autoscala it seems to be independent of both the period and the solar activity.

Focusing the attention on the accuracy of the autoscaling of ionograms of subset 1, this is very high for Autoscala, regardless of both the considered period and the solar activity, while for UDIDA it decreases for disturbed periods. Moreover, Tables 1, 2, 3, 4, 5, 6 and 7 show also that the accuracy of the UDIDA autoscaling seems to slightly decrease for nighttime ionograms. As it is shown in Fig. 7, this is mainly due to the occurrence of multiple reflections of the sporadic E (Es) layer and the presence of noise, which mislead the algorithm, causing as a consequence large absolute and relative errors.

The high accuracy characterizing the autoscaling made by Autoscala can be also inferred by looking at Figs. 2 and 3, where the linear fits related to Autoscala are all very good, with values of the Pearson correlation coefficient ranging from 0.87 to 0.99. This means that when Autoscala gives a value of $foF2$ as output, this is very reliable.

Instead, the linear fits related to UDIDA, except the cases of September 2000 and March 2004, are worse than those related to Autoscala, showing a more significant spread of the points, with values of the Pearson correlation coefficient that are lower, especially for disturbed periods.

The main causes of error for UDIDA are then: (1) the presence of spread-F (Fig. 6); (2) the presence of noise (Fig. 7a); (3) the appearance of multiple reflections of the sporadic E layer (Fig. 7b); (4) the appearance of multiple reflections of the F layer (Fig. 8). Concerning (1), UDIDA

Table 7 Same as Table 1 but for the 5-min ionograms recorded during 17–19 March 2015, a disturbed period characterized by a D_{st} of -223 nT
St. Patrick storm 17–19 March 2015 ($D_{st} = -223$ nT) (864 ionograms)

| Autoscaling statistics | | | | | | |
|--|---------------------------|-------------------------|---------------------------|-----------------------|-------------------------|---------------------------|
| | Autoscala | | | UDIDA | | |
| Subset 1 | 269 (~46%) | | | 613 | | |
| Subset 2 | 216 (~86%) | | | – | | |
| Subset 3 | 35 (~14%) | | | 251 | | |
| Subset 4 | 344 (~57%) | | | – | | |
| Autoscaling accuracy subset 1 | | | | | | |
| $x = \text{lautoscaled value} - \text{validated value}$ (MHz) | Autoscala (269 ionograms) | | | UDIDA (613 ionograms) | | |
| | Total ionograms (269) | Daytime ionograms (135) | Nighttime ionograms (134) | Total ionograms (613) | Daytime ionograms (352) | Nighttime ionograms (261) |
| $x \leq 0.5$ | 250 (~93%) | 125 (~93%) | 125 (~93%) | 238 (~39%) | 143 (~41%) | 95 (~36%) |
| $0.5 < x \leq 1.0$ | 16 (~6%) | 7 (~5%) | 9 (~7%) | 110 (~18%) | 59 (~17%) | 51 (~20%) |
| $1.0 < x \leq 1.5$ | 3 (~1%) | 3 (~2%) | – | 97 (~16%) | 68 (~19%) | 29 (~11%) |
| $x > 1.5$ | – | – | – | 168 (~27%) | 82 (~23%) | 86 (~33%) |
| $x = \text{lautoscaled value} - \text{validated value} / \text{validated value}$ | Autoscala (269 ionograms) | | | UDIDA (613 ionograms) | | |
| | Total ionograms (269) | Daytime ionograms (135) | Nighttime ionograms (134) | Total ionograms (613) | Daytime ionograms (352) | Nighttime ionograms (261) |
| $x \leq 0.1$ | 252 (~94%) | 129 (~96%) | 123 (~92%) | 357 (~58%) | 229 (~65%) | 128 (~49%) |
| $0.1 < x \leq 0.3$ | 17 (~6%) | 6 (~4%) | 11 (~8%) | 146 (~24%) | 84 (~24%) | 62 (~24%) |
| $0.3 < x \leq 0.5$ | – | – | – | 50 (~8%) | 24 (~7%) | 26 (~10%) |
| $x > 0.5$ | – | – | – | 60 (~10%) | 15 (~4%) | 45 (~17%) |

considers the trace definite enough even though it is characterized by a strong spread-F phenomenon which induces unreliable f_oF2 values; concerning (2), the presence of clusters of points due to noise causes a wrong identification of the trace and a consequent overestimation of the value given as output; concerning (3) and (4), UDIDA is misled by traces that are not those of the real ionogram. A significant effort has been invested by the UDIDA research team to smooth such errors (Pillat et al. 2015).

Autoscala instead can manage quite well the presence of multiple reflections (Scotto and Pezzopane 2008), and the main causes of error are: (1) a pronounced weakness of the trace (Fig. 4); (2) the presence of spread-F (Fig. 6); (3) the appearance of additional stratifications of the F layer (Lynn et al. 2000; Fagundes et al. 2007) (Fig. 9). Concerning (1), as it was previously mentioned, this malfunctioning of Autoscala might be mitigated by working on thresholds used by the algorithm; concerning (2), it is worth

highlighting that most of the ionograms characterized by spread-F phenomena are correctly identified by Autoscala as not identifiable; concerning (3), the problem is known (Scotto 2009) and more complex to be fixed, because it implies a reconsideration of the empirical curves used by Autoscala to identify the different parts of the ionogram; at the same time, Fig. 9 shows that UDIDA can manage better than Autoscala the presence of additional stratifications of the F layer.

Conclusions

By considering ionograms recorded by the CADI ionosonde installed at SJC, for quiet and disturbed periods, two autoscaling systems, Autoscala and UDIDA, were compared. UDIDA gives a value of f_oF2 for each analyzed ionogram, while Autoscala tends to reject ionograms for which the digital information is considered

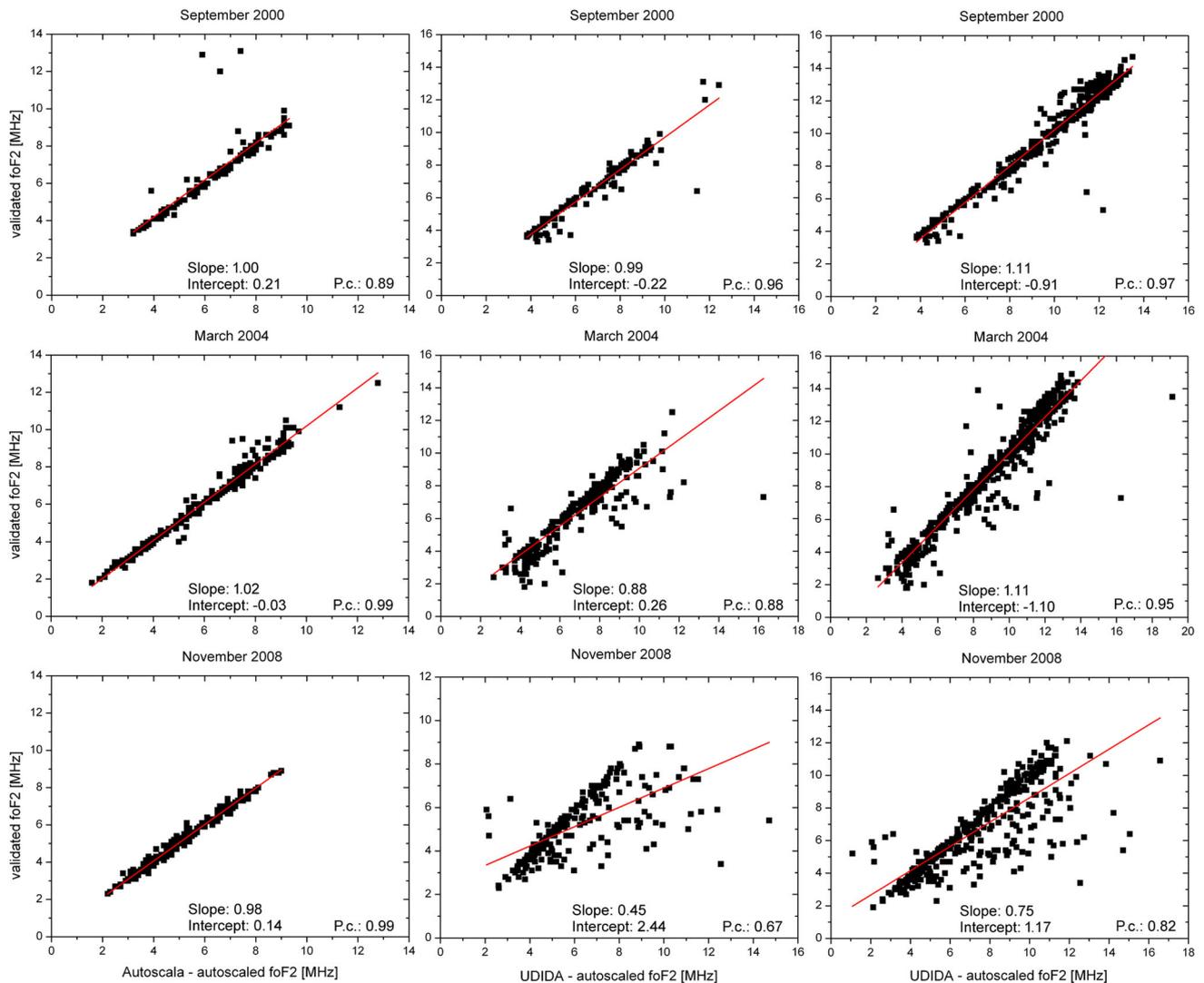


Fig. 2 Linear fits (in red the regression line, the slope and the intercept values of which are reported at the *bottom* of each plot) for the whole subset 1 (see the text for the description of the four different subsets) of Autoscala (*left column*), for the whole subset 1 of UDIDA (*right column*), and for the part of subset 1 of UDIDA, which is

coincident with the subset 1 of Autoscala (*middle column*), for the quiet periods of September 2000, March 2004, and November 2008. The value of the Pearson correlation coefficient is reported at the *bottom right corner* of each plot

insufficient, and no value of $foF2$ is given as output. This means that UDIDA can give more $foF2$ values as output than Autoscala, and that Autoscala can correctly discard a large number of ionograms for which the operator does not validate any $foF2$ value. With regard to the dependence of the autoscaling on the period and the solar activity, when the operator validates a value of $foF2$, the

capacity of both autoscaling systems in giving a value as output does not present any kind of pattern. Instead, the rejection made by Autoscala for ionograms validated by an operator seems to increase for disturbed conditions. The same pattern seems to characterize the number of ionograms not validated by an operator and incorrectly autoscaled by both systems. The accuracy of the

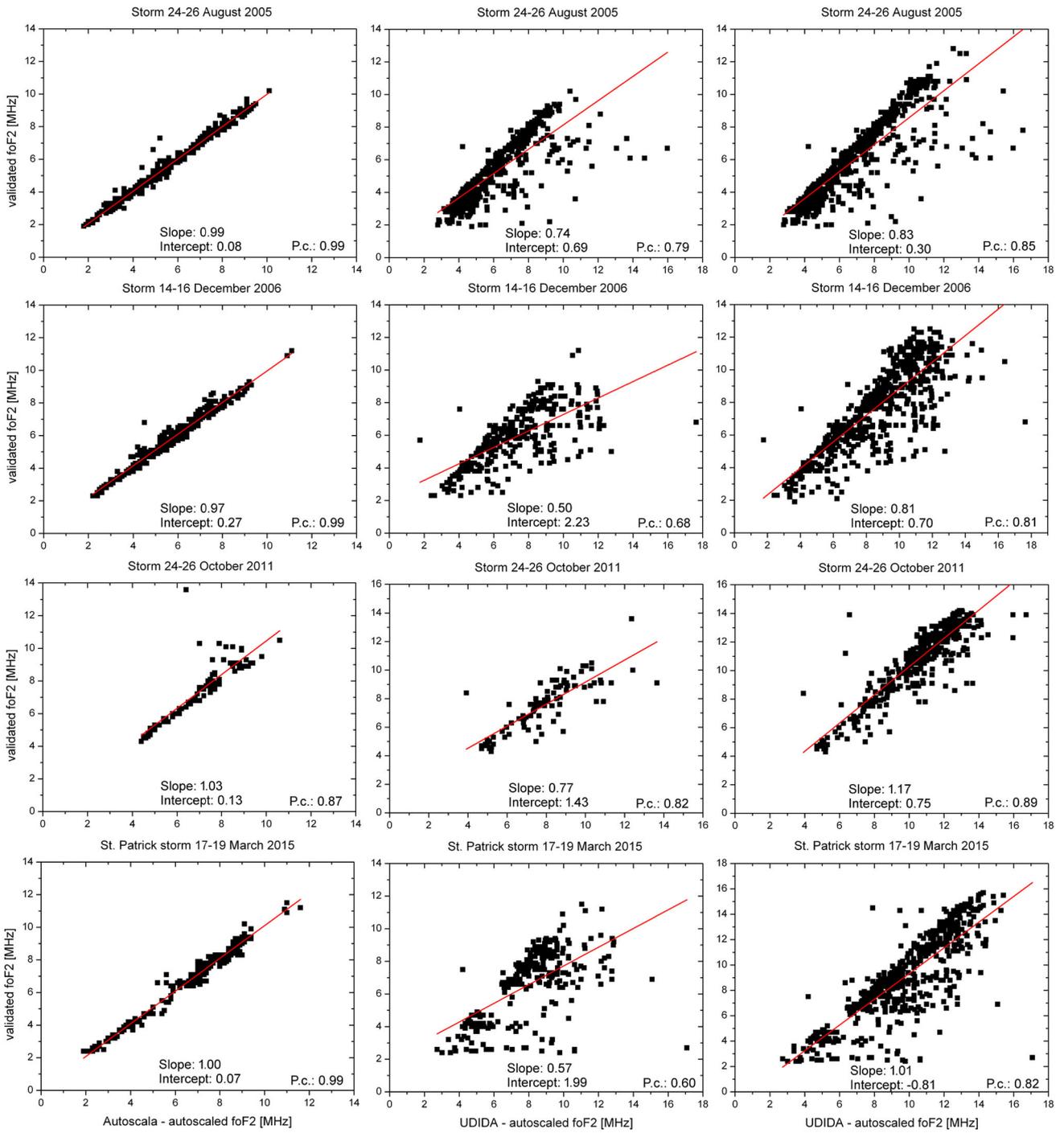


Fig. 3 Same as Fig. 2 but for the disturbed periods that occurred during 24–26 August 2005, 14–16 December 2006, 24–26 October 2011, and 17–19 March 2015

Fig. 4 Example of ionogram for which the operator scales a value of 11.0 MHz for f_oF_2 , while Autoscala does not give any value as output

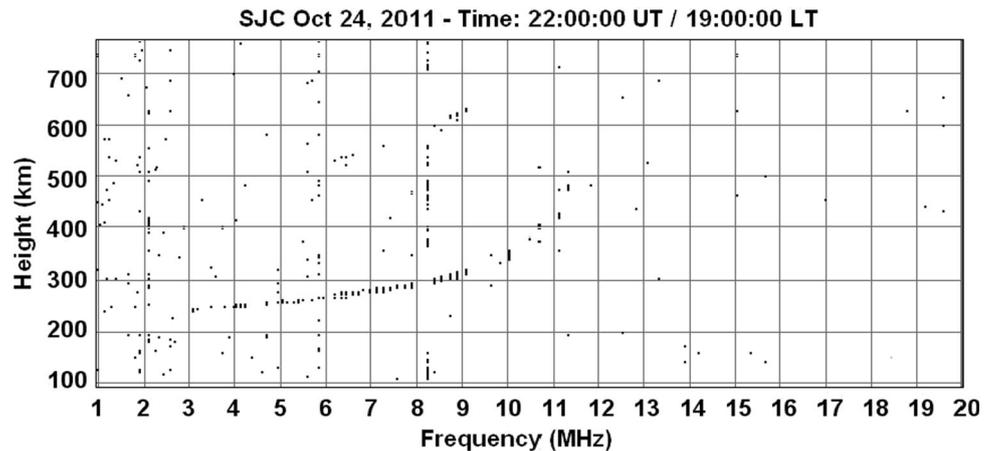
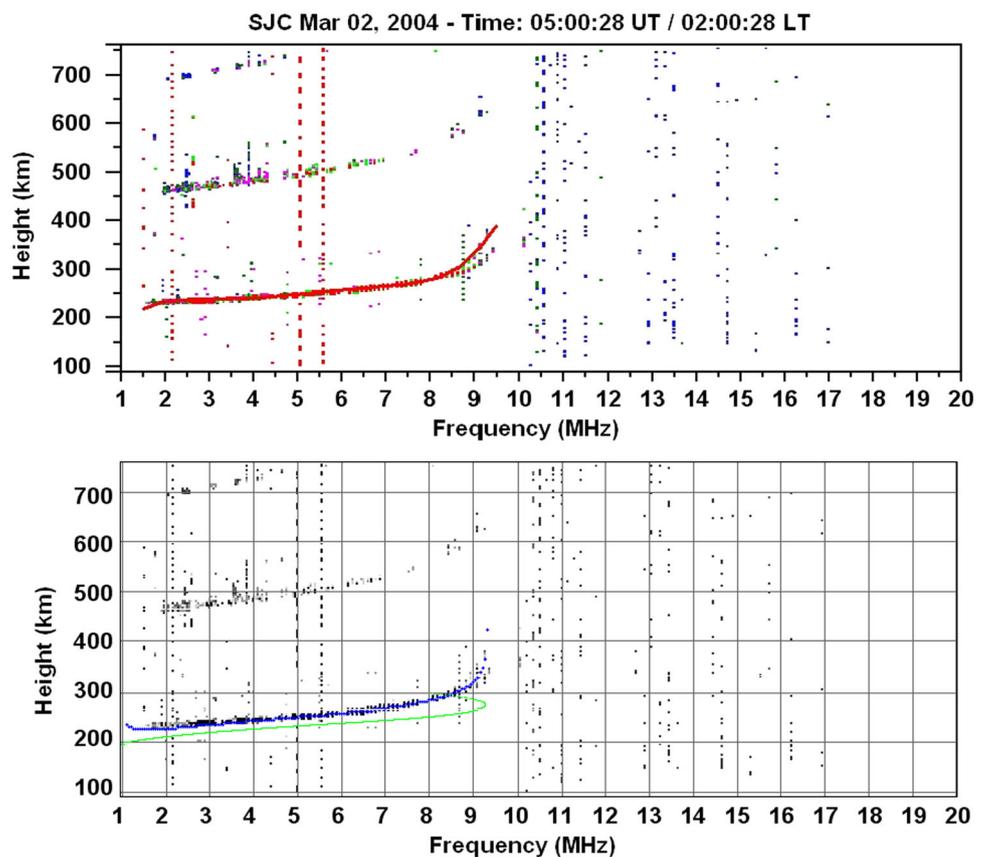


Fig. 5 Example of ionogram of subset 1 as autoscaled by (top) UDIDA and (bottom) Autoscala. The red curve in the top panel is the ionogram reconstructed by UDIDA, while the blue and the green curves in the bottom panel are, respectively, the ionogram reconstructed and the corresponding vertical electron density profile given by Autoscala



autoscaling is very high for Autoscala, independent of both the considered period and solar activity, while UDIDA shows an accuracy decrease both for disturbed periods, mainly because of the presence of multiple reflections of the sporadic E layer, multiple reflections of the F layer, and spread-F, and for nighttime ionograms because of noise. The main shortcomings showed by

Autoscala are instead related to the inability in identifying the ionogram when this is characterized by a pronounced weakness of the trace, the presence of spread-F, and the appearance of additional stratifications of the F layer.

To smooth the aforementioned autoscaling errors made by both systems, so that they can produce more

Fig. 6 Same as Fig. 5 but for an ionogram of subset 3

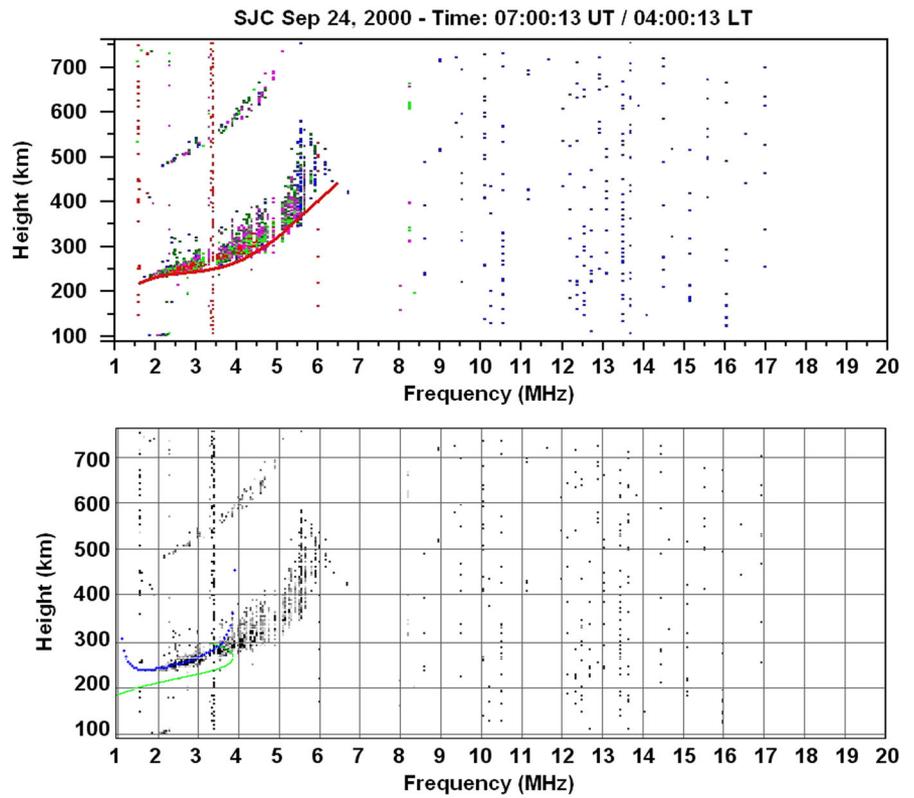


Fig. 7 Example of nighttime ionograms for which UDIDA was misled by: **a** the presence of noise characterizing the ionogram trace, and **b** the appearance of multiple reflections of the Es layer

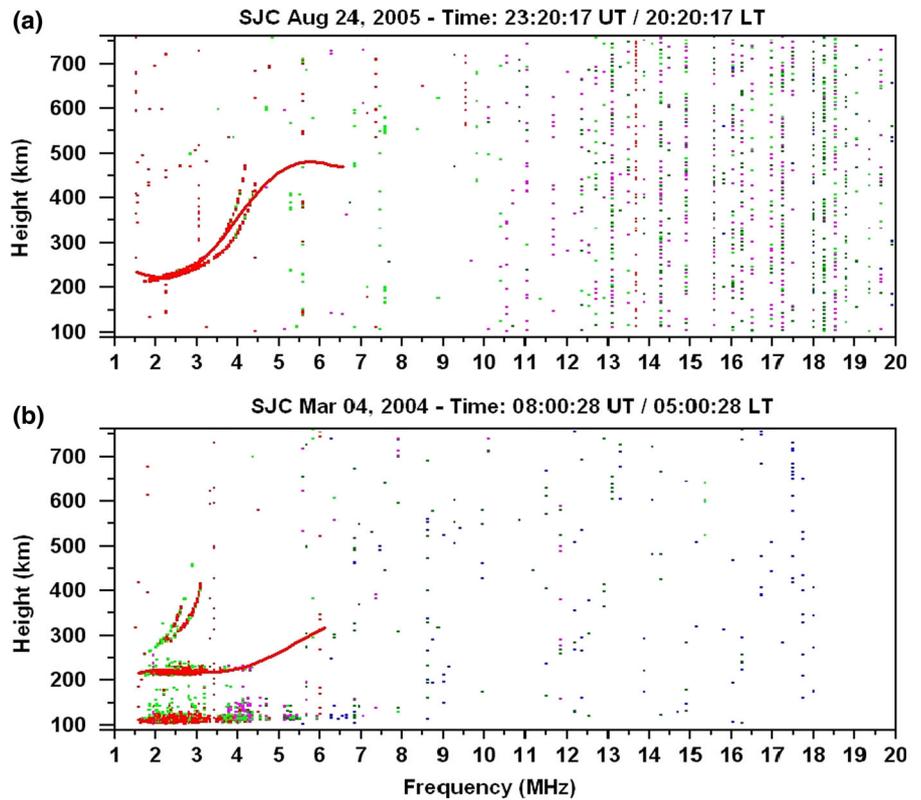


Fig. 8 Example of ionogram for which UDIDA was misled by the appearance on the ionograms of multiple reflections of the F layer

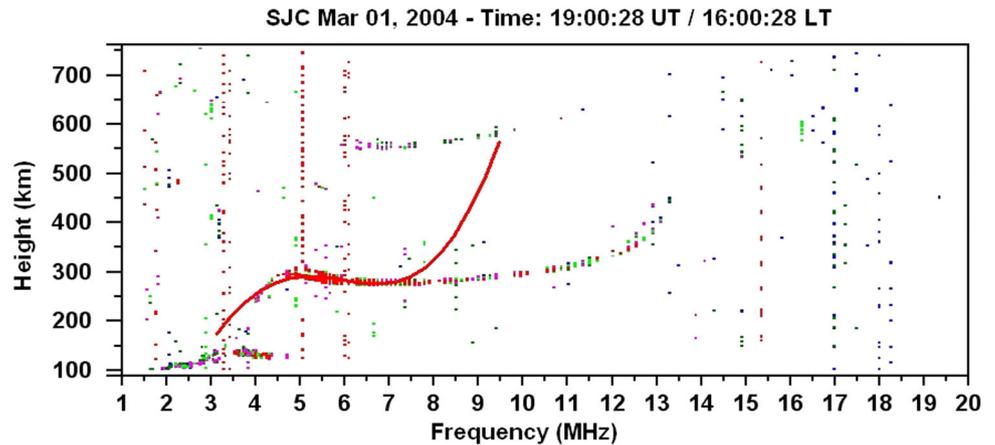
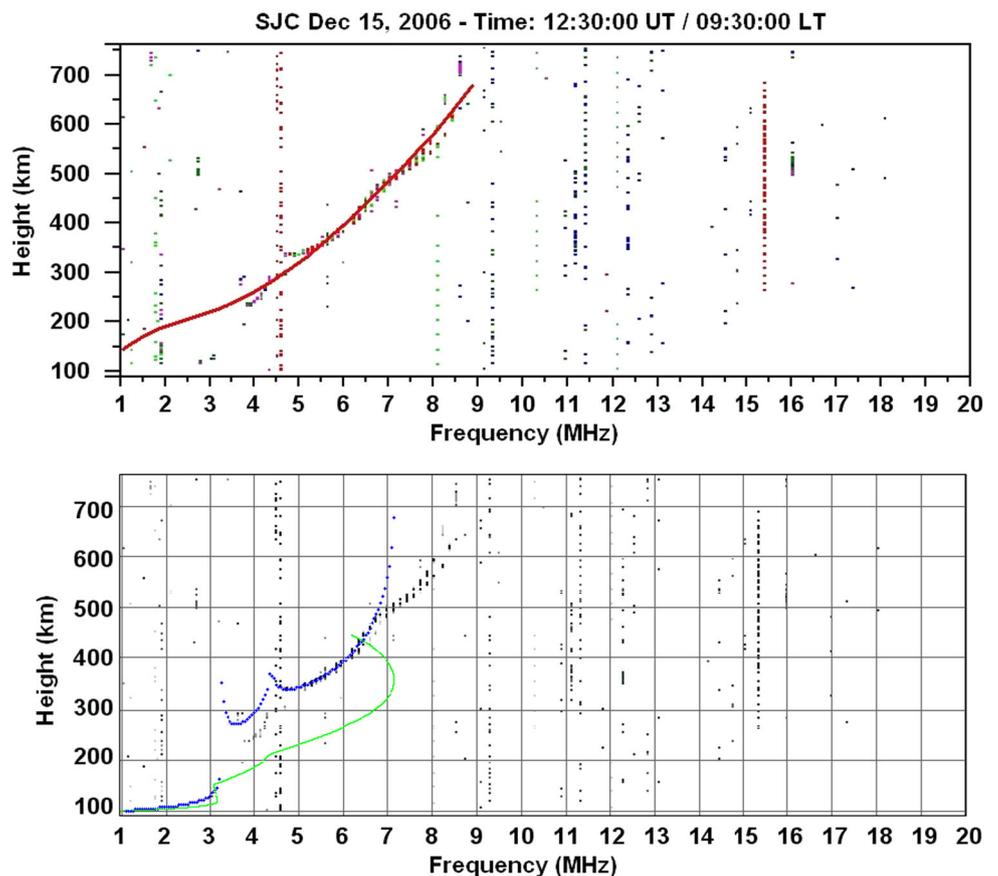


Fig. 9 Example of ionogram characterized by the appearance of additional stratifications of the F layer for which (top) UDIDA gives as output a reliable value of foF_2 , while (bottom) Autoscala performs a strong underestimation of it



reliable foF_2 values, further fine-tuning analysis is needed. This can lead to the development of a reliable real-time monitoring system of the foF_2 ionospheric characteristic at SJC.

References

- Astafyeva E, Zakharenkova I, Förster M (2015) Ionospheric response to the 2015 St. Patrick's Day storm: a global multi-instrumental overview. *J Geophys Res* 120(10):9023–9037. doi:[10.1002/2015JA02629](https://doi.org/10.1002/2015JA02629)
- Bamford RA, Stamper R, Cander LR (2008) A comparison between the hourly autoscaled and manually scaled characteristics from the Chilton ionosonde from 1996 to 2004. *Radio Sci* 43(1):RS1001. doi:[10.1029/2005RS003401](https://doi.org/10.1029/2005RS003401)
- Bullett T, Malagnini A, Pezzopane M, Scotto C (2010) Application of Autoscala to ionograms recorded by the VIPIR ionosonde. *Adv Space Res* 45(9):1156–1172. doi:[10.1016/j.asr.2010.01.024](https://doi.org/10.1016/j.asr.2010.01.024)
- Carter BA, Yizengaw E, Pradipta R, Retterer JM, Groves K, Valladares C, Caton R, Bridgwood C, Norman R, Zhang K (2016) Global equatorial plasma bubble occurrence during the

- 2015 St. Patrick's Day storm. *J Geophys Res* 121(1):894–905. doi:[10.1002/2015JA022194](https://doi.org/10.1002/2015JA022194)
- Chartier AT, Matsuo T, Anderson JL, Collins N, Hoar TJ, Lu G, Mitchell CN, Coster AJ, Paxton LJ, Bust GS (2016) Ionospheric data assimilation and forecasting during storms. *J Geophys Res* 121(1):764–778. doi:[10.1002/2014JA020799](https://doi.org/10.1002/2014JA020799)
- Chen Z, Wang S, Zhang S, Fang G, Wang J (2013) Automatic scaling of F layer from ionograms. *Radio Sci* 48(3):334–343. doi:[10.1002/rds.20038](https://doi.org/10.1002/rds.20038)
- Conkright RO, McNamara LF (1997) Quality control of automatically scaled vertical incidence ionogram data. *Radio Sci* 32(5):1997–2002. doi:[10.1029/97RS01031](https://doi.org/10.1029/97RS01031)
- De Michelis P, Consolini G, Tozzi R, Marcucci MF (2016) Observations of high-latitude geomagnetic field fluctuations during St. Patrick's Day storm: Swarm and SuperDARN measurements. *Earth Planets Space* 68(105):1–16. doi:[10.1186/s40623-016-0476-3](https://doi.org/10.1186/s40623-016-0476-3)
- Ding Z, Ning B, Wan W, Liu L (2007) Automatic scaling of F2-layer parameters from ionograms based on the empirical orthogonal function (EOF) analysis of ionospheric electron density. *Earth Planets Space* 59(1):51–58. doi:[10.1186/BF03352022](https://doi.org/10.1186/BF03352022)
- Enell C-F, Kozlovsky A, Turunen T, Ulich T, Väitalo S, Scotto C, Pezzopane M (2016) Comparison between manual scaling and Autoscala automatic scaling applied to Sodankylä Geophysical Observatory ionograms. *Geosci Instrum Method Data Syst* 5:53–64. doi:[10.5194/gi-5-53-2016](https://doi.org/10.5194/gi-5-53-2016)
- Fagundes PR, Klausner V, Sahai Y, Pillat VG, Becker-Guedes F, Bertoni FCP, Bolzan MJA, Abalde JR (2007) Observations of daytime F2-layer stratification under the southern crest of the equatorial ionization anomaly region. *J Geophys Res* 112(A4):A04302. doi:[10.1029/2006JA011888](https://doi.org/10.1029/2006JA011888)
- Fox MW, Blundell C (1989) Automatic scaling of digital ionograms. *Radio Sci* 24(6):747–761. doi:[10.1029/RS024i006p00747](https://doi.org/10.1029/RS024i006p00747)
- Galkin IA, Reinisch BW (2008) The new ARTIST 5 for all digisondes, Ionosonde Netw Adv Group Bull 69. <http://www.ips.gov.au/IPSHosted/INAG/web-69/2008/artist5-inag.pdf>
- Galkin IA, Reinisch BW, Huang X, Bilitza D (2012) Assimilation of GIRO data into a real-time IRI. *Radio Sci* 47(4):RS0L07. doi:[10.1029/2011RS004952](https://doi.org/10.1029/2011RS004952)
- Gardner LC, Schunk RW, Scherliess L, Sojka JJ, Zhu L (2014) Global assimilation of ionospheric measurements-Gauss Markov model: improved specifications with multiple data types. *Space Weather* 12(12):675–688. doi:[10.1002/2014SW001104](https://doi.org/10.1002/2014SW001104)
- Gilbert JD, Smith RW (1988) A comparison between the automatic ionogram scaling system ARTIST and the standard manual method. *Radio Sci* 23(6):968–974. doi:[10.1029/RS023i006p00968](https://doi.org/10.1029/RS023i006p00968)
- Igi S, Nozaki K, Nagayama M, Ohtani A, Kato H, Igarashi K (1993) Automatic ionogram processing systems in Japan. In: Proceedings of Session G6, 24th General Assembly of the International Union of Radio Science (URSI), 25 August–2 September 1993, Kyoto
- Jacobs L, Poole AWV, McKinnell LA (2004) An analysis of automatically scaled F1 layer data over Grahamstown, South Africa. *Adv Space Res* 34(9):1949–1952. doi:[10.1016/j.asr.2004.06.009](https://doi.org/10.1016/j.asr.2004.06.009)
- Jiang C, Yang G, Zhao Z, Zhang Y, Zhu P, Sun H (2013) An automatic scaling technique for obtaining F2 parameters and F1 critical frequency from vertical incidence ionograms. *Radio Sci* 48(6):739–751. doi:[10.1002/2013RS005223](https://doi.org/10.1002/2013RS005223)
- Jiang C, Yang G, Zhao Z, Zhang Y, Zhu P, Sun H, Zhou C (2014) A method for the automatic calculation of electron density profiles from vertical incidence ionograms. *J Atmos Sol Terr Phys* 107:20–29. doi:[10.1016/j.jastp.2013.10.012](https://doi.org/10.1016/j.jastp.2013.10.012)
- Jiang C, Zhang Y, Yang G, Zhu P, Sun H, Cui X, Song H, Zhao Z (2015a) Automatic scaling of the sporadic E layer and removal of its multiple reflection and backscatter echoes for vertical incidence ionograms. *J Atmos Sol Terr Phys* 129:41–48. doi:[10.1016/j.jastp.2015.04.005](https://doi.org/10.1016/j.jastp.2015.04.005)
- Jiang C, Yang G, Lan T, Zhu P, Song H, Zhou C, Cui X, Zhao Z, Zhang Y (2015b) Improvement of automatic scaling of vertical incidence ionograms by simulated annealing. *J Atmos Sol Terr Phys* 133:178–184. doi:[10.1016/j.jastp.2015.09.002](https://doi.org/10.1016/j.jastp.2015.09.002)
- Krasheninnikov I, Pezzopane M, Scotto C (2010) Application of Autoscala to ionograms recorded by the AIS-Parus ionosonde. *Comp Geosci* 36(5):628–635. doi:[10.1016/j.cageo.2009.09.013](https://doi.org/10.1016/j.cageo.2009.09.013)
- Lee IT, Matsuo T, Richmond AD, Liu JY, Wang W, Lin CH, Anderson JL, Chen MQ (2012) Assimilation of FORMOSAT-3/COSMIC electron density profiles into a coupled thermosphere/ionosphere model using ensemble Kalman filtering. *J Geophys Res* 117(A10):A10318. doi:[10.1029/2012JA017700](https://doi.org/10.1029/2012JA017700)
- Lynn KJW, Harris TJ, Sjarifudin M (2000) Stratification of the F2 layer observed in Southeast Asia. *J Geophys Res* 105(A12):27147–27156. doi:[10.1029/2000JA900056](https://doi.org/10.1029/2000JA900056)
- MacDougall JW, Grant IF, Shen X (1993) The Canadian advanced digital ionosonde: design and results. In: Proceedings of 24th General Assembly of the International Union of Radio Science (URSI), 25 August–2 September 1993, Kyoto
- McNamara LF (2006) Quality figures and error bars for autoscaled Digisonde vertical incidence ionograms. *Radio Sci* 41(4):RS4011. doi:[10.1029/2005RS003440](https://doi.org/10.1029/2005RS003440)
- McNamara LF, Angling MJ, Elvidge S, Fridman SV, Hausman MA, Nickisch LJ, McKinnell LA (2013) Assimilation procedures for updating ionospheric profiles below the F2 peak. *Radio Sci* 48(2):143–157. doi:[10.1002/rds.20020](https://doi.org/10.1002/rds.20020)
- Nava B, Rodríguez-Zuluaga J, Alazo-Cuartas K, Kashcheyev A, Migoya-Orué Y, Radicella SM, Amory-Mazaudier C, Fleury R (2016) Middle-and low-latitude ionosphere response to 2015 St. Patrick's Day geomagnetic storm. *J Geophys Res* 121(4):3421–3438. doi:[10.1002/2015JA022299](https://doi.org/10.1002/2015JA022299)
- Nayak C, Tsai L-C, Su S-Y, Galkin IA, Tan ATK, Nofri E, Jamjareegulgarn P (2016) Peculiar features of the low-latitude and midlatitude ionospheric response to the St. Patrick's Day geomagnetic storm of 17 March 2015. *J Geophys Res* 121(8):7941–7960. doi:[10.1002/2016JA022489](https://doi.org/10.1002/2016JA022489)
- Pezzopane M (2004) Interpre: a Windows software for semiautomatic scaling of ionospheric parameters from ionograms. *Comp Geosci* 30(1):125–130. doi:[10.1016/j.cageo.2003.09.009](https://doi.org/10.1016/j.cageo.2003.09.009)
- Pezzopane M, Scotto C (2004) Software for the automatic scaling of critical frequency foF2 and MUF(3000)F2 from ionograms applied at the Ionospheric Observatory of Gibilmanna. *Ann Geophys* 47(6):1783–1790
- Pezzopane M, Scotto C (2005) The INGV software for the automatic scaling of foF2 and MUF(3000)F2 from ionograms: a performance comparison with ARTIST 4.01 from Rome data. *J Atmos Sol Terr Phys* 67(12):1063–1073. doi:[10.1016/j.jastp.2005.02.022](https://doi.org/10.1016/j.jastp.2005.02.022)
- Pezzopane M, Scotto C (2007) Automatic scaling of critical frequency foF2 and MUF(3000)F2: a comparison between Autoscala and ARTIST 4.5 on Rome data. *Radio Sci* 42(4):RS4003. doi:[10.1029/2006RS003581](https://doi.org/10.1029/2006RS003581)
- Pezzopane M, Scotto C (2010) Highlighting the F2 trace on an ionogram to improve Autoscala performance. *Comp Geosci* 36(9):1168–1177. doi:[10.1016/j.cageo.2010.01.010](https://doi.org/10.1016/j.cageo.2010.01.010)
- Pezzopane M, Scotto C, Tomasik Ł, Krasheninnikov I (2010) Autoscala: an aid for different ionosondes. *Acta Geophys* 58(3):513–526. doi:[10.2478/s11600-009-0038-1](https://doi.org/10.2478/s11600-009-0038-1)
- Pezzopane M, Pietrella M, Pignatelli A, Zolesi B, Cander LR (2011) Assimilation of autoscaled data and regional and local ionospheric models as input sources for real-time 3-D International Reference Ionosphere modeling. *Radio Sci* 46(5):RS5009. doi:[10.1029/2011RS004697](https://doi.org/10.1029/2011RS004697)

- Pezzopane M, Pietrella M, Pignatelli A, Zolesi B, Cander LR (2013) Testing the three-dimensional IRI-SIRMUP-P mapping of the ionosphere for disturbed periods. *Adv Space Res* 52(10):1726–1736. doi:[10.1016/j.asr.2012.11.028](https://doi.org/10.1016/j.asr.2012.11.028)
- Piggott WR, Rawer K (1972) U.R.S.I. Handbook of Ionogram Interpretation and Reduction. World Data Center A for Solar-Terrestrial Physics, NOAA, Boulder
- Pignalberi A, Pezzopane M, Tozzi R, De Michelis P, Coco I (2016) Comparison between IRI and preliminary Swarm Langmuir probe measurements during the St. Patrick storm period. *Earth Planets Space* 68(93):1–18. doi:[10.1186/s40623-016-0466-5](https://doi.org/10.1186/s40623-016-0466-5)
- Pillat VG, Guimaraes LNF, Fagundes PR, da Silva JDS (2013) A computational tool for ionosonde CADI's ionogram analysis. *Comp Geosci* 52:372–378. doi:[10.1016/j.cageo.2012.11.009](https://doi.org/10.1016/j.cageo.2012.11.009)
- Pillat VG, Fagundes PR, Guimaraes LNF (2015) Automatically identification of Equatorial Spread-F occurrence on ionograms. *J Atmos Sol Terr Phys* 135:118–125. doi:[10.1016/j.jastp.2015.10.015](https://doi.org/10.1016/j.jastp.2015.10.015)
- Reinisch BW, Huang X (1983) Automatic calculation of electron density profiles from digital ionograms: 3. Processing of bottomside ionograms. *Radio Sci* 18(3):477–492. doi:[10.1029/RS018i003p00477](https://doi.org/10.1029/RS018i003p00477)
- Reinisch BW, Huang X, Galkin IA, Paznukhov V, Kozlov A (2005) Recent advances in real-time analysis of ionograms and ionospheric drift measurements with digisondes. *J Atmos Sol Terr Phys* 67(12):1054–1062. doi:[10.1016/j.jastp.2005.01.009](https://doi.org/10.1016/j.jastp.2005.01.009)
- Sabbagh D, Scotto C, Sgrigna V (2016) A regional adaptive and assimilative three-dimensional ionospheric model. *Adv Space Res* 57(5):1241–1257. doi:[10.1016/j.asr.2015.12.038](https://doi.org/10.1016/j.asr.2015.12.038)
- Scotto C (2009) Electron density profile calculation technique for Autoscala ionogram analysis. *Adv Space Res* 44(6):756–766. doi:[10.1016/j.asr.2009.04.037](https://doi.org/10.1016/j.asr.2009.04.037)
- Scotto C, MacDougall J (2012) Application of Autoscala software to the Canadian advanced digital ionosonde. *Int J Remote Sens* 33(17):5574–5582. doi:[10.1080/01431161.2012.661097](https://doi.org/10.1080/01431161.2012.661097)
- Scotto C, Pezzopane M (2008) Removing multiple reflections from the F2 layer to improve Autoscala performance. *J Atmos Sol Terr Phys* 70(15):1929–1934. doi:[10.1016/j.jastp.2008.05.012](https://doi.org/10.1016/j.jastp.2008.05.012)
- Spogli L, Cesaroni C, Di Mauro D, Pezzopane M, Alfonsi L, Musicò E, Povero G, Pini M, Dovis F, Romero R, Linty N, Abadi P, Nuraeni F, Husin A, Le Huy M, Thi Lan T, Vihn La T, Pillat VG, Flourey N (2016) Formation of ionospheric irregularities over Southeast Asia during the 2015 St. Patrick's Day storm. *J Geophys Res* 121(12):12211–12233. doi:[10.1002/2016JA023222](https://doi.org/10.1002/2016JA023222)
- Stankov SM, Jodogne JC, Kutiev I, Stegen K, Warnant R (2012) Evaluation of automatic ionogram scaling for use in real-time ionospheric density profile specification: dourbes DGS-256/ARTIST-4 performance. *Ann Geophys* 55(2):283–291. doi:[10.4401/ag-4976](https://doi.org/10.4401/ag-4976)
- Su F, Zhao Z, Li S, Yao M, Chen G, Zhou Y (2012) Signal identification and trace extraction for the vertical ionogram. *IEEE Geosci Remote Sens Lett* 9(6):1031–1035. doi:[10.1109/LGRS.2012.2189350](https://doi.org/10.1109/LGRS.2012.2189350)
- Tsai LC, Berkey FT (2000) Ionogram analysis using fuzzy segmentation and connectedness techniques. *Radio Sci* 35(5):1173–1186. doi:[10.1029/1999RS002170](https://doi.org/10.1029/1999RS002170)
- Tulasi Ram S, Yokoyama T, Otsuka Y, Shiokawa K, Sripathi S, Veenadhari B, Heelis R, Ajith KK, Gowtam VS, Gurubaran S, Supnithi P, Le Huy M (2016) Duskside enhancement of equatorial zonal electric field response to convection electric fields during the St. Patrick's Day storm on 17 March 2015. *J Geophys Res* 121(1):538–548. doi:[10.1002/2015JA021932](https://doi.org/10.1002/2015JA021932)
- Zabotin NA, Wright JW, Zhubankov GA (2006) NeXtYZ: three-dimensional electron density inversion for dynasonde ionograms. *Radio Sci* 41(6):RS6S32. doi:[10.1029/2005RS003352](https://doi.org/10.1029/2005RS003352)
- Zheng H, Ji G, Wang G, Zhao Z, He S (2013) Automatic scaling of F layer from ionograms based on image processing and analysis. *J Atmos Sol Terr Phys* 105–106:110–118. doi:[10.1016/j.jastp.2013.09.007](https://doi.org/10.1016/j.jastp.2013.09.007)
- Zhong J, Wang W, Yue X, Burns AG, Dou X, Lei J (2016) Long-duration depletion in the topside ionospheric total electron content during the recovery phase of the March 2015 strong storm. *J Geophys Res* 121(5):4733–4747. doi:[10.1002/2016JA022469](https://doi.org/10.1002/2016JA022469)

Magnetoelastic shear wave propagation in pre-stressed anisotropic media under gravity

Nirmala Kumari¹ · Amares Chattopadhyay¹ · Abhishek K. Singh¹ · Sanjeev A. Sahu¹

Received: 23 January 2017 / Accepted: 27 January 2017 / Published online: 15 February 2017
© Institute of Geophysics, Polish Academy of Sciences & Polish Academy of Sciences 2017

Abstract The present study investigates the propagation of shear wave (horizontally polarized) in two initially stressed heterogeneous anisotropic (magnetoelastic transversely isotropic) layers in the crust overlying a transversely isotropic gravitating semi-infinite medium. Heterogeneities in both the anisotropic layers are caused due to exponential variation (case-I) and linear variation (case-II) in the elastic constants with respect to the space variable pointing positively downwards. The dispersion relations have been established in closed form using Whittaker's asymptotic expansion and were found to be in the well-agreement to the classical Love wave equations. The substantial effects of magnetoelastic coupling parameters, heterogeneity parameters, horizontal compressive initial stresses, Biot's gravity parameter, and wave number on the phase velocity of shear waves have been computed and depicted by means of a graph. As a special case, dispersion equations have been deduced when the two layers and half-space are isotropic and homogeneous. The comparative study for both cases of heterogeneity of the layers has been performed and also depicted by means of graphical illustrations.

Keywords Heterogeneity · Magnetoelastic · Transversely isotropic · Shear wave · Initial stress

Introduction

The crust is heterogeneous, which makes the study of wave propagation much practical considering the heterogeneous layers. Different types of vertical heterogeneity persist in crustal layers in the form of exponential function, linear function, etc. The study of wave propagation in layered elastic media with different boundaries helps to understand and predict the seismic behavior at the different margins of earth, which makes it applicable in the field of geophysics, civil, mechanical, and other engineering branches. Many researchers had widely studied the theory of Love wave/shear wave propagation in a medium where the velocity, rigidity and density are functions of depth (e.g., Bhattacharya 1962; Dutta 1963; Chattopadhyay 1975; Gogna 1976; Chattopadhyay and Kar 1977a). Scattering of SH-waves in multi-layered media with irregular interfaces has been discussed by Ding and Dravinski (1996). Guz (2002) has analyzed the three-dimensional linearised theory of elastic waves propagating in initially stressed solids. He formulated surface waves along planar and curvilinear boundaries and interfaces, waves in layers and cylinders, waves in composite materials, waves in hydroelastic systems, and dynamic problems for moving loads. Chattopadhyay et al. (2009) discussed the propagation of shear waves in viscoelastic medium at irregular boundaries. Chattopadhyay et al. (2010) described the effects of point source and heterogeneity on the propagation of SH-waves. Sahu et al. (2014) have studied the shear wave propagation in a heterogeneous fiber-reinforced layer over a half-space.

✉ Nirmala Kumari
nirmala.ism@gmail.com

Amares Chattopadhyay
amares.c@gmail.com

Abhishek K. Singh
abhi.5700@gmail.com

Sanjeev A. Sahu
ism.sanjeev@gmail.com

¹ Department of Applied Mathematics, Indian School of Mines, Dhanbad, Jharkhand, India

Chattopadhyay and Choudhury (1990) have studied the propagation of reflection and transmission of magnetoelastic shear waves in a self-reinforced medium. Due to the presence of many physical factors, a large amount of initial stress evolves in a medium which has a pronounced influence on the propagation of waves, as shown by Biot (1965). These factors may include the overburden layer, variation in temperature, slow process of creep, gravitational field, etc. The Earth is a highly initially stressed medium. Chattopadhyay and Kar (1977b) have discussed the dispersion curves of Love-type waves in an internal stratum of finite thickness under initial stress lying between two semi-infinite isotropic elastic media. Dey and Addy (1978) have shown the effect of initial stresses on the propagation of Love waves by considering the layer and the half-space to be isotropic elastic in one case and viscoelastic in another case. Propagation of Love waves in a non-homogeneous orthotropic elastic layer under changeable initial stress overlying semi-infinite medium was given by Abd-Alla and Ahmad (1999).

Transversely isotropic materials exhibit hexagonal symmetry. The hexagonal crystals like cadmium, cobalt, beryl, and zinc belong to the class of transversely isotropic solids. Backus' upscaling (Backus 1962) is often used to determine the effective transversely isotropic elastic constants of layered media for long wavelength seismic waves. Sur (1963) documented a note on stresses produced by a shearing force moving over the boundary of a semi-infinite transversely isotropic solid. Elastodynamic equations for transversely isotropic solids have been employed by Acharya et al. (2009) to investigate the general theory of transversely isotropic magnetoelastic shear waves in conducting media under initial hydrostatic tension or compression. Bieńkowski et al. (2010) have established that magnetoelastic effect can be used in development of force sensors. This effect was used for sensors in civil engineering, monitoring of large diesel engines in locomotives, for monitoring of ball valves, and for biomedical monitoring. Magnetoelastic effect has to be also considered as a side effect of accidental application of mechanical stresses to the magnetic core of inductive component. The magnetisable body placed in an external magnetic field becomes magnetized by induction. The magnetization is measured by the magnetic polarization, also called magnetization intensity or dipole moment. Chattopadhyay and Singh (2014) have discussed the propagation of a crack due to magnetoelastic shear waves in a self-reinforced medium. Keeping in mind the existence of different types of heterogeneity in the crustal layers and motivated by the fact that Earth is an initially stressed body, we considered the present problem.

In this paper, we have discussed two problems (cases) relating to the propagation of shear waves in two initially

stressed non-homogeneous magnetoelastic transversely isotropic media. Both media are lying over a transversely isotropic half-space which is under gravity. The material constants of both magnetoelastic transversely isotropic layers (M_1 and M_2) are varying exponentially in the first problem (case-I) and varying linearly in the second problem (case-II) as a function of depth (the space variable pointing positively downwards). No attempt has been made to discuss the shear wave propagation in two initially stressed magnetoelastic transversely isotropic layers having vertical heterogeneity and lying over a transversely isotropic half-space under gravity. The dispersion relations have been obtained in closed form using Whittaker's asymptotic expansion. The validation of dispersion relations with the classical case has been made. The pronounced effects of magnetoelastic coupling parameters, heterogeneity parameters, horizontal compressive initial stresses, and Biot's gravity parameter on the phase velocity of shear wave have also been studied numerically. Moreover, comparative study and graphical illustrations have been carried out to reveal some of the important facts.

Basic equations

If u , v , and w are the displacement components in x_1 , x_2 , and x_3 directions, respectively, then for shear wave propagating in the x_1 -direction and causing displacement in the x_2 -direction only, we shall assume that

$$u = w = 0 \quad \text{and} \quad v = v(x_1, x_3, t). \quad (1)$$

The equations governing the propagation of small elastic disturbances in a perfectly conducting transversely isotropic medium having electromagnetic force $\vec{J} \times \vec{B}$ (the Lorentz force, \vec{J} being the electric current density, and \vec{B} being the magnetic induction vector) as the only body force are

$$\tau_{ij,j} + (\vec{J} \times \vec{B})_i = \rho \frac{\partial^2 u_i}{\partial t^2}, \quad (2)$$

where $(\vec{J} \times \vec{B})_i$ is the x_i -component of the force $(\vec{J} \times \vec{B})$ and ρ is the density of the medium.

Here, interaction of mechanical and electromagnetic fields is considered.

Let $u_i = (u, v, w)$, then Eq. 2 can be written as

$$\left. \begin{aligned} \frac{\partial \tau_{11}}{\partial x_1} + \frac{\partial \tau_{12}}{\partial x_2} + \frac{\partial \tau_{13}}{\partial x_3} + (\vec{J} \times \vec{B})_{x_1} &= \rho \frac{\partial^2 u}{\partial t^2} \\ \frac{\partial \tau_{12}}{\partial x_1} + \frac{\partial \tau_{22}}{\partial x_2} + \frac{\partial \tau_{23}}{\partial x_3} + (\vec{J} \times \vec{B})_{x_2} &= \rho \frac{\partial^2 v}{\partial t^2} \\ \frac{\partial \tau_{13}}{\partial x_1} + \frac{\partial \tau_{23}}{\partial x_2} + \frac{\partial \tau_{33}}{\partial x_3} + (\vec{J} \times \vec{B})_{x_3} &= \rho \frac{\partial^2 w}{\partial t^2} \end{aligned} \right\}, \quad (3)$$

where

$$\begin{aligned}
 \tau_{11} &= A \frac{\partial u}{\partial x_1} + (A - 2N) \frac{\partial v}{\partial x_2} + F \frac{\partial w}{\partial x_3}, \\
 \tau_{22} &= (A - 2N) \frac{\partial u}{\partial x_1} + A \frac{\partial v}{\partial x_2} + F \frac{\partial w}{\partial x_3}, \\
 \tau_{33} &= F \frac{\partial u}{\partial x_1} + F \frac{\partial v}{\partial x_2} + C \frac{\partial w}{\partial x_3}, \\
 \tau_{12} &= N \left(\frac{\partial u}{\partial x_2} + \frac{\partial v}{\partial x_1} \right), \\
 \tau_{23} &= L \left(\frac{\partial w}{\partial x_2} + \frac{\partial v}{\partial x_3} \right), \quad \tau_{13} = L \left(\frac{\partial u}{\partial x_3} + \frac{\partial w}{\partial x_1} \right),
 \end{aligned}
 \tag{4}$$

and $A, N, L, F,$ and C are material constants.

With the help of Eq. 1, Eqs. in 3 and 4 become

$$\frac{\partial \tau_{12}}{\partial x_1} + \frac{\partial \tau_{23}}{\partial x_3} + (\vec{J} \times \vec{B})_{x_2} = \rho \frac{\partial^2 v}{\partial t^2},
 \tag{5}$$

where

$$\tau_{12} = N \left(\frac{\partial v}{\partial x_1} \right), \quad \text{and} \quad \tau_{23} = L \left(\frac{\partial v}{\partial x_3} \right).
 \tag{6}$$

The well-known Maxwell’s equations governing the electromagnetic field are

$$\begin{aligned}
 \vec{\nabla} \times \vec{B} &= 0, \quad \vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \quad \vec{\nabla} \times \vec{H} = \vec{J}, \\
 \vec{B} &= \mu_e \vec{H}, \quad \vec{J} = \sigma \left(\vec{E} + \frac{\partial \vec{u}_i}{\partial t} \times \vec{B} \right),
 \end{aligned}
 \tag{7}$$

where \vec{E} is the induced electric field, \vec{J} is the current density vector, and magnetic field \vec{H} includes both primary and induced magnetic fields; μ_e and σ are the induced permeability and conduction coefficient, respectively.

The linearized Maxwell’s stress tensor $(\tau_{ij}^0)^{M_{x_1}}$ due to the magnetic field is given by $(\tau_{ij}^0)^{M_{x_1}} = \mu_e (H_i b_j + H_j b_i - H_k b_k \delta_{ij})$.

Let $\vec{H} = (H_{x_1}, H_{x_2}, H_{x_3})$ and $b_i = (b_1, b_2, b_3)$; b_i is the change in the magnetic field.

In writing the above equations, we have neglected the displacement current.

From Eq. 7, we get

$$\nabla^2 \vec{H} = \mu_e \sigma \left\{ \frac{\partial \vec{H}}{\partial t} + \vec{\nabla} \times \left(\frac{\partial \vec{u}_i}{\partial t} \times \vec{H} \right) \right\}.
 \tag{8}$$

In component form, Eq. 8 can be written as

$$\left. \begin{aligned}
 \frac{\partial H_{x_1}}{\partial t} &= \frac{1}{\mu_e \sigma} \nabla^2 H_{x_1}, \\
 \frac{\partial H_{x_3}}{\partial t} &= \frac{1}{\mu_e \sigma} \nabla^2 H_{x_3}, \quad \text{and} \\
 \frac{\partial H_{x_2}}{\partial t} &= \frac{1}{\mu_e \sigma} \nabla^2 H_{x_2} + \frac{\partial \left(H_{x_1} \frac{\partial v}{\partial t} \right)}{\partial x_1} + \frac{\partial \left(H_{x_3} \frac{\partial v}{\partial t} \right)}{\partial x_3}.
 \end{aligned} \right\}
 \tag{9}$$

Now, for a perfectly conducting medium (i.e., $\sigma \rightarrow \infty$), it can be seen that Eqs. in 9 become

$$\frac{\partial H_{x_1}}{\partial t} = \frac{\partial H_{x_3}}{\partial t} = 0,
 \tag{10}$$

and

$$\frac{\partial H_{x_2}}{\partial t} = \frac{\partial \left(H_{x_1} \frac{\partial v}{\partial t} \right)}{\partial x_1} + \frac{\partial \left(H_{x_3} \frac{\partial v}{\partial t} \right)}{\partial x_3}
 \tag{11}$$

assuming that the primary magnetic field is uniform throughout the space.

It is clear from Eq. 10 that there is no perturbation in H_{x_1} and H_{x_3} , but Eq. 11 shows that there may be perturbation in H_{x_2} . Therefore, taking small perturbation, say b_2 in H_{x_2} , we have $H_{x_1} = H_{01}$, $H_{x_2} = H_{02} + b_2$, and $H_{x_3} = H_{03}$, where H_{01}, H_{02} , and H_{03} are components of the initial magnetic field \vec{H}_o .

We can write: $\vec{H}_o = (H_o \cos \phi, 0, H_o \sin \phi)$, where $H_o = |\vec{H}_o|$ and ϕ is the angle at which the wave crosses the magnetic field. Thus, we have

$$\vec{H} = (H_o \cos \phi, b_2, H_o \sin \phi).
 \tag{12}$$

We shall consider the initial value of b_2 to be zero.

Using Eq. 12 in Eq. 11, we obtain

$$\frac{\partial b_2}{\partial t} = \frac{\partial (H_o \cos \phi \frac{\partial v}{\partial t})}{\partial x_1} + \frac{\partial (H_o \sin \phi \frac{\partial v}{\partial t})}{\partial x_3}.
 \tag{13}$$

Integrating Eq. 13 with respect to t , we get

$$b_2 = H_o \cos \phi \frac{\partial v}{\partial x_1} + H_o \sin \phi \frac{\partial v}{\partial x_3}.
 \tag{14}$$

Considering $\vec{\nabla} \left(\frac{H^2}{2} \right) = -(\vec{\nabla} \times \vec{H}) \times \vec{H} + (\vec{H} \times \vec{\nabla}) \vec{H}$ and Eq. 7, we have

$$\vec{J} \times \vec{B} = \mu_e \left\{ -\vec{\nabla} \left(\frac{H^2}{2} \right) + (\vec{H} \times \vec{\nabla}) \vec{H} \right\}.
 \tag{15}$$

In view of Eqs. 5, 6, and 15, we obtain the only non-vanishing equation of motion for the propagation of SH-wave in transversely isotropic magnetoelastic medium as

$$\begin{aligned}
 \frac{\partial}{\partial x_1} \left\{ N \frac{\partial v}{\partial x_1} \right\} + \frac{\partial}{\partial x_3} \left\{ L \frac{\partial v}{\partial x_3} \right\} + (\mu_e H_o^2 \cos^2 \phi) \frac{\partial^2 v}{\partial x_1^2} \\
 + (\mu_e H_o^2 \sin^2 \phi) \frac{\partial^2 v}{\partial x_3^2} + (\mu_e H_o^2 \sin 2\phi) \frac{\partial^2 v}{\partial x_1 \partial x_3} = \rho \frac{\partial^2 v}{\partial t^2}.
 \end{aligned}
 \tag{16}$$

In view of Biot (1965), equation of motion for propagation of in transversely isotropic magnetoelastic medium under horizontal initial stress (acting along x_1 -direction) is given as

$$\begin{aligned} & \frac{\partial}{\partial x_1} \left[N \frac{\partial v}{\partial x_1} \right] + \frac{\partial}{\partial x_3} \left[L \frac{\partial v}{\partial x_3} \right] \\ & + \left(\frac{P}{2} + \mu_e H_0^2 \cos^2 \varphi \right) \frac{\partial^2 v}{\partial x_1^2} + \mu_e H_0^2 \sin^2 \varphi \frac{\partial^2 v}{\partial x_3^2} \\ & + \mu_e H_0^2 \sin 2\varphi \frac{\partial^2 v}{\partial x_1 \partial x_3} = \rho \frac{\partial^2 v}{\partial t^2}. \end{aligned} \tag{17}$$

The only non-vanishing equations of motion for the propagation of shear wave in the uppermost layer (M_1), intermediate layer (M_2), and lowermost half-space (M_3) as per geometry (Fig. 1) are given by

$$\begin{aligned} & \frac{\partial}{\partial x_1} \left[N_i \frac{\partial v_i}{\partial x_1} \right] + \frac{\partial}{\partial x_3} \left[L_i \frac{\partial v_i}{\partial x_3} \right] + \left(\frac{P_i}{2} + \mu_e^{(i)} H_0^2 \cos^2 \varphi \right) \frac{\partial^2 v_i}{\partial x_1^2} \\ & + \mu_e^{(i)} H_0^2 \sin^2 \varphi \frac{\partial^2 v_i}{\partial x_3^2} + \mu_e^{(i)} H_0^2 \sin 2\varphi \frac{\partial^2 v_i}{\partial x_1 \partial x_3} = \rho_i \frac{\partial^2 v_i}{\partial t^2}, \end{aligned} \tag{18}$$

where the index $i = 1, 2,$ and 3 correspond to the quantities for medium $M_1, M_2,$ and $M_3,$ respectively.

Boundary conditions

The following conditions are identified as boundary conditions of the present problem:

1. The upper surface is stress free ($\tau_{yz}^{(1)} = 0$) at the uppermost surface $x_3 = -H_2$,
2. The displacement components are continuous ($v_1 = v_2$) as well as tractions (stresses) are continuous

($\tau_{yz}^{(1)} = \tau_{yz}^{(2)}$) at the common interface ($x_3 = -H_1$) of uppermost layer and intermediate layer,

3. The displacement components are continuous ($v_2 = v_3$) as well as tractions (stresses) are continuous ($\tau_{yz}^{(2)} = \tau_{yz}^{(3)}$) at the common interface ($x_3 = 0$) of uppermost layer and half-space.

Formulation and solution of the problem

The present model consists of two layers of finite width lying over a half-space (Fig. 2). The uppermost layer (M_1) and intermediate layer (M_2) comprise two distinct heterogeneous magnetoelastic transversely isotropic materials under initial stresses and the lowermost half-space (M_3) is constituted of an initially stressed transversely isotropic material due to gravity. Thicknesses of the uppermost layer and the intermediate layers are $H_2 - H_1$ and H_1 , respectively. Horizontal initial stresses acting in the uppermost and the intermediate layer are P_1 and P_2 , respectively. Initial stress due to gravity in transversely isotropic half-space is given by $P_3 = -\rho_3 g x_3$. Heterogeneity in layers is considered in exponential as well as in linear form, which is discussed in two different cases.

Case-I: when heterogeneity in the layers is of exponential form

In this case, we have considered the vertical heterogeneity in both the layers (M_1 and M_2) in exponential form.

Fig. 1 Geometry of the problem

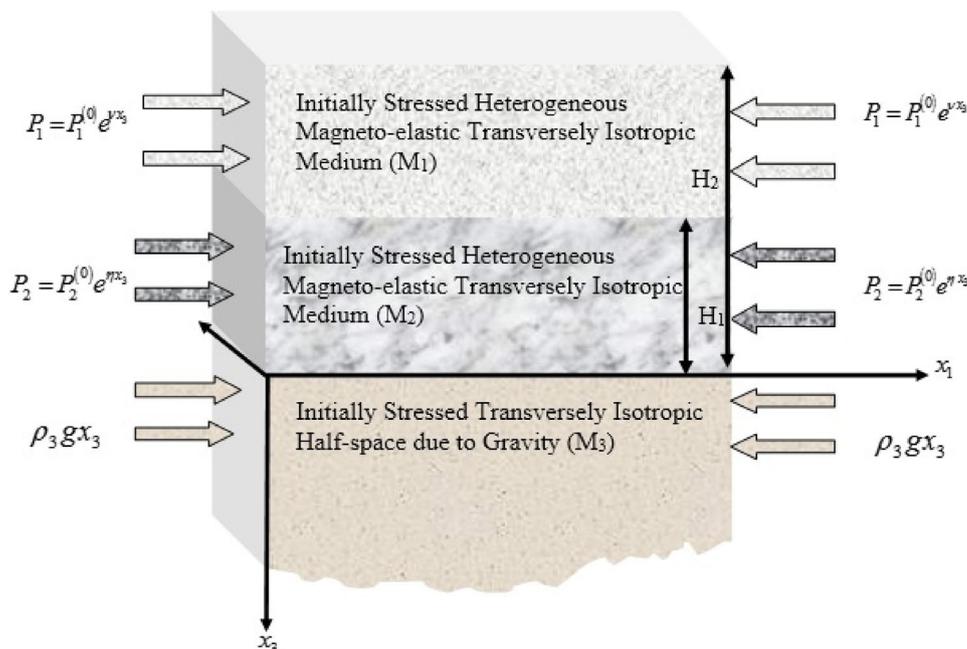
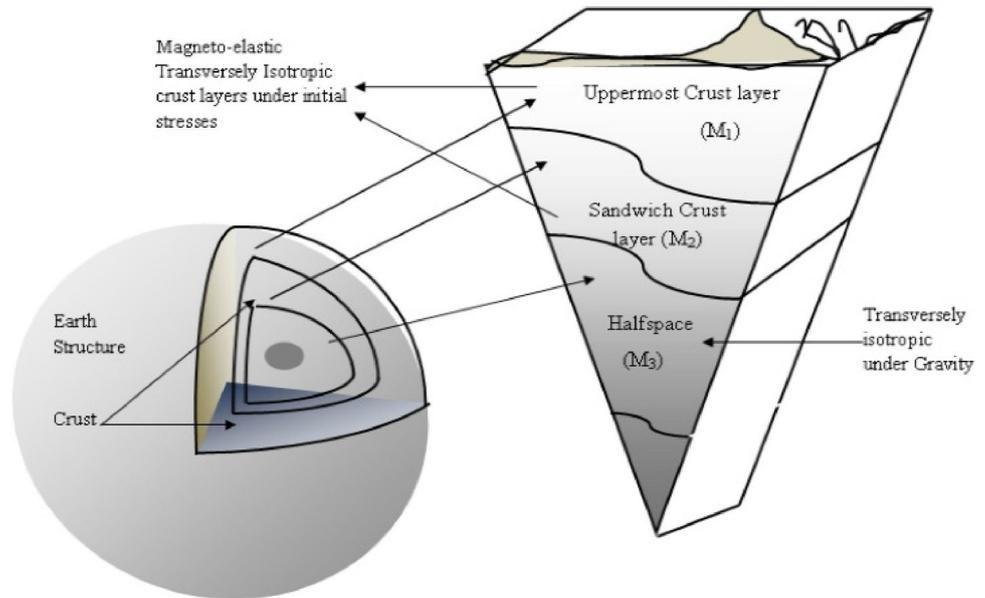


Fig. 2 Layered structure of the problem



Dynamics for uppermost layer (M₁)

Let us consider heterogeneity in the uppermost layer (M₁) as

$$\left. \begin{aligned} N_1 &= N_1^{(0)} e^{vx_3}, \quad L_1 = L_1^{(0)} e^{vx_3}, \quad \rho_1 = \rho_1^{(0)} e^{vx_3}, \\ \mu_e^{(1)} &= \mu_{e1}^{(0)} e^{vx_3}, \quad \text{and} \quad P_1 = P_1^{(0)} e^{vx_3} \end{aligned} \right\} \quad (19)$$

where v is the heterogeneity parameter having dimension inverse of length and $\{N_1^{(0)}, L_1^{(0)}, \rho_1^{(0)}, P_1^{(0)}, \mu_{e1}^{(0)}\}$ are the constants with dimension of density, shear elastic modulus in the transverse direction, shear elastic modulus in longitudinal direction, initial stress and magnetic permeability of the heterogeneous magnetoelastic transversely isotropic upper most layer (M₁).

In view of Eq. 18 for $i = 1$ and Eq. 19, we get the non-vanishing equation of motion for the propagation of shear wave in uppermost layer (M₁) as

$$Q_1 \frac{\partial^2 v_1}{\partial x_1^2} + R_1 \frac{\partial^2 v_1}{\partial x_3^2} + S_1 \frac{\partial^2 v_1}{\partial x_1 \partial x_3} + v \frac{\partial v_1}{\partial x_3} = \frac{1}{\beta_1^2} \frac{\partial^2 v_1}{\partial t^2}, \quad (20)$$

where

$$Q_1 = \frac{N_1^{(0)}}{L_1^{(0)}} + \frac{P_1^{(0)}}{2L_1^{(0)}} + t_H^{(1)} \cos^2 \varphi, \quad R_1 = 1 + t_H^{(1)} \sin^2 \varphi,$$

$$S_1 = t_H^{(1)} \sin 2\varphi, \quad t_H^{(1)} = \frac{\mu_{e1}^{(0)} H_0^2}{L_1^{(0)}}$$

is the transversely isotropic magnetoelastic coupling parameter; $\beta_1 = \sqrt{L_1^{(0)}/\rho_1^{(0)}}$ is the shear wave velocity of medium (M₁) and v_1 is the non-vanishing displacement component for propagation of shear wave in M₁ as per considered geometry (Fig. 1).

Assuming solution of Eq. 20 as $v_1(x_1, x_3, t) = V_1(x_3)e^{ik(x_1-ct)}$ which leads Eq. 20 to the following equation in $V_1(x_3)$

$$R_1 \frac{d^2 V_1(x_3)}{dx_3^2} + (ikS_1 + v) \frac{dV_1(x_3)}{dx_3} + k^2(c^2/\beta_1^2 - Q_1)V_1(x_3) = 0, \quad (21)$$

where k and c are the wave number and common wave velocity, respectively.

Now the solution of Eq. 21 may be given as

$$V_1(x_3) = e^{-(\alpha_1/2)x_3} [A_1 \cos(T_1 x_3) + B_1 \sin(T_1 x_3)], \quad (22)$$

where

$$\alpha_1 = ikS_1 + v \quad \text{and} \quad T_1 = k \sqrt{\frac{S_1^2}{4R_1^2} + \frac{(c^2/\beta_1^2 - Q_1)}{R_1}}.$$

Therefore, the non-vanishing displacement components for the propagation of shear wave in medium (M₁) results in

$$v_1(x_1, x_3, t) = e^{-(\alpha_1/2)x_3} [A_1 \cos(T_1 x_3) + B_1 \sin(T_1 x_3)] e^{ik(x_1-ct)}, \quad (23)$$

where A_1 and B_1 are arbitrary constants.

Dynamics for intermediate layer (M₂)

Let us consider heterogeneity in the intermediate layer (M₂) as

$$\left. \begin{aligned} N_2 &= N_2^{(0)} e^{\eta x_3}, \quad L_2 = L_2^{(0)} e^{\eta x_3}, \quad \rho_2 = \rho_2^{(0)} e^{\eta x_3}, \\ P_2 &= P_2^{(0)} e^{\eta x_3}, \quad \text{and} \quad \mu_e^{(2)} = \mu_{e2}^{(0)} e^{\eta x_3}, \end{aligned} \right\} \quad (24)$$

where η is the heterogeneity parameter having dimension inverse of length and $\{N_2^{(0)}, L_2^{(0)}, \rho_2^{(0)}, P_2^{(0)}, \mu_{e2}^{(0)}\}$ are the constants with dimension of density, shear elastic modulus in the transverse direction, shear elastic modulus in longitudinal direction, initial stress, and magnetic permeability of the heterogeneous magnetoelastic transversely isotropic intermediate layer (M_2).

In view of Eq. 18 for $i = 2$ and Eq. 24, we get the non-vanishing equation of motion for the propagation of shear wave in intermediate layer (M_2) as

$$Q_2 \frac{\partial^2 v_2}{\partial x_1^2} + R_2 \frac{\partial^2 v_2}{\partial x_3^2} + S_2 \frac{\partial^2 v_2}{\partial x_1 \partial x_3} + \eta \frac{\partial v_2}{\partial x_3} = \frac{1}{\beta_2^2} \frac{\partial^2 v_2}{\partial t^2}, \quad (25)$$

where

$$Q_2 = \frac{N_2^{(0)}}{L_2^{(0)}} + \frac{P_2^{(0)}}{2L_2^{(0)}} + t_H^{(2)} \cos^2 \varphi, \quad R_2 = 1 + t_H^{(2)} \sin^2 \varphi, \quad S_2 = t_H^{(2)} \sin 2\varphi, \quad t_H^{(2)} = \frac{\mu_{e2}^{(0)} H_0^2}{L_2^{(0)}}$$

is the transversely isotropic magnetoelastic coupling parameter; $\beta_2 = \sqrt{L_2^{(0)} / \rho_2^{(0)}}$ is the shear wave velocity of medium (M_2) and $v_2(x_1, x_3, t)$ is the non-vanishing displacement component for propagation of shear wave in M_2 as per considered geometry.

Considering the solution of Eq. 25 as $v_2(x_1, x_3, t) = V_2(x_3)e^{ik(x_1-ct)}$ and following the same methodology as that used to solve Eq. 20, we result in the following expression of non-vanishing displacement components for the propagation of horizontally polarized shear wave in intermediate layer (M_2)

$$v_2(x_1, x_3, t) = e^{-(\alpha_2/2)x_3} [A_2 \cos(T_2 x_3) + B_2 \sin(T_2 x_3)] e^{ik(x_1-ct)}, \quad (26)$$

where A_2 and B_2 are arbitrary constants. Parameters α_2 and T_2 are provided in Appendix 1.

Dynamics for half-space (M_3)

The equation of motion for propagation of horizontally polarized shear wave in the transversely isotropic half-space under initial stress due to gravity (Biot 1965) may be written as

$$\left[N_3 - \frac{1}{2} \rho_3 g x_3 \right] \frac{\partial^2 v_3}{\partial x_1^2} + \left[L_3 - \frac{1}{2} \rho_3 g x_3 \right] \frac{\partial^2 v_3}{\partial x_3^2} - \frac{1}{2} \rho_3 g \frac{\partial v_3}{\partial x_3} = \rho_3 \frac{\partial^2 v_1}{\partial t^2}, \quad (27)$$

where $N_3, L_3, \rho_3, v_3,$ and g are the shear elastic modulus in transverse direction, shear elastic modulus in longitudinal directions, the density of the medium, non-vanishing displacement component associated with the half-space, and

acceleration due to gravity (generally, its value on the Earth surface is 9.8 m/s²).

Using $v_3(x_1, x_3, t) = V_3(x_3)e^{ik(x_1-ct)}$ in the above Eq. 27, we have

$$(a_1 + b_1 x_3) \frac{d^2 V_3(x_3)}{dx_3^2} + b_1 \frac{dV_3(x_3)}{dx_3} + k^2 \left(c^2 - \frac{N_3}{\rho_3} - b_1 x_3 \right) V_3(x_3) = 0, \quad (28)$$

where

$$a_1 = \frac{L_3}{\rho_3} \quad \text{and} \quad b_1 = \left(-\frac{g}{2} \right). \quad (29)$$

On simplifying, Eq. 28 takes the form

$$\frac{d^2 V_3(x_3)}{dx_3^2} + \frac{b_1}{(a_1 + b_1 x_3)} \frac{dV_3(x_3)}{dx_3} + k^2 \left(\frac{c_1^2}{a_1 + b_1 x_3} - 1 \right) V_3(x_3) = 0, \quad (30)$$

where

$$c_1^2 = c^2 + a_1(1 - N_3/L_3).$$

Using the transformation $V_3(x_3) = \psi(x_3) / \sqrt{a_1 + b_1 x_3}$ in the above Eq. 30, we have

$$\frac{d^2 \psi(x_3)}{dx_3^2} + \left\{ \frac{b_1^2}{4(a_1 + b_1 x_3)^2} + k^2 \left(\frac{c_1^2}{a_1 + b_1 x_3} - 1 \right) \right\} \psi(x_3) = 0. \quad (31)$$

Considering $\sigma_1 = -\frac{2k}{b_1}(a_1 + b_1 x_3)$ and $s = -\frac{c_1^2 k}{2b_1}$, the above Eq. 31 reduces to Whittaker's equation of the form

$$\frac{d^2 \psi(\sigma_1)}{d\sigma_1^2} + \left[-\frac{1}{4} + \frac{s}{\sigma_1} + \frac{1}{4\sigma_1^2} \right] \psi(\sigma_1) = 0. \quad (32)$$

The solution of Eq. 32 may be written as

$$\psi(\sigma_1) = A_3 W_{-s,0}(-\sigma_1) + B_3 W_{s,0}(\sigma_1), \quad (33)$$

where A_3 and B_3 are arbitrary constants; $W_{s,0}(\sigma_1)$ and $W_{-s,0}(-\sigma_1)$ are Whittaker's functions of first and second kind of order s and 0.

In view of condition $\psi(\sigma_1) \rightarrow 0$ as $x_3 \rightarrow \infty$ the appropriate solution of Eq. 33 becomes

$$\psi(\sigma_1) = A_3 W_{-s,0}(-\sigma_1). \quad (34)$$

Hence, the non-vanishing displacement components for the propagation of horizontally polarized shear wave in half-space (M_3) reduces to

$$v_3(x_1, x_3, t) = A_3 \left(\frac{2a_1 - gx_3}{2} \right)^{-1/2} W_{-s,0} \left[-\left(\frac{4}{G} - 2kx_3 \right) \right] e^{ik(x_1-ct)}, \quad (35)$$

where $G = g/ka_1$ is the Biot’s gravity parameter.

Dispersion equation and particular cases

Using the expressions of non-vanishing displacement components $v_1, v_2,$ and v_3 from Eqs. 23, 26, and 35 in the previously mentioned boundary conditions (“Boundary conditions” section), we results in the following five equations:

$$\left\{ -\frac{\alpha_1}{2} \cos(T_1 H_2) + T_1 \sin(T_1 H_2) \right\} A_1 + \left\{ \frac{\alpha_1}{2} \sin(T_1 H_2) + T_1 \cos(T_1 H_2) \right\} B_1 = 0, \tag{36}$$

$$e^{(\frac{21-22}{2})H_1} \{ A_1 \cos(T_1 H_1) - B_1 \sin(T_1 H_1) \} - \{ A_2 \cos(T_2 H_1) - B_2 \sin(T_2 H_1) \} = 0, \tag{37}$$

$$\frac{L_1^{(0)}}{L_2^{(0)}} e^{\{(\eta-\nu)+(\frac{21-22}{2})\}H_1} \left[\left\{ -\frac{\alpha_1}{2} \cos(T_1 H_1) + T_1 \sin(T_1 H_1) \right\} A_1 + \left\{ \frac{\alpha_1}{2} \sin(T_1 H_1) + T_1 \cos(T_1 H_1) \right\} B_1 \right] = \left[\left\{ -\frac{\alpha_2}{2} \cos(T_2 H_1) + T_2 \sin(T_2 H_1) \right\} A_2 + \left\{ \frac{\alpha_2}{2} \sin(T_2 H_1) + T_2 \cos(T_2 H_1) \right\} B_2 \right], \tag{38}$$

$$A_2 = \frac{A_3}{\sqrt{a_1}} \left\{ W_{-s,0} \left[-\left(\frac{4}{G} - 2kz \right) \right] \right\}_{z=0}, \tag{39}$$

and

$$\left(\frac{L_2^{(0)}}{L_3} \right) \left\{ -\frac{\alpha_2}{2} A_2 + T_2 B_2 \right\} = A_3 \left[\frac{d}{dz} \left\{ \left(\frac{2a_1 - gz}{2} \right)^{-1/2} W_{-s,0} \left[-\left(\frac{4}{G} - 2kz \right) \right] \right\} \right]_{z=0}. \tag{40}$$

Elimination of arbitrary constants $A_1, B_1, A_2, B_2,$ and A_3 from Eqs. 36–40 yields the dispersion equation as where

$$X_1 = \left\{ W_{-s,0} \left[-\left(\frac{4}{G} - 2kx_3 \right) \right] \right\}_{x_3=0} \quad \text{and} \\ X_2 = \left[\frac{d}{dx_3} \left\{ \left(\frac{2a_1 - gx_3}{2} \right)^{-1/2} W_{-s,0} \left[-\left(\frac{4}{G} - 2kx_3 \right) \right] \right\} \right]_{x_3=0}. \tag{42}$$

The asymptotic expansion (Whittaker and Watson 1991) of Whittaker’s function $W_{-s,0} \left[-\left(\frac{4}{G} - 2kx_3 \right) \right]$ for large argument by retaining up to the second term yields

$$W_{-s,0} \left[-\left(\frac{4}{G} - 2kx_3 \right) \right] \sim e^{-(kx_3 - \frac{2}{G})} \left(2kx_3 - \frac{4}{G} \right)^{-s} \times \left\{ 1 - \frac{(s + 0.5)^2}{(2kx_3 - \frac{4}{G})} \right\}. \tag{43}$$

In view of Eq. 43, X_1 and X_2 may be expressed as

$$X_1 = \left\{ W_{-s,0} \left(2kx_3 - \frac{4}{G} \right) \right\}_{x_3=0} \sim e^{\frac{2}{G}} \left(-\frac{4}{G} \right)^{-s} \left[1 + \frac{G}{4} (s + 0.5)^2 \right], \tag{44}$$

and

$$X_2 = \left[\frac{d}{dx_3} \left\{ \left(\frac{2a_1 - gx_3}{2} \right)^{-1/2} W_{-s,0} \left(2kx_3 - \frac{4}{G} \right) \right\} \right]_{x_3=0} \sim \frac{e^{\frac{2}{G}}}{\sqrt{a}} \left(-\frac{4}{G} \right)^{-s} \left[1 + \frac{G}{4} (s + 0.5)^2 \right] \times \left[\left\{ \frac{g}{4a_1} - k + \frac{kGs}{2} \right\} + \frac{kG^2 (s + 0.5)^2}{2 \left[4 + G(s + 0.5)^2 \right]} \right]. \tag{45}$$

$$e^{(\eta-\nu)H_1} \left[\frac{T_2 - \left(\alpha_2/2 + \left(L_3/L_2^{(0)} \right) \sqrt{a_1} \frac{X_2}{X_1} \right) \tan(T_2 H_1)}{1 - \left\{ \frac{\alpha_1/2 - T_1 \tan(T_1 H_2)}{\alpha_1/2 \tan(T_1 H_2) + T_1} \right\} \tan(T_1 H_1)} \right] = \left(L_2^{(0)} / L_1^{(0)} \right) \left[\frac{\{ T_2 \tan(T_2 H_1) - \alpha_2/2 \} T_2 + \left(\alpha_2/2 + \left(L_3/L_2^{(0)} \right) \sqrt{a_1} \frac{X_2}{X_1} \right) \{ T_2 + \alpha_2/2 \tan(T_2 H_1) \}}{T_1 \{ \tan(T_1 H_1) - \tan(T_1 H_2) \}} \right], \tag{41}$$

Now, using Eqs. 44 and 45 in Eq. 41, we get

$$\begin{aligned}
 & e^{(\eta-v)H_1} \left[\frac{1 - \{(\alpha_2/2 + Q) \tan(T_2 H_1)\} / T_2}{1 - \left\{ \frac{(\alpha_1/2) - T_1 \tan(T_1 H_2)}{(\alpha_1/2) \tan(T_1 H_2) + T_1} \right\} \tan(T_1 H_1)} \right] \\
 &= \left(\frac{L_2^{(0)}}{L_1^{(0)}} \right) \left[\frac{\left(T_2 \tan(T_2 H_1) + \left\{ (\alpha_2/2)^2 \tan(T_2 H_1) / T_2 \right\} \right) + Q(1 + \{(\alpha_2/2) \tan(T_2 H_1) / T_2\})}{T_1 \{ \tan(T_1 H_1) - \tan(T_1 H_2) \}} \right], \tag{46}
 \end{aligned}$$

where Q is defined in ‘‘Appendix 1’’.

Equation 46 is the dispersion equation for propagation of horizontally polarized shear wave in magnetoelastic transversely isotropic double layers with exponential heterogeneity under initial stresses lying over a transversely isotropic half-space under gravity.

When both the layers (M_1 and M_2) and the half-space (M_3) are isotropic homogeneous, *i.e.*, $N_1 = L_1 = \mu^{(1)}$, $N_2 = L_2 = \mu^{(2)}$, $N_3 = L_3 = \mu^{(3)}$, $P_1 \rightarrow 0$, $P_2 \rightarrow 0$, $\mu_e^{(1)} \rightarrow 0$, $\mu_e^{(2)} \rightarrow 0$, $v \rightarrow 0$, $\eta \rightarrow 0$, $G \rightarrow 0$ the dispersion Eq. 4 reduces to

$$\tan\{s_1(H_2 - H_1)\} = \frac{s_2}{r_1} \left[\frac{r_2 + \tan(s_2 H_1)}{r_2 \tan(s_2 H_1) - 1} \right], \tag{47}$$

where

$$\begin{aligned}
 r_1 &= \frac{s_1 \mu_1}{\mu_2}, \quad r_2 = -\frac{s_3 \mu_3}{s_2 \mu_2}, \quad s_1 = k \sqrt{\frac{c^2}{\beta_1^2} - 1}, \\
 s_2 &= k \sqrt{\frac{c^2}{\beta_2^2} - 1}, \quad s_3 = k \sqrt{1 - \frac{c^2}{\beta_3^2}}, \\
 \beta_2^2 &= \mu^{(2)} / \rho_2, \quad \text{and} \quad \beta_3^2 = \mu^{(3)} / \rho_3.
 \end{aligned}$$

Equation 47 is the dispersion equation for the propagation of Love-type wave in two homogeneous layers lying over an isotropic homogeneous half-space (Kumari et al. 2015).

In addition to the conditions imposed above, if we impose $H_2 = H_1 = H$, $\mu^{(1)} = \mu^{(2)}$, and $\rho_1 = \rho_2$, the dispersion Eq. 47 further reduces to the classical Love wave equation as

$$\tan \left(kH \sqrt{\frac{c^2}{\beta_2^2} - 1} \right) = \frac{\mu^{(3)} \sqrt{1 - \frac{c^2}{\beta_3^2}}}{\mu^{(2)} \sqrt{\frac{c^2}{\beta_2^2} - 1}}. \tag{48}$$

Case-II: when heterogeneity in the layers is of linear form

In this problem, the uppermost layer (M_1) and intermediate layer (M_2) are considered in the crust part of the Earth with

the material constants as a linear function of space variable pointing vertically downwards as per considered structure (Fig. 2).

Dynamics for uppermost layer (M_1)

Let us consider heterogeneity in the uppermost layer (M_1) as

$$\begin{aligned}
 N_1 &= N_1^{(0)}(1 + ax_3), \quad L_1 = L_1^{(0)}(1 + ax_3), \quad \rho_1 = \rho_1^{(0)}(1 + ax_3), \\
 P_1 &= P_1^{(0)}(1 + ax_3), \quad \text{and} \quad \mu_e^{(1)} = \mu_e^{(0)}(1 + ax_3), \tag{49}
 \end{aligned}$$

where a is the heterogeneity parameter having dimension inverse of length.

In view of Eq. 18 for $i = 1$ and Eq. 49, the non-vanishing equation of motion for the propagation of horizontally polarized shear wave in upper most layer (M_1) is obtained, which on further substitution, $v_1(x_1, x_3, t) = V_1(x_3)e^{ik(x_1 - ct)}$, leads to

$$\frac{d^2 V_1(x_3)}{dx_3^2} + \chi_1(x_3) \frac{dV_1(x_3)}{dx_3} + \zeta_1 V_1(x_3) = 0, \tag{50}$$

where

$$\chi_1(x_3) = \frac{1}{R_1} \left\{ ikS_1 + \frac{a}{(1 + ax_3)} \right\}.$$

Now applying transformation $V_1(x_3) = \Psi_1(x_3) \times \Phi_1(x_3)$ in the above Eq. 50, we have

$$\begin{aligned}
 & \frac{d^2 \Phi_1(x_3)}{dx_3^2} + \left\{ \chi_1(x_3) + \frac{2}{\Psi_1(x_3)} \frac{d\Psi_1(x_3)}{dx_3} \right\} \frac{d\Phi_1(x_3)}{dx_3} \\
 &+ \frac{1}{\Psi_1(x_3)} \left(\frac{d^2 \Psi_1(x_3)}{dx_3^2} + \chi_1(x_3) \frac{d\Psi_1(x_3)}{dx_3} + \zeta_1 \Psi_1(x_3) \right) \Phi_1(x_3) = 0, \tag{51}
 \end{aligned}$$

where $V_1(x_3)$, $\Psi_1(x_3)$, and $\Phi_1(x_3)$ are the differentiable functions.

To remove the first order derivative $d\Phi_1(x_3)/dx_3$ from Eq. 51, we consider

$$\chi_1(x_3) + \frac{2}{\Psi_1(x_3)} \frac{d\Psi_1(x_3)}{dx_3} = 0, \tag{52}$$

where

$$\Psi_1(x_3) = e^{-1/2 \int \chi_1(x_3) dx_3} = (1 + ax_3)^{-(1/2R_1)} e^{-(h_1 x_3/2)}. \tag{53}$$

Now using the value of $\Psi_1(x_3)$ from Eq. 53 in Eq. 51, we get

$$\frac{d^2 \Phi_1(x_3)}{dx_3^2} + \left(\frac{\{a/(1 + ax_3)\}^2}{2R_1} - \frac{1}{4} \left(h_1 + \frac{a/(1 + ax_3)}{R_1} \right)^2 + \xi_1 \right) \Phi_1(x_3) = 0. \tag{54}$$

On simplification, the above Eq. 54 can be written as

$$\frac{d^2 \Phi_1(t_1)}{dt_1^2} + \left(-\frac{1}{4} + \frac{k_1}{t_1} + \frac{1/4 - \mu_1^2}{t_1^2} \right) \Phi_1(t_1) = 0, \tag{55}$$

where

$$t_1 = (1 + ax_3)q_1 \text{ with } q_1 = \frac{\sqrt{h_1^2 - 4\xi_1}}{a}.$$

Equation 55 is in the form of Whittaker’s equation and its solution can be written as

$$\Phi_1(t_1) = C_1 W_{k_1, \mu_1}(t_1) + D_1 W_{-k_1, \mu_1}(-t_1), \tag{56}$$

where C_1 and D_1 are arbitrary constants; k_1 , μ_1 , ξ_1 , and h_1 are defined in “Appendix 2”.

With the help of Eqs. 53 and 56, the solution of Eq. 50 can be written as

$$V_1(x_3) = (1 + ax_3)^{-(1/2R_1)} e^{-\frac{h_1 x_3}{2}} \times [C_1 W_{k_1, \mu_1}(t_1) + D_1 W_{-k_1, \mu_1}(-t_1)], \tag{57}$$

where $W_{k_1, \mu_1}(t_1)$ and $W_{-k_1, \mu_1}(-t_1)$ are the Whittaker’s functions of first and second kinds of order (k_1, μ_1) , respectively.

Therefore, non-vanishing displacement component for propagation of horizontally polarized shear wave in medium M_1 can be written as

$$v_1(x_1, x_3, t) = (1 + ax_3)^{-1/2R_1} e^{-\frac{h_1 x_3}{2}} [C_1 W_{k_1, \mu_1}(t_1) + D_1 W_{-k_1, \mu_1}(-t_1)] e^{ik(x_1 - ct)}. \tag{58}$$

Dynamics for Intermediate layer (M₂)

Let us consider heterogeneity in intermediate layer (M_2) as

$$\left. \begin{aligned} N_2 &= N_2^{(0)}(1 + bx_3), \quad L_2 = L_2^{(0)}(1 + bx_3), \quad \rho_2 = \rho_2^{(0)}(1 + bx_3), \\ P_2 &= P_2^{(0)}(1 + bx_3), \quad \text{and } \mu_e^{(2)} = \mu_e^{(0)}(1 + bx_3), \end{aligned} \right\} \tag{59}$$

where b is the heterogeneity parameter having dimension inverse of length.

In view of Eq. 18 for $i = 2$ and Eq. 59, the non-vanishing equation of motion for the propagation of horizontally polarized shear wave in intermediate layer (M_2) is obtained, which on further substitution $v_2(x_1, x_3, t) = V_2(x_3)e^{ik(x_1 - ct)}$ and application of the same transformation used in “Case-II to obtained the dynamics for uppermost layer M_1 ” will lead towards the non-vanishing displacement component for propagation of horizontally polarized shear wave in medium M_2 as

$$v_2(x_1, x_3, t) = (1 + bx_3)^{-(1/2R_2)} e^{-\frac{h_2 x_3}{2}} \times [C_2 W_{k_2, \mu_2}(t_2) + D_2 W_{-k_2, \mu_2}(-t_2)] e^{ik(x_1 - ct)}, \tag{60}$$

where C_2 and D_2 are arbitrary constants; $t_2 = (1 + bx_3)q_2$ with $q_2 = \frac{\sqrt{h_2^2 - 4\xi_2}}{b}$, $W_{k_2, \mu_2}(t_2)$, and $W_{-k_2, \mu_2}(-t_2)$ are the Whittaker’s functions of first and second kinds of order (k_2, μ_2) , respectively. Moreover, newly introduced symbols, ξ_2 , h_2 , $\chi_2(x_3)$, μ_2 , and k_2 , are defined in “Appendix 2”.

Dynamics for half-space (M₃)

Then non-vanishing displacement component for the propagation of horizontally polarized shear wave in half-space (M_3) is given by Eq. 35.

Dispersion equation and particular cases

Now using the asymptotic expansion (Whittaker and Watson 1991) of Whittaker’s function for large argument and retaining up to the second term, the following approximations may be made:

$$\left. \begin{aligned} W_{k_1, \mu_1}(t_1) &\sim e^{-\frac{1}{2}(t_1)} (t_1)^{k_1} \left[1 + \frac{m_1}{t_1} \right], \\ W_{-k_1, \mu_1}(-t_1) &\sim e^{\frac{1}{2}(t_1)} (-t_1)^{-k_1} \left[1 - \frac{m_2}{t_1} \right], \\ W_{k_2, \mu_2}(t_2) &\sim e^{-\frac{1}{2}(t_2)} (t_2)^{k_2} \left[1 + \frac{m_3}{t_2} \right], \\ W_{-k_2, \mu_2}(-t_2) &\sim e^{\frac{1}{2}(t_2)} (-t_2)^{-k_2} \left[1 - \frac{m_4}{t_2} \right], \\ \text{and } W_{-s,0} \left(2kx_3 - \frac{4}{G} \right) &\sim e^{-(kx_3 - \frac{2}{G})} \\ &\times \left(2kx_3 - \frac{4}{G} \right)^{-s} \left[1 - \frac{(s + 0.5)^2}{\left(2kx_3 - \frac{4}{G} \right)} \right]. \end{aligned} \right\} \tag{61}$$

where $m_i (i = 1$ to 4) is defined in “Appendix 2”.

In view of approximation provided in Eq. 61, the non-vanishing displacement components given in Eqs. 58, 60, and 35 reduce to

$$v_1(x_1, x_3, t) \sim (1 + ax_3)^{-1/2} R_1 e^{-\frac{h_1 x_3}{2}} \left[C_1 e^{-\frac{i}{2}(t_1)^{k_1}} (1 + m_1/t_1) + D_1 e^{\frac{i}{2}(-t_1)^{-k_1}} (1 - m_2/t_1) \right] e^{ik(x_1 - ct)}, \tag{62}$$

$$v_2(x_1, x_3, t) \sim (1 + bx_3)^{-1/2} R_2 e^{-\frac{h_2 x_3}{2}} \left[C_2 e^{-\frac{i}{2}(t_2)^{k_2}} (1 + m_3/t_2) + D_2 e^{\frac{i}{2}(-t_2)^{-k_2}} (1 - m_4/t_2) \right] e^{ik(x_1 - ct)}, \tag{63}$$

and

$$v_3(x_1, x_3, t) \sim A_3 \left(\frac{2a_1 - gx_3}{2} \right)^{-1/2} e^{-(kx_3 - \frac{t}{G})} \left(2kx_3 - \frac{4}{G} \right)^{-s} \left[1 - \frac{(s + 0.5)^2}{(2kx_3 - \frac{4}{G})} \right] e^{ik(x_1 - ct)}. \tag{64}$$

Using the expressions of non-vanishing displacement components v_1 , v_2 , and v_3 from Eqs. 62–64 in the previously mentioned boundary conditions (Sect. 3), we result in the following five equations:

$$f_1 C_1 + f_2 D_1 = 0, \tag{65}$$

$$f_3 C_1 + f_4 D_1 = f_5 C_2 + f_6 D_2, \tag{66}$$

$$f_7 C_1 + f_8 D_1 = f_9 C_2 + f_{10} D_2, \tag{67}$$

$$f_{11} C_2 + f_{12} D_2 = A_3 f_{13}, \tag{68}$$

$$f_{14} C_2 + f_{15} D_2 = A_3 f_{16}. \tag{69}$$

Eliminating arbitrary constants C_1 , D_1 , C_2 , D_2 , and A_3 from the above Eqs. 65–69, we get the dispersion equation as

$$\frac{f_3 \left(-\frac{f_2}{f_1} \right) + f_4}{f_7 \left(-\frac{f_2}{f_1} \right) + f_8} = \frac{f_5 \left(\frac{f_{15}/f_{16} - f_{12}/f_{13}}{f_{11}/f_{13} - f_{14}/f_{16}} \right) + f_6}{f_9 \left(\frac{f_{15}/f_{16} - f_{12}/f_{13}}{f_{11}/f_{13} - f_{14}/f_{16}} \right) + f_{10}}, \tag{70}$$

where f_i ($i = 1$ to 16) is defined in ‘‘Appendix 2’’.

Equation 70 is the dispersion equation of horizontally polarized shear wave propagating in the linear heterogeneous magnetoelastic, transversely isotropic media under initial stresses lying over a transversely isotropic half-space under gravity.

When both the layers (M_1 and M_2) and the half-space (M_3) are isotropic homogeneous, *i.e.*, $N_1 = L_1 = \mu^{(1)}$, $N_2 = L_2 = \mu^{(2)}$, $N_3 = L_3 = \mu^{(3)}$, $P_1 \rightarrow 0$, $P_2 \rightarrow 0$, $\mu_e^{(1)} \rightarrow 0$, $\mu_e^{(2)} \rightarrow 0$, $v \rightarrow 0$, $\eta \rightarrow 0$, $G \rightarrow 0$, the dispersion Eq. 70 reduces to Eq. 47. If in addition to the above, we further impose $H_2 = H_1 = H$, $\mu^{(1)} = \mu^{(2)}$, and $\rho_1 = \rho_2$, the classical Love wave Eq. 48 is obtained.

Numerical results and discussions

To unravel the effects of dimensionless affecting parameters on the propagation of horizontally polarized shear wave in double initially stressed heterogeneous anisotropic (magnetoelastic transversely isotropic) layers overlying a transversely isotropic gravitational half-space, numerical computation of dispersion Eqs. 46 and 70 has been carried out. For the computational purpose, we consider the following data of elastic coefficients for equivalent transversely isotropic models of the upper mantle (olivine model and petrofabric model) from Anderson (1989):

- (i) For uppermost crustal layer (M_1)

$$L_1^{(0)} = 0.659 \text{ Mbar}, \quad N_1^{(0)} = 0.824 \text{ Mbar}, \quad \rho_1^{(0)} = 3.298 \text{ gm/cm}^3;$$

- (ii) For intermediate crustal layer (M_2)

$$L_2^{(0)} = 0.723 \text{ Mbar}, \quad N_2^{(0)} = 0.623 \text{ Mbar}, \quad \rho_2^{(0)} = 3.298 \text{ gm/cm}^3;$$

- (iii) For lowermost half-space (M_3)

$$L_3 = 0.770 \text{ Mbar}, \quad N_3 = 0.784 \text{ Mbar}, \quad \rho_3^{(0)} = 3.305 \text{ gm/cm}^3.$$

Moreover, the following data have also been considered:

$$\left\{ P_1^{(0)}/2L_1^{(0)}, P_2^{(0)}/2L_2^{(0)} \right\} = 0.1, 0.2, 0.3, 0.4;$$

$$\left\{ t_H^{(1)}, t_H^{(2)} \right\} = 0.1, 0.3, 0.5, 0.7; \quad G = 0.1, 0.3, 0.5, 0.7;$$

$$\{ vH_1, aH_1 \} = 0.05, 0.1, 0.15, 0.2;$$

$$\{ \eta H_1, bH_1 \} = 0.50, 0.51, 0.52, 0.53 \text{ and } \phi = \pi/18.$$

The dispersion curves delineating the variation of dimensionless phase velocity (c/β_1) against dimensionless wave number (kH_1) have been plotted in Figs. 3, 4, 5, 6, 7, 8, 9 when both heterogeneity parameter of uppermost layer (M_1) and heterogeneity parameter of intermediate layer (M_2) are supposed to vary exponentially with depth (case-I) and are also supposed to vary linearly with depth (case-II). The substantial influence of various affecting parameters viz. heterogeneity parameters (vH_1 and ηH_1) of the uppermost layer and intermediate layer associated with case-I, heterogeneity parameter (aH_1 and bH_1) of the uppermost layer and intermediate layer corresponding to case-II, magnetoelastic coupling parameters $t_H^{(1)}$ and $t_H^{(2)}$ of layers M_1 and M_2 , respectively, horizontal compressive initial stresses $P_1^{(0)}/2L_1^{(0)}$ and $P_2^{(0)}/2L_2^{(0)}$ of layers M_1 and

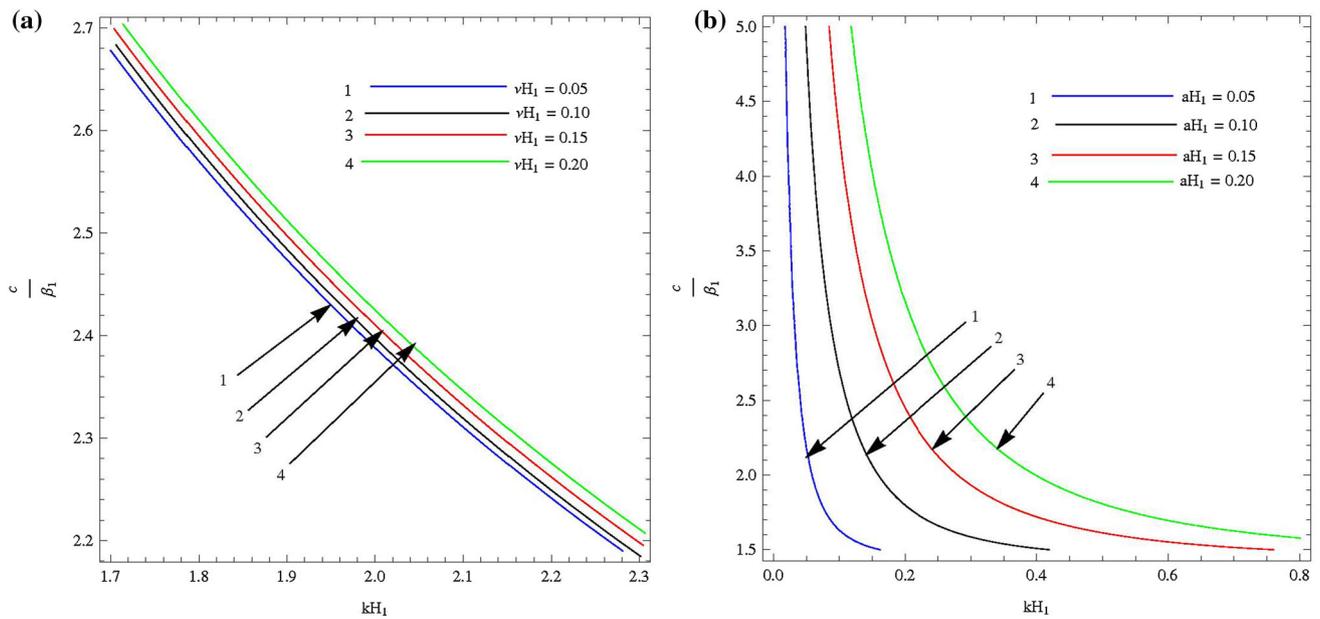


Fig. 3 Variation in dimensionless phase velocity (c/β_1) against dimensionless wave number (kH_1) for different values of **a** heterogeneity parameter (vH_1) of uppermost layer M_1 (case-I), **b** heterogeneity parameter (aH_1) of uppermost layer M_1 (case-II)

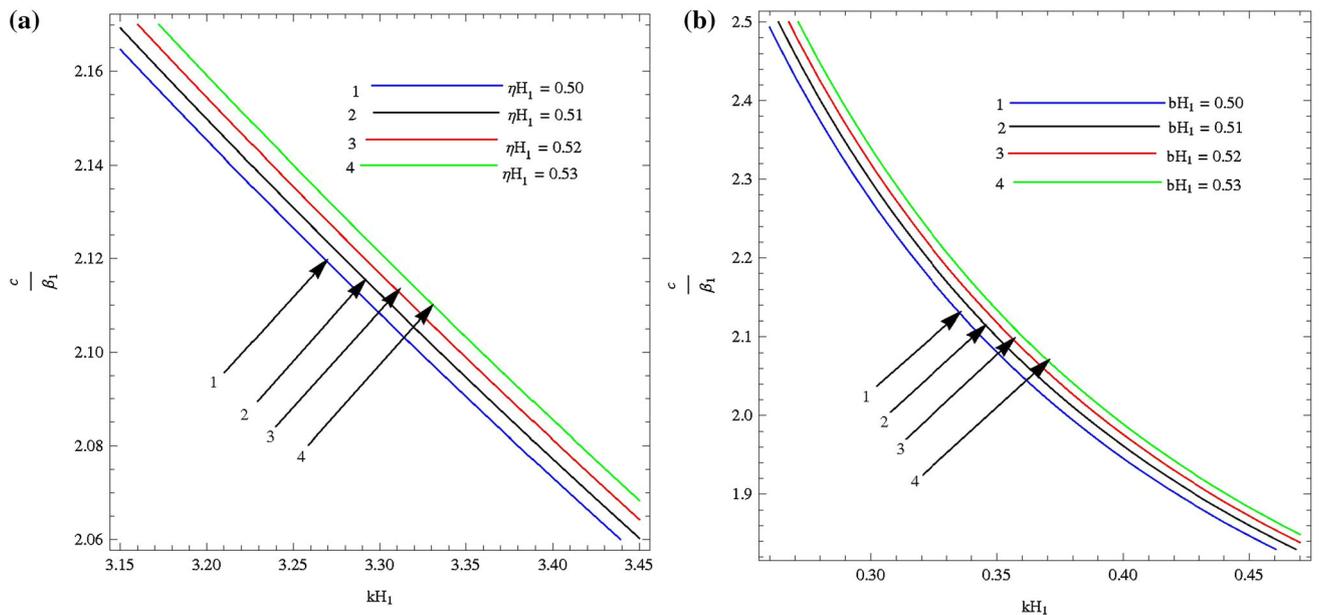


Fig. 4 Variation in dimensionless phase velocity (c/β_1) against dimensionless wave number (kH_1) for different values of (a) heterogeneity parameter (ηH_1) of intermediate layer M_2 (case-I), (b) heterogeneity parameter (bH_1) of intermediate layer M_2 (case-II)

M_2 , respectively, and Biot’s gravity parameter (G) of the lowest half-space (M_3) on the phase velocity of horizontally polarized shear waves have been traced out by means of numerical simulations and graphical illustrations. It is evident from all the figures (Figs. 3, 4, 5, 6, 7, 8, 9) that phase velocity of shear waves diminishes with an increase in wave number and finally approaches its minimum value for a particular wave number.

Figure 3a depicts the effect of heterogeneity parameter (vH_1) of the uppermost layer (M_1) for case-I, whereas Fig. 3b illustrates the impact of heterogeneity parameter (aH_1) of the uppermost layer (M_1) for case-II when both the magnetoelastic layers M_1 and M_2 are initially stressed and the lowest half-space (M_3) is under the action of gravity. From these two figures (Fig. 3a, b) it is concluded that as the heterogeneity parameters vH_1 and aH_1 steepen

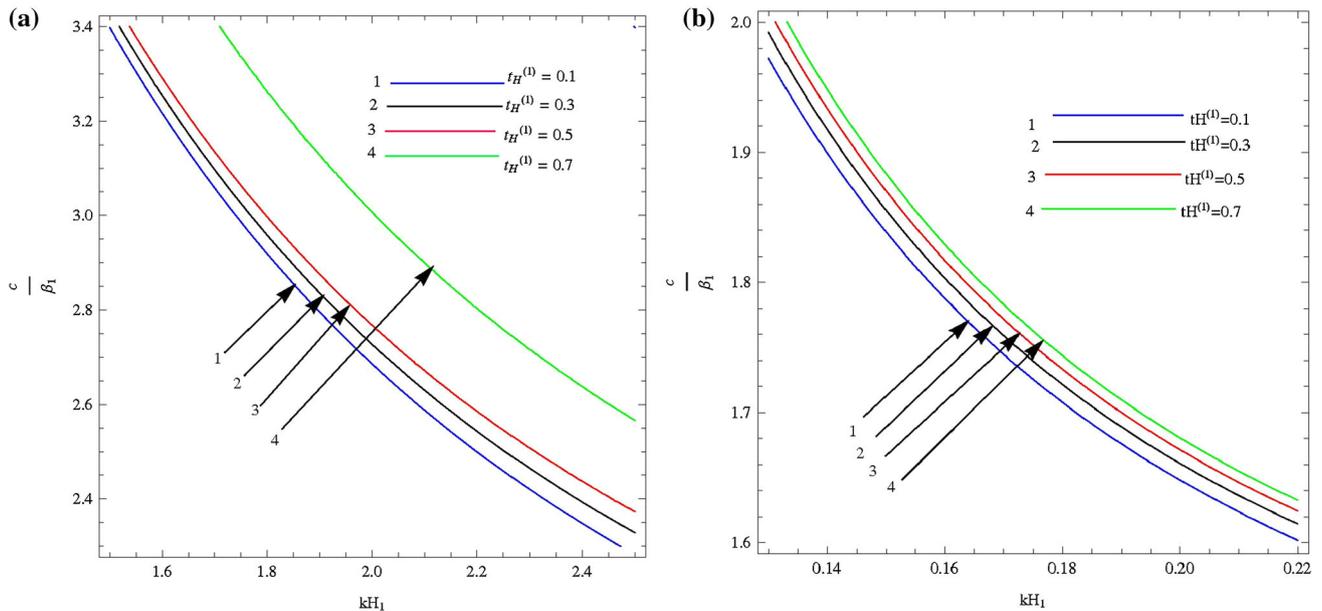


Fig. 5 Variation in dimensionless phase velocity (c/β_1) against dimensionless wave number (kH_1) for different values of **a** magnetoelastic coupling parameter ($t_H^{(1)}$) of uppermost layer M_1 (case-I), **b** magnetoelastic coupling parameter ($t_H^{(1)}$) of uppermost layer M_1 (case-II)

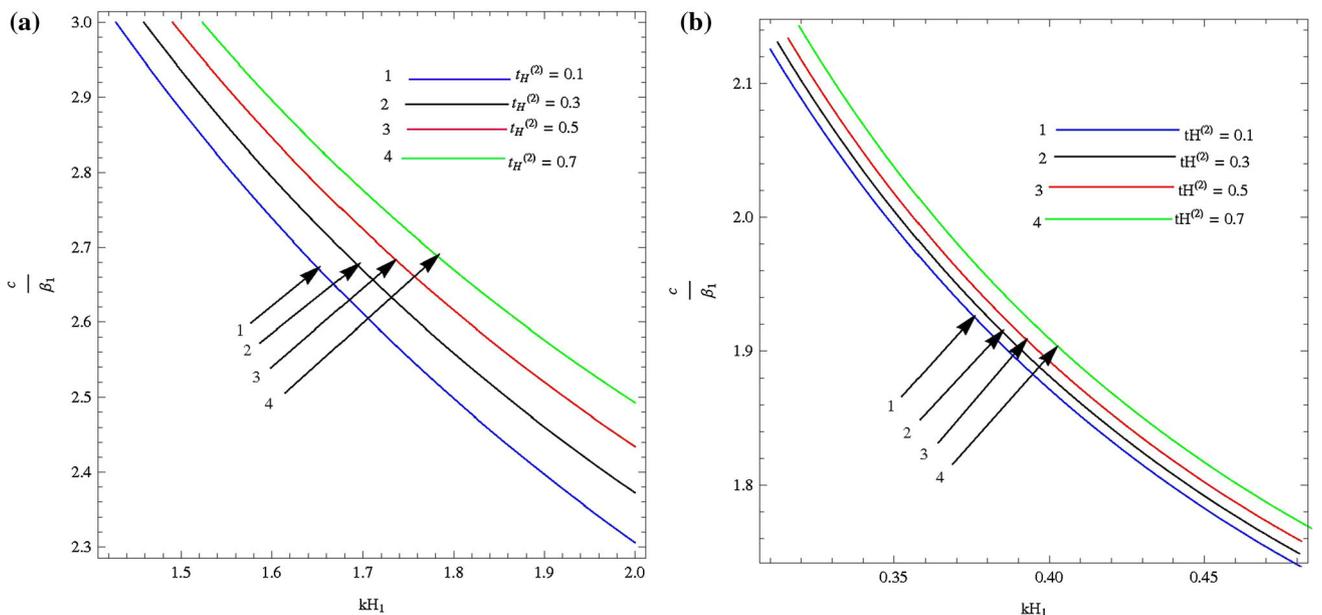


Fig. 6 Variation in dimensionless phase velocity (c/β_1) against dimensionless wave number (kH_1) for different values of **a** magnetoelastic coupling parameter ($t_H^{(2)}$) of intermediate layer M_2 (case-I), **b** magnetoelastic coupling parameter ($t_H^{(2)}$) of intermediate layer M_2 (case-II)

in the uppermost layer, phase velocity of shear wave also steps up, but the heterogeneity parameter (νH_1) caused due to exponential variation in the uppermost layer (case-I) favors more the phase velocity of shear wave as compared to the heterogeneity parameter (aH_1) caused due to linear variation in the uppermost layer (case-II). In a similar way, the influence of heterogeneity parameter (ηH_1) of intermediate layer for case-I has been delineated in Fig. 4a

while the significant effect of heterogeneity parameter (bH_1) of intermediate layer for case-II has been outlined in Fig. 4b. Both figures suggest that phase velocity of shear wave steps down with a rise in heterogeneity parameters ηH_1 and bH_1 in the intermediate layer. It is also noticed that the heterogeneity parameter (ηH_1) caused due to exponential variation in the intermediate layer (case-I) has a more favorable impact on the phase velocity of shear

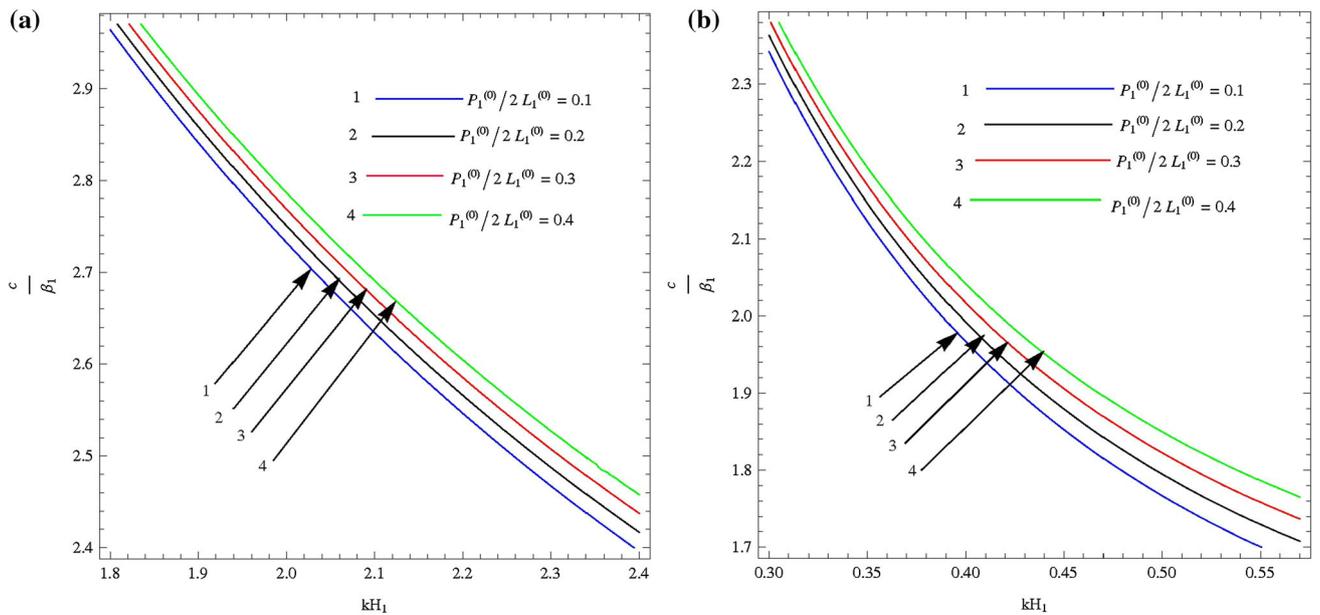


Fig. 7 Variation in dimensionless phase velocity (c/β_1) against dimensionless wave number (kH_1) for different values of **a** horizontal compressive initial stress ($P_1^{(0)}/2L_1^{(0)}$) of uppermost layer M_1 (case-I),

b horizontal compressive initial stress ($P_1^{(0)}/2L_1^{(0)}$) of uppermost layer M_1 (case-II)

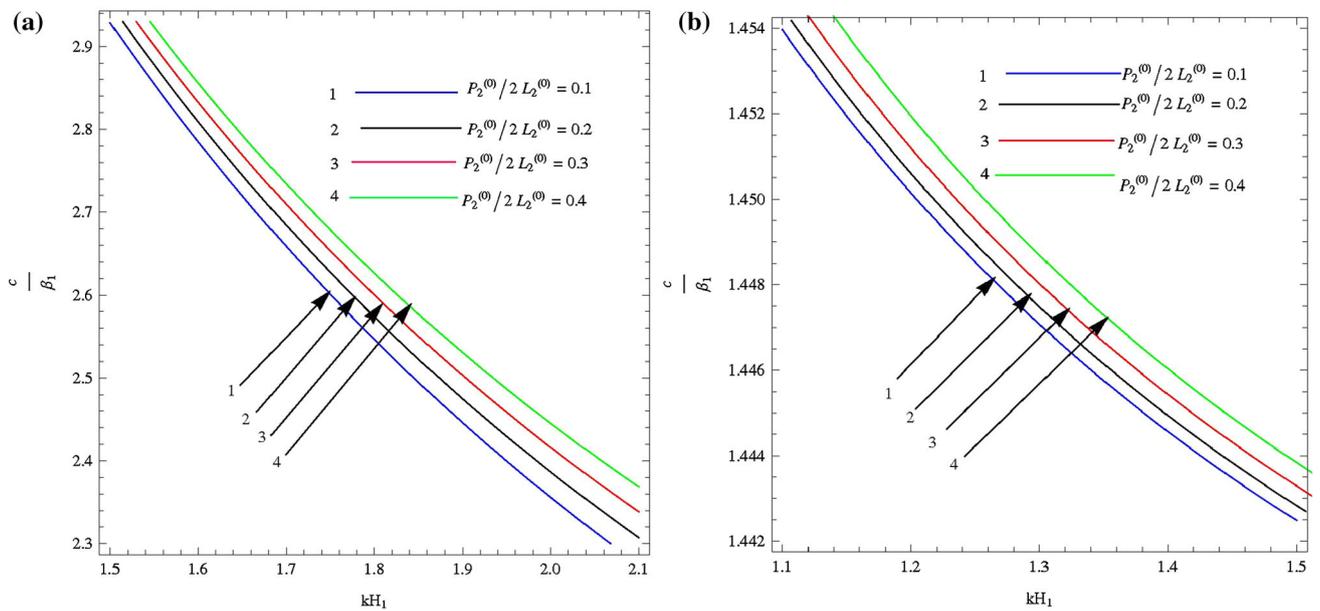


Fig. 8 Variation in dimensionless phase velocity (c/β_1) against dimensionless wave number (kH_1) for different values of **a** horizontal compressive initial stress ($P_2^{(0)}/2L_2^{(0)}$) of intermediate layer M_2 (case-

I), **b** horizontal compressive initial stress ($P_2^{(0)}/2L_2^{(0)}$) of intermediate layer M_2 (case-II)

wave as compared to the heterogeneity parameter (bH_1) caused due to linear variation in the intermediate layer (case-II).

As a further study, Fig. 5 demonstrates the impact of magnetoelastic coupling parameter $t_H^{(1)}$ of uppermost layer for case-I and case-II, respectively, while Fig. 6 presents

the influence of magnetoelastic coupling parameter $t_H^{(2)}$ of intermediate layer for case-I and case-II, respectively. It is manifested from these four figures that as magnetoelastic coupling parameter surmounts in both the heterogeneous layers M_1 and M_2 , phase velocity of shear wave also gets increased. It may be due to the fact that as the value of

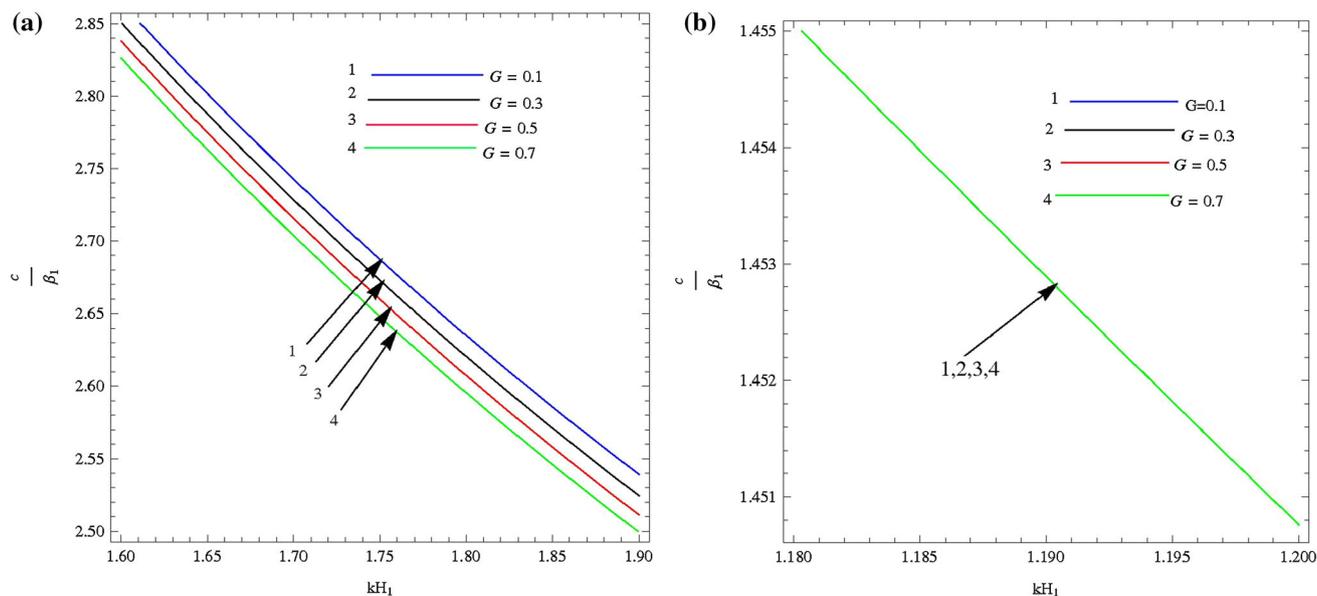


Fig. 9 Variation in dimensionless phase velocity (c/β_1) against dimensionless wave number (kH_1) for different values of **a** Biot's gravity parameter (G) of lowermost half-space M_3 (case-I), **b** Biot's gravity parameter (G) of lowermost half-space M_3 (case-II)

these magnetoelastic coupling parameter in the magnetoelastic transversely isotropic medium increase, the molecules of the medium get more attracted toward each other which, in turn, lessen the void in the medium and hence steepen the phase velocity of shear wave. Moreover, the graphical delineation studying the effect of horizontal compressive initial stress $P_1^{(0)}/2L_1^{(0)}$ acting on the uppermost layer for both the cases, i.e., case-I and case-II, has been remarkably traced out in Fig. 7, respectively, while the impact of horizontal compressive initial stress $P_2^{(0)}/2L_2^{(0)}$ acting on the intermediate layer for case-I and case-II has been studied in Fig. 8, respectively. These four panels (Figs. 7, 8) reveal that both the initial stresses, i.e., horizontal compressive initial stresses $P_1^{(0)}/2L_1^{(0)}$ and $P_2^{(0)}/2L_2^{(0)}$ acting on the uppermost and intermediate heterogeneous magnetoelastic layers, respectively, favor the phase velocity of shear wave. The reason behind this nature of variation may lie in the fact that as these horizontal compressive initial stresses increase, the geomaterials may become more compact which further enhances the phase velocity of shear wave propagating in the medium. Furthermore, the effect of Biot's gravity parameter for case-I has been reflected in Fig. 9a whereas for case-II it has been depicted in Fig. 9b. Figure 9a signifies that phase velocity of shear wave falls down with an increase in Biot's gravity parameter. A subtle examination of Fig. 9b leads to the fact that there is a negligible effect of Biot's gravity parameter on the phase velocity of shear wave.

The comparative study of all the figures (Figs. 3, 4, 5, 6, 7, 8, 9) establishes that the phase velocity of shear wave is

found to be predominant in the case when exponential heterogeneity prevails in both the magnetoelastic layers (case-I) than the case when linear heterogeneity prevails in both the magnetoelastic layers (case-II).

Conclusions

The present work investigates the propagation behavior of horizontally polarized shear waves in double initially stressed heterogeneous magnetoelastic transversely isotropic layers overlying a transversely isotropic gravitating half-space. The heterogeneities in both the initially stressed heterogeneous magnetoelastic transversely isotropic layers are due to exponential variation (case-I) and linear variation (case-II) in the material constants with respect to the space variable pointing positively downwards. The closed form solutions for the displacement components of layers and half-space have been derived separately. The dispersion relations have been obtained in closed form for both the cases (exponential heterogeneity and linear heterogeneity) by using Whittaker's asymptotic expansion. Also the dispersion equations of shear waves have been obtained for the case when the two layers along with the half-space are isotropic and homogeneous. Moreover, the comparative studies have been made to analyze the effect of different type of heterogeneity (in crust layers) on the dispersion curve. To observe the profound effects of the various affecting parameters on the dispersion curves, numerical computation and graphical demonstration have been

carried out. The following outcomes can be outlined as the salient features of the study:

- Phase velocity of shear wave falls down with an increase in wave number and finally settles down for a specific wave number in both exponential and linear case of heterogeneity.
- It has been observed that as heterogeneity prevails in the uppermost layer in either cases (exponential heterogeneity and linear heterogeneity), the phase velocity of shear wave gets increased, whereas as heterogeneity surmounts in the intermediate layer in either cases, the phase velocity of shear wave diminishes.
- It is worth mentioning that magnetoelastic coupling parameters of both the layers favor the phase velocity of shear wave in both exponential and linear case of heterogeneity.
- It can be pointed out that phase velocity of shear wave steps up with an increment in the horizontal compressive initial stresses in both the layers in either cases of heterogeneity.
- It can be remarkably quoted that Biot’s gravity parameter of the lowermost half-space disfavors the phase velocity of shear wave.
- For shear wave propagating in the assumed geometry, the following conditions must be satisfied:
 $\beta_1 < \beta_2 < c < \beta_3$ or $\beta_2 < \beta_1 < c < \beta_3$.
- The above condition signifies that phase velocity of shear wave must be less than the shear wave velocity in the lowermost half-space and greater than the shear wave velocity in both the layers.
- Comparative study suggests that phase velocity of shear wave is found to be predominant in the case when exponential heterogeneity prevails in both the magnetoelastic layers (case-I) than the case when linear heterogeneity prevails in both the magnetoelastic layers (case-II).
- Deduced dispersion equations for both the cases of heterogeneity are in well-agreement to the classical Love wave equation.

Acknowledgements The authors convey their sincere thanks to Indian School of Mines, Dhanbad for providing JRF to Ms. Nirmala Kumari and also facilitating us with its best facility. Authors extend their sincere thanks to Mr. Santan Kumar for his valuable suggestion in development of the manuscript.

Appendix 1

$$\alpha_2 = ikS_2 + \eta, \quad T_2 = k\sqrt{\frac{S_2^2}{4R_2^2} + \frac{(c^2/\beta_2^2 - Q_2)}{R_2}},$$

$$Q = \left(L_3/L_2^{(0)}\right) \left\{ \frac{g}{4a_1} + k \left(\frac{Gs}{2} - 1 + \frac{G^2(s+0.5)^2}{2[4+G(s+0.5)^2]} \right) \right\}.$$

Appendix 2

$$h_1 = \frac{ik t_H^{(1)} \sin 2\varphi_1}{\left(1 + t_H^{(1)} \sin^2 \varphi_1\right)},$$

$$\xi_1 = \frac{k^2}{R_1} \left[c^2/\beta_1^2 - \left(\frac{N_1^{(0)}}{L_1^{(0)}} + \frac{P_1^{(0)}}{2L_1^{(0)}} + t_H^{(1)} \cos^2 \varphi \right) \right],$$

$$k_1 = -\frac{h_1/2}{R_1 \sqrt{h_1^2 - 4\xi_1}}, \quad \mu_1 = \frac{1}{4} \times \sqrt{1 + 1/R_1^2 - 2/R_1},$$

$$h_2 = \frac{ik \{t_H^{(2)} \sin 2\varphi\}}{R_2},$$

$$\chi_2(z) = \frac{1}{R_2} \left\{ ik \left(t_H^{(2)} \sin 2\varphi \right) + \frac{b}{(1+bz)} \right\},$$

$$\xi_2 = \frac{k^2}{R_2} \left[c^2/\beta_2^2 - \left(\frac{N_2^{(0)}}{L_2^{(0)}} + \frac{P_2^{(0)}}{2L_2^{(0)}} + t_H^{(2)} \cos^2 \varphi \right) \right],$$

$$k_2 = -\frac{h_2/2}{R_2 \sqrt{h_2^2 - 4\xi_2}}, \quad \mu_2 = \frac{1}{4} \times \sqrt{1 + 1/R_2^2 - 2/R_2},$$

$$m_1 = \left\{ \mu_1^2 - (k_1 - 1/2)^2 \right\}, \quad m_2 = \left\{ \mu_1^2 - (k_1 + 1/2)^2 \right\},$$

$$m_3 = \left\{ \mu_2^2 - (k_2 - 1/2)^2 \right\}, \quad m_4 = \left\{ \mu_2^2 - (k_2 + 1/2)^2 \right\},$$

$$f_1 = \frac{(1 - aH_2)^{k_1+1-1/2R_1}}{q_1^{-k_1}} e^{\frac{1}{2}\{(h_1+aq_1)H_2-q_1\}} L_1^{(0)}$$

$$\times \left[\left(\frac{ikS_1}{2} - \frac{R_1 h_1}{2} - \frac{a}{2(1-aH_2)} \right) \left(1 + \frac{m_1}{(1-aH_2)q_1} \right) \right.$$

$$+ aq_1 R_1 \left\{ \left(\frac{k_1}{(1-aH_2)q_1} - \frac{1}{2} \right) \left(1 + \frac{m_1}{(1-aH_2)q_1} \right) \right.$$

$$\left. \left. - \frac{m_1}{(1-aH_2)^2 q_1^2} \right\} \right],$$

$$\begin{aligned}
f_2 &= \frac{(1 - aH_2)^{1-k_1-1/2R_1}}{[-q_1]^{k_1}} e^{\frac{1}{2}\{(h_1-aq_1)H_2+q_1\}} L_1^{(0)} \times \left[\left(\frac{ikS_1}{2} - \frac{R_1h_1}{2} - \frac{a}{2(1-aH_2)} \right) \left(1 - \frac{m_2}{(1-aH_2)q_1} \right) \right. \\
&\quad \left. + aq_1R_1 \left\{ \left(\frac{-k_1}{(1-aH_2)q_1} + \frac{1}{2} \right) \left(1 - \frac{m_2}{(1-aH_2)q_1} \right) + \frac{m_2}{(1-aH_2)^2q_1^2} \right\} \right], \\
f_3 &= \frac{(1 - aH_1)^{k_1-1/2R_1}}{q_1^{-k_1}} e^{\frac{1}{2}\{(h_1+aq_1)H_1-q_1\}} \left(1 + \frac{m_1}{(1-aH_1)q_1} \right), \\
f_4 &= \frac{(1 - aH_1)^{-(k_1+1/2R_1)}}{[-q_1]^{k_1}} e^{\frac{1}{2}\{(h_1-aq_1)H_1+q_1\}} \left(1 - \frac{m_2}{(1-aH_1)q_1} \right), \\
f_5 &= \frac{(1 - bH_1)^{k_2-1/2R_2}}{[q_2]^{-k_2}} e^{\frac{1}{2}\{(h_2+bq_2)H_1-q_2\}} \left(1 + \frac{m_3}{(1-bH_1)q_2} \right), \\
f_6 &= \frac{(1 - bH_1)^{-(k_2+1/2R_2)}}{[-q_2]^{k_2}} e^{\frac{1}{2}\{(h_2-bq_2)H_1+q_2\}} \left(1 - \frac{m_4}{(1-bH_1)q_2} \right), \\
f_7 &= \frac{(1 - aH_1)^{1+k_1-1/2R_1}}{[q_1]^{-k_1}} e^{\frac{1}{2}\{(h_1+aq_1)H_1-q_1\}} L_1^{(0)} \times \left[\left(\frac{ikS_1}{2} - \frac{R_1h_1}{2} - \frac{a}{2(1-aH_1)} \right) \left(1 + \frac{m_1}{(1-aH_1)q_1} \right) \right. \\
&\quad \left. + aq_1R_1 \left\{ \left(\frac{k_1}{(1-aH_1)q_1} - \frac{1}{2} \right) \left(1 + \frac{m_1}{(1-aH_1)q_1} \right) - \frac{m_1}{(1-aH_1)^2q_1^2} \right\} \right], \\
f_8 &= \frac{(1 - aH_1)^{1-(k_1+1/2R_1)}}{[-q_1]^{k_1}} e^{\frac{1}{2}\{(h_1-aq_1)H_1+q_1\}} L_1^{(0)} \times \left[\left(\frac{ikS_1}{2} - \frac{R_1h_1}{2} - \frac{a}{2(1-aH_1)} \right) \left(1 - \frac{m_2}{(1-aH_1)q_1} \right) \right. \\
&\quad \left. + aq_1R_1 \left\{ \left(\frac{-k_1}{(1-aH_1)q_1} + \frac{1}{2} \right) \left(1 - \frac{m_2}{(1-aH_1)q_1} \right) + \frac{m_2}{(1-aH_1)^2q_1^2} \right\} \right], \\
f_9 &= \frac{(1 - bH_1)^{1+k_2-1/2R_2}}{[q_2]^{-k_2}} e^{\frac{1}{2}\{(h_2+bq_2)H_1-q_2\}} L_2^{(0)} \times \left[\left(\frac{ikS_2}{2} - \frac{R_2h_2}{2} - \frac{b}{2(1-bH_1)} \right) \left(1 + \frac{m_3}{(1-bH_1)q_2} \right) \right. \\
&\quad \left. + bq_2R_2 \left\{ \left(\frac{k_2}{(1-bH_1)q_2} - \frac{1}{2} \right) \left(1 + \frac{m_3}{(1-bH_1)q_2} \right) - \frac{m_3}{(1-bH_1)^2q_2^2} \right\} \right], \\
f_{10} &= \frac{(1 - bH_1)^{1-(k_2+1/2R_2)}}{[-q_2]^{k_2}} e^{\frac{1}{2}\{(h_2-bq_2)H_1+q_2\}} L_2^{(0)} \times \left[\left(\frac{ikS_2}{2} - \frac{R_2h_2}{2} - \frac{b}{2(1-bH_1)} \right) \left(1 - \frac{m_4}{(1-bH_1)q_2} \right) \right. \\
&\quad \left. + bq_2R_2 \left\{ \left(\frac{-k_2}{(1-bH_1)q_2} + \frac{1}{2} \right) \left(1 - \frac{m_4}{(1-bH_1)q_2} \right) + \frac{m_4}{(1-bH_1)^2q_2^2} \right\} \right], \\
f_{11} &= \frac{e^{-q_2/2}}{(q_2)^{-k_2}} \left(1 + \frac{m_3}{q_2} \right), \quad f_{12} = \frac{e^{q_2/2}}{(-q_2)^{k_2}} \left(1 - \frac{m_4}{q_2} \right), \quad f_{13} = \frac{e^{\frac{2}{\sigma}}}{\sqrt{a_1}} \left[-\frac{4}{G} \right]^{-s} \left(1 + \frac{G(s+0.5)^2}{4} \right), \\
f_{14} &= L_2^{(0)} e^{-\frac{q_2}{2}} (q_2)^{k_2} \left[\left(\frac{ikS_2}{2} - \frac{R_2h_2}{2} - \frac{b}{2} \right) \left(1 + \frac{m_3}{q_2} \right) + bq_2R_2 \left\{ \left(\frac{k_2}{q_2} - \frac{1}{2} \right) \left(1 + \frac{m_3}{q_2} \right) - \frac{m_3}{q_2^2} \right\} \right], \\
f_{15} &= L_2^{(0)} e^{\frac{q_2}{2}} (-q_2)^{-k_2} \left[\left(\frac{ikS_2}{2} - \frac{R_2h_2}{2} - \frac{b}{2} \right) \left(1 - \frac{m_4}{q_2} \right) + bq_2R_2 \left\{ \left(\frac{-k_2}{q_2} + \frac{1}{2} \right) \left(1 - \frac{m_4}{q_2} \right) + \frac{m_4}{q_2^2} \right\} \right], \\
f_{16} &= L_3 \sqrt{a_1} e^{\frac{2}{\sigma}} \left[-\frac{4}{G} \right]^{-s} \left[\frac{g}{4a_1} - k \left(1 - \frac{sG}{2} \right) \left(1 + \frac{G(s+0.5)^2}{4} \right) + \frac{kG^2(s+0.5)^2}{8} \right].
\end{aligned}$$

References

- Abd-Alla AM, Ahmad SM (1999) Propagation of Love waves in a non-homogeneous orthotropic elastic layer under initial stress overlying semi-infinite medium. *Appl Math Comput* 106(2–3):265–275. doi:[10.1016/S0096-3003\(98\)10128-5](https://doi.org/10.1016/S0096-3003(98)10128-5)
- Acharya DP, Roy I, Sengupta S (2009) Effect of magnetic field and initial stress on the propagation of interface waves in transversely isotropic perfectly conducting media. *Acta Mech* 202(1–4):35–45. doi:[10.1007/s00707-008-0027-5](https://doi.org/10.1007/s00707-008-0027-5)
- Anderson DL (1989) *Theory of the Earth*. Blackwell Scientific Publ, Boston
- Backus GE (1962) Long-wave elastic anisotropy produced by horizontal layering. *J Geophys Res* 67(11):4427–4440. doi:[10.1029/JZ067i011p04427](https://doi.org/10.1029/JZ067i011p04427)
- Bhattacharya J (1962) On the dispersion curve for Love wave due to irregularity in the thickness of the transversely isotropic crustal layer. *Gerl Beitr Geophys* 6:324–334
- Bieńkowski A, Szewczyk R, Salach J (2010) Industrial application of magnetoelastic force and torque sensors. *Acta Phys Pol A* 118(5):1008–1009
- Biot MA (1965) *Mechanics of incremental deformations*. Wiley, New York
- Chattopadhyay A (1975) On the propagation of Love type waves in an intermediate non-homogeneous layers lying between two semi-infinite homogeneous elastic media. *Gerl Beitr Geophys* 84(3/4):327–334
- Chattopadhyay A, Choudhury S (1990) Propagation, reflection and transmission of magnetoelastic shear waves in a self-reinforced medium. *Int J Eng Sci* 28(6):485–495. doi:[10.1016/0020-7225\(90\)90051-J](https://doi.org/10.1016/0020-7225(90)90051-J)
- Chattopadhyay A, Kar BK (1977a) On the propagation of Love-type waves in a non-homogeneous internal stratum of finite thickness lying between two semi-infinite isotropic media. *Gerl Beitr Geophys* 86(5):407–412
- Chattopadhyay A, Kar BK (1977b) On the dispersion curves of Love type waves in an internal stratum of finite thickness under initial stress lying between two semi-infinite isotropic elastic media. *Gerl Beitr Geophys* 86:493–497
- Chattopadhyay A, Singh AK (2014) Propagation of a crack due to magneto-elastic shear waves in a self-reinforced medium. *J Vib Cont* 20(3):406–420. doi:[10.1177/1077546312458134](https://doi.org/10.1177/1077546312458134)
- Chattopadhyay A, Gupta S, Sharma VK, Kumari P (2009) Propagation of shear waves in viscoelastic medium at irregular boundaries. *Acta Geophys* 58(2):195–214. doi:[10.2478/s11600-009-0060-3](https://doi.org/10.2478/s11600-009-0060-3)
- Chattopadhyay A, Gupta S, Sharma VK, Kumari P (2010) Effect of point source and heterogeneity on the propagation of SH waves. *Int J Appl Math Mech* 6(9):76–89
- Dey S, Addy SK (1978) Love waves under initial stresses. *Acta Geophys Pol* 26(1):47–54
- Ding G, Dravinski M (1996) Scattering of in multi-layered media with irregular interfaces. *Earthq Eng Struct Dyn* 25(12):1391–1404. doi:[10.1002/\(SICI\)1096-9845\(199612\)25:12<1391::AID-EQE617>3.0.CO;2-W](https://doi.org/10.1002/(SICI)1096-9845(199612)25:12<1391::AID-EQE617>3.0.CO;2-W)
- Dutta S (1963) Love wave in a non-homogeneous internal stratum lying between two semi-infinite isotropic media. *Geophysics* 28(2):156–160. doi:[10.1190/1.1439162](https://doi.org/10.1190/1.1439162)
- Gogna ML (1976) On Love waves in heterogeneous layered media. *Geophys J Roy Astr Soc* 45(2):357–370. doi:[10.1111/j.1365-246X.1976.tb00331.x](https://doi.org/10.1111/j.1365-246X.1976.tb00331.x)
- Guz AN (2002) Elastic waves in bodies with initial (residual) stresses. *Int Appl Mech* 38(1):23–59. doi:[10.1023/A:1015379824503](https://doi.org/10.1023/A:1015379824503)
- Kumari N, Sahu SA, Chattopadhyay A, Singh AK (2015) Influence of heterogeneity on the propagation behavior of Love-type waves in a layered isotropic structure. *Int J Geomech* 16(2):04015062. doi:[10.1061/\(ASCE\)GM.1943-5622.0000541](https://doi.org/10.1061/(ASCE)GM.1943-5622.0000541)
- Sahu SA, Saroj PK, Paswan B (2014) Shear waves in a heterogeneous fiber-reinforced layer over a half-space under gravity. *Int J Geomech* 15(2):04014048. doi:[10.1061/\(ASCE\)GM.1943-5622.0000404](https://doi.org/10.1061/(ASCE)GM.1943-5622.0000404)
- Sur SP (1963) Note on stresses produced by a shearing force moving over the boundary of a semi-infinite transversely isotropic solid. *Pure Appl Geophys* 55(1):72–76. doi:[10.1007/BF02011217](https://doi.org/10.1007/BF02011217)
- Whittaker ET, Watson GN (1991) *A Course of Modern Analysis*. Universal Book Stall, New Delhi

Ground deformation effects from the ~M6 earthquakes (2014–2015) on Cephalonia–Ithaca Islands (Western Greece) deduced by GPS observations

Vassilis Sakkas¹  · Evangelos Lagios¹

Received: 24 January 2017 / Accepted: 7 February 2017 / Published online: 2 March 2017
© Institute of Geophysics, Polish Academy of Sciences & Polish Academy of Sciences 2017

Abstract The implications of the earthquakes that took place in the central Ionian Islands in 2014 (Cephalonia, M_w 6.1, M_w 5.9) and 2015 (Lefkas, M_w 6.4) are described based on repeat measurements of the local GPS networks in Cephalonia and Ithaca, and the available continuous GPS stations in the broader area. The Lefkas earthquake occurred on a branch of the Cephalonia Transform Fault, affecting Cephalonia with SE displacements gradually decreasing from north (~ 100 mm) to south (~ 10 mm). This earthquake revealed a near N–S dislocation boundary separating Paliki Peninsula in western Cephalonia from the rest of the island, as well as another NW–SE trending fault that separates kinematically the northern and southern parts of Paliki. Strain field calculations during the interseismic period (2014–2015) indicate compression between Ithaca and Cephalonia, while extension appears during the following co-seismic period (2015–2016) including the 2015 Lefkas earthquake. Additional tectonically active zones with differential kinematic characteristics were also identified locally.

Keywords Tectonic motions · Ground deformation · GPS measurements · Cephalonia · Ithaca Islands · Ionian Sea

Introduction

During the last few years, an increased seismic activity has taken place in the central Ionian Islands (Fig. 1), one of the most tectonically active areas of Europe. It is an area with major seismogenic zones that includes the cephalonia transform fault (CTF), the dominant faulting system in the area, as well as several other faulting features which combine to generate the largest seismic energy release in Europe (e.g., Le Pichon et al. 1995).

The broader area exhibits a complex kinematic state that changes from north to south. The collision between Apulian microplate and Eurasia in the north, the right lateral CTF in the central part and the oceanic subduction in the south creates a complicated stress field that affects both regional and local ground deformation, which makes difficult the understanding of the ongoing local tectonic activity in the area (e.g., van Hinsbergen et al. 2006; Vassilakis et al. 2011). Velocities increase systematically from north to south (e.g., Ganas et al. 2013). In addition, a clockwise rotational pattern in the Ionian region complicates the kinematics even more (e.g., Lagios et al. 2007; Chousianitis et al. 2015).

Large-magnitude earthquakes (M_w 5.9 and M_w 6.1) occurred in the western part of Cephalonia (Paliki Peninsula) in late January–early February 2014 (e.g., Papadopoulos et al. 2015; Karastathis et al. 2015; Karakostas et al. 2015), causing severe damage in this part of the island (e.g., Lekkas and Mavroulis 2015). After the occurrence of these earthquakes, a long seismic sequence followed (see red and yellow colored epicenters in Fig. 2) that lasted several months in 2014. These epicenters were mainly located along a ruptured NNE–SSW trending faulting zone in Paliki, as modeled from local GPS data (Sakkas and

✉ Vassilis Sakkas
vsakkas@geol.uoa.gr

¹ Laboratory of Geophysics, Department of Geology and Geoenvironment, National and Kapodistrian University of Athens, Panepistimiopolis, Ilissia, 157 84 Athens, Greece

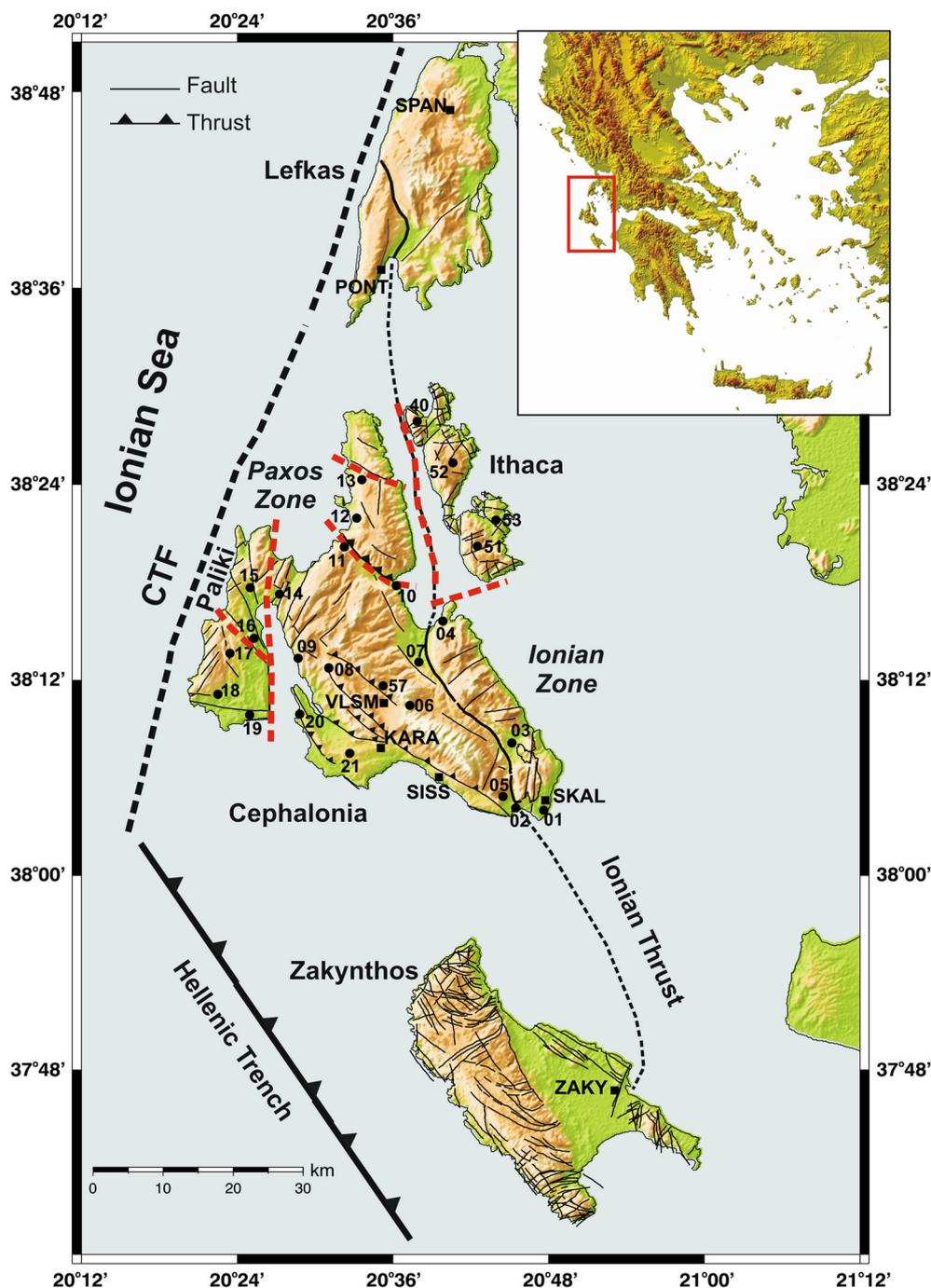


Fig. 1 Map of the broader area of Central Ionian Islands, showing the main faulting and thrusting zones, together with the distribution of GPS benchmarks (circles) and continuous GPS stations (rectangles). CTF: Cephalonia Transform Fault. Red dashed lines indicate the

approximate locations of inferred major kinematic boundaries deduced from the GPS observations as described in the text. Datum: WGS'84

Lagios 2015) and radar interferometric data (Briole et al. 2015).

An M_w 6.4 earthquake took place on November 17, 2015 in the southern part of Lefkas Island (Chousianitis et al. 2016; Ganas et al. 2016) just north of Cephalonia (Fig. 2). The earthquake was followed by a significant seismic

activity that also extended southwards into the sea to the west of the northern peninsula of Cephalonia (see blue epicenters in Fig. 2). This southern extension of activity covered the same area in the sea as that after the occurrence of an earlier M_w 6.3 earthquake in August 2003, which was located further to the north than the 2015 event, and just

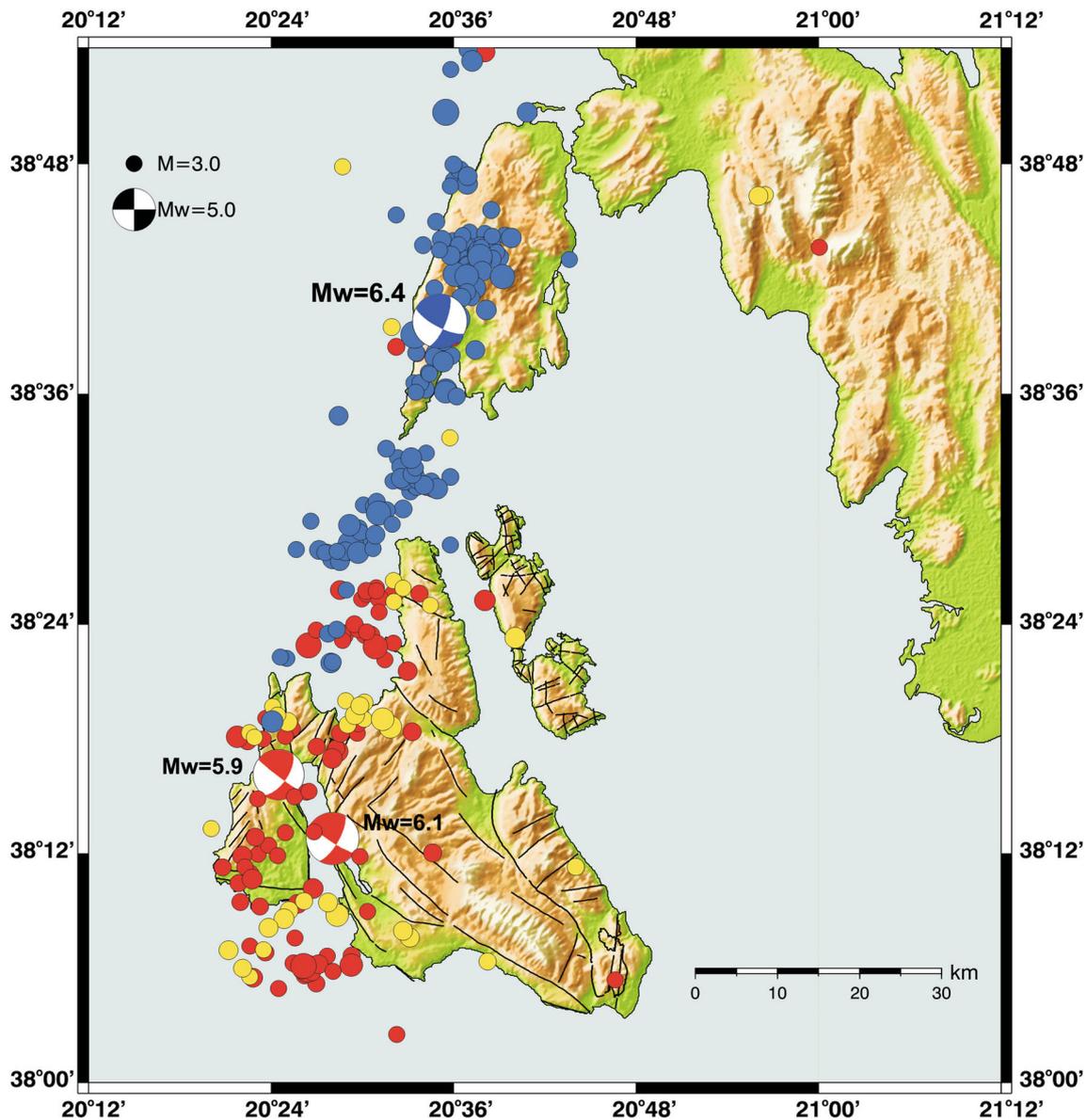


Fig. 2 Distribution of earthquake epicenters ($M > 3$) in Cephalonia, Ithaca and Lefkas islands in the periods April 2014 to March 2015 (red circles), March 2015 to November 16, 2015 (yellow circles), and November 17, 2015 to February 2016 (blue circles) showing also the focal mechanisms of the earthquakes in Cephalonia ($M_w = 5.9$ and

$M_w = 6.1$) in early 2014, and in Lefkas ($M_w = 6.4$) on November 17, 2015. The relocated epicenters are from the NOA (Greece) earthquake catalog, and the focal mechanisms are from the Seismological Laboratory (NKUA)

offshore of the western coast of Lefkas Island (e.g., Karakostas et al. 2004; Papadimitriou et al. 2006; Lagios et al. 2007; Saltogianni and Stiros 2015).

The present paper provides a detailed description of the ground deformation in Cephalonia after the intense seismic activity in 2014, up to early 2016, including also the tectonic implications of the 2015 Lefkas earthquake on the broader area. These accounts are based on the repeat measurement of our Cephalonia–Ithaca GPS benchmarks, and the available continuous GPS stations in the broader area.

GPS measurements

An extensive GPS network consisting of 27 stations (benchmarks) was installed on Cephalonia in 2001 and expanded to Ithaca in 2004 (Fig. 1). The locations of the stations were selected to better describe and study the kinematics of the recognized fault-blocks of these islands. This network was remeasured several times up to 2010 (Lagios et al. 2007, 2012), as well as in 2014, after the early 2014 earthquake sequence that occurred in

Cephalonia (Sakkas and Lagios 2015). The network was subsequently remeasured in early March 2015, and also in February 2016, after the recent event in the southern part of Lefkas, which occurred just 20 km north of Cephalonia. Data from continuous GPS (cGPS) recording stations located in the broader area were also processed and included in the present analysis.

Time series for the period January 2014 to February 2016 were compiled from four cGPS stations in Cephalonia—(1) Valsamata (VLISM), (2) Karavados (KARA), (3) Sissia (SISS), and (4) Skala (SKAL)—as well as two stations in Lefkas, (1) Spanochori (SPAN) and (2) Ponti (PONT) (Fig. 1). The Cephalonia–Ithaca benchmark network was occupied using Leica receivers of types SR399, SR9500 and GRX1200. For each daily session, the occupied stations were tied to all the Cephalonia-based cGPS stations that served as local reference stations. The motion at VLISM is well known (Table 1), as this station has been operating since 2005. Each roving station was occupied at least twice at a sampling rate of 15 s, and at a nominal duration time of about 48–56 h for each occupational period.

All the available GPS data were homogeneously processed using the Bernese Software version 5.2 (Dach et al. 2015). In addition to the regional cGPS network, data from seventeen IGS, EUREF, and cGPS stations in Greece (<http://www.epncb.oma.be>) were included in the processing. These additional sites were used to define the local reference frame to calculate the coordinates of the stations, and eventually the horizontal and vertical displacements with respect to IGB08 reference frame (<http://igsceb.jpl.nasa.gov/network/refframe.html>).

The GPS data were evaluated using a double-difference approach based on daily sessions. The modeled observable was the ionosphere-free linear combination. Ambiguities were solved using the quasi-ionosphere-free (QIF) strategy. For each station, hourly zenith path delays and daily gradient parameters were estimated. The Neill mapping function was used. Ocean tide loading corrections (<http://holt.oso.chalmers.se/loading>) according to model FES2004 and absolute antenna phase center corrections were

applied. Precise orbits, earth orientation parameters (EOP), satellite clock corrections and GPS monthly P1–C1 code bias solutions were obtained from the CODE analysis center (<ftp://unibe.ch/aiub/CODE>), grids of the Vienna Mapping Function (<http://ggosatm.hg.tuwien.ac.at>), and atmospheric coefficients (<http://geophy.uni.lu/>) were used to calculate a set of high-precision coordinates for the stations. The long-time span of the data significantly reduced uncertainties (at a 90% confidence level) for the majority of the stations to overall errors of about 1.4–7.1 and 2.2–8.1 mm for the horizontal and vertical solutions, respectively. Since the observations at each station last for several days, each day was treated as a separate 24 h long session. The errors, expressed as standard deviation, quantify the amount of variation of each daily session coordinate solution from the mean value.

The variation of the coordinates in the IGB08 reference frame for the cGPS stations in the area is shown in Figs. 3 and 4. Until the occurrence of the seismic events in Cephalonia (2014) and Lefkas (2015), these stations exhibited the anticipated regional motion of the area with respect to IGB08 (Ganas et al. 2013), which represents an NNE horizontal motion with small variations from north to south (Table 1). The seismic events in early 2014 caused a WSW horizontal motion accompanied by subsidence only at the station located in Cephalonia, while the 2015 event in Lefkas affected a much broader area extending as far as Zakynthos Island to the south, and AGRI station in the western part of Central Greece (Fig. 4). At the VLISM station in Cephalonia, the overall directional displacement during the early 2014 earthquakes was ~ 28 mm in the WSW horizontal direction, and about -11 mm in the vertical component (Sakkas and Lagios 2015); however, the rest of the central Ionian Islands were not affected. The 2015 Lefkas event caused a relatively large SW horizontal displacement at the cGPS stations in Lefkas (~ 424 and ~ 89 mm in its southern and northern parts, respectively; Fig. 4a), and southward displacements in Cephalonia and Zakynthos, decreasing from ~ 17 mm in the central part of Cephalonia (VLISM) to ~ 10 mm in its southern part (SKAL), and ~ 6 mm in Zakynthos (ZAKY) (Fig. 4b).

Table 1 Continuous GPS station velocities with respect to IGB08

| Station | Period | V_{N-S} (mm/year) | V_{E-W} (mm/year) | V_{Up} (mm/year) |
|---------|-----------------|---------------------|---------------------|--------------------|
| VLISM | 01/2009–12/2013 | 4.25 ± 0.02 | 17.10 ± 0.02 | -0.63 ± 0.05 |
| | 01/2014–11/2015 | 1.4 ± 0.1 | 17.1 ± 0.2 | 1.6 ± 0.4 |
| KARA | 05/2014–02/2016 | 2.1 ± 0.2 | 16.8 ± 0.2 | 2.5 ± 0.7 |
| SISS | 03/2014–02/2015 | 2.2 ± 0.3 | 15.0 ± 0.4 | 6.5 ± 1.2 |
| SKAL | 03/2015–02/2016 | -4.5 ± 0.5 | 15.4 ± 0.4 | -8.5 ± 2.0 |
| SPAN | 01/2009–11/2015 | 3.83 ± 0.02 | 20.86 ± 0.02 | -0.5 ± 0.1 |
| PONT | 01/2009–11/2015 | 6.81 ± 0.02 | 20.44 ± 0.03 | -0.7 ± 0.1 |

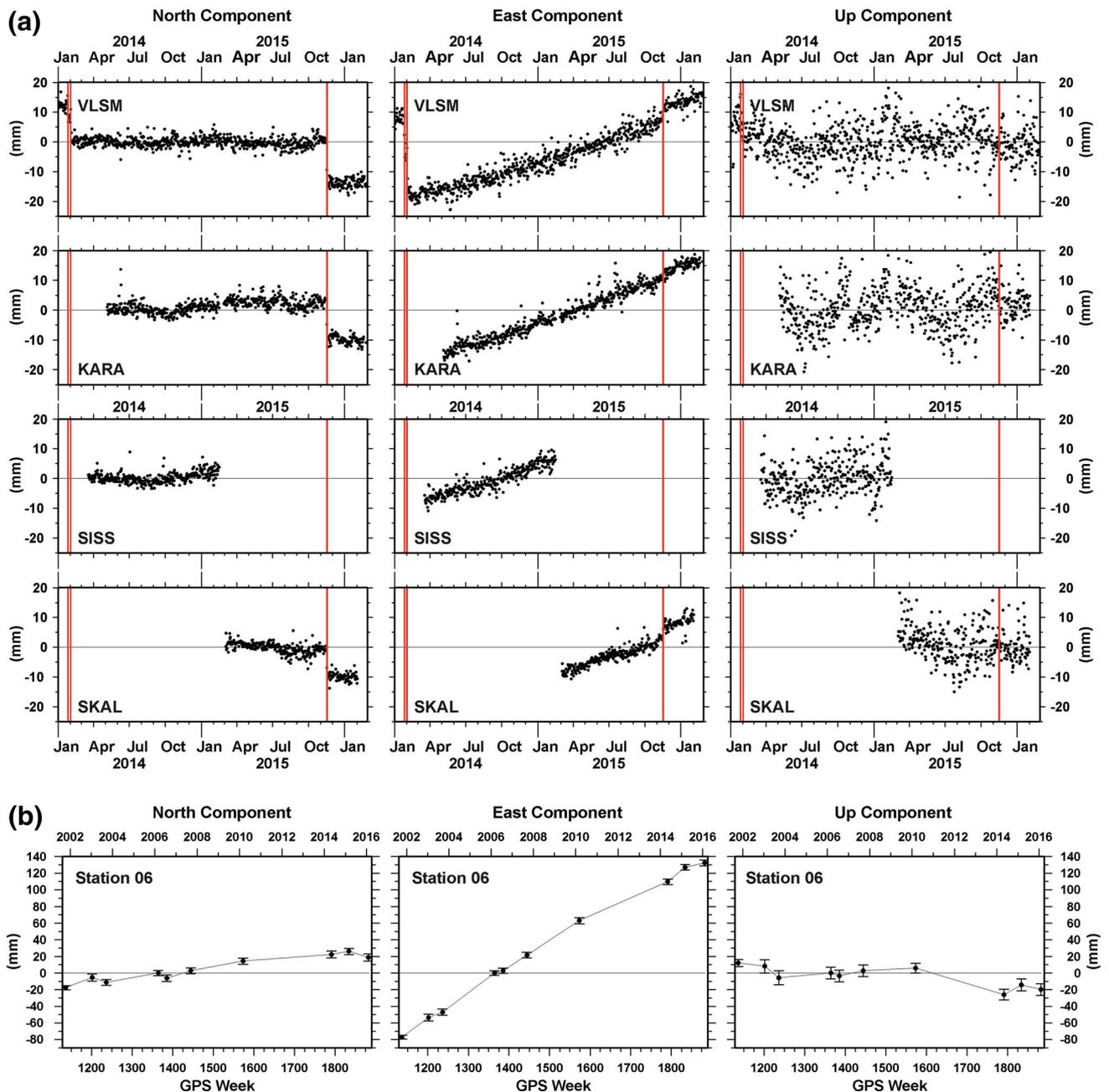


Fig. 3 Time-series of the displacements observed at the continuous GPS stations in Cephalonia for the period Jan 2014 to Feb 2016 (a), together with the components of GPS station 06 (b), which has been used earlier as a network reference station in Aenos Mt. for the period

2001–2016 (Reference Frame: IGB08). Red vertical lines mark the occurrence of the main seismic events in the broader area (2014–2016)

Displacement of the vertical component during the same event was observed only at station PONT (subsidence ~ -56 mm), which is located very close to the epicentral area, while there were no clear vertical displacements for the rest of the cGPS stations. A nearby benchmark to VLSM, station 06, which served as a local reference station from 2001 to 2010, showed a similar pattern to VLSM for all the observational periods (Fig. 3b).

Period March 2014–March 2015

Figure 5 represents the displacement vectors for the period interseismic period March 2014 to March 2015 (Table 2). The overall character of the ground deformation was of an ENE displacement with small horizontal amplitudes that describes the anticipated motion of the area with respect to the IGB08 reference frame, as depicted by the cGPS

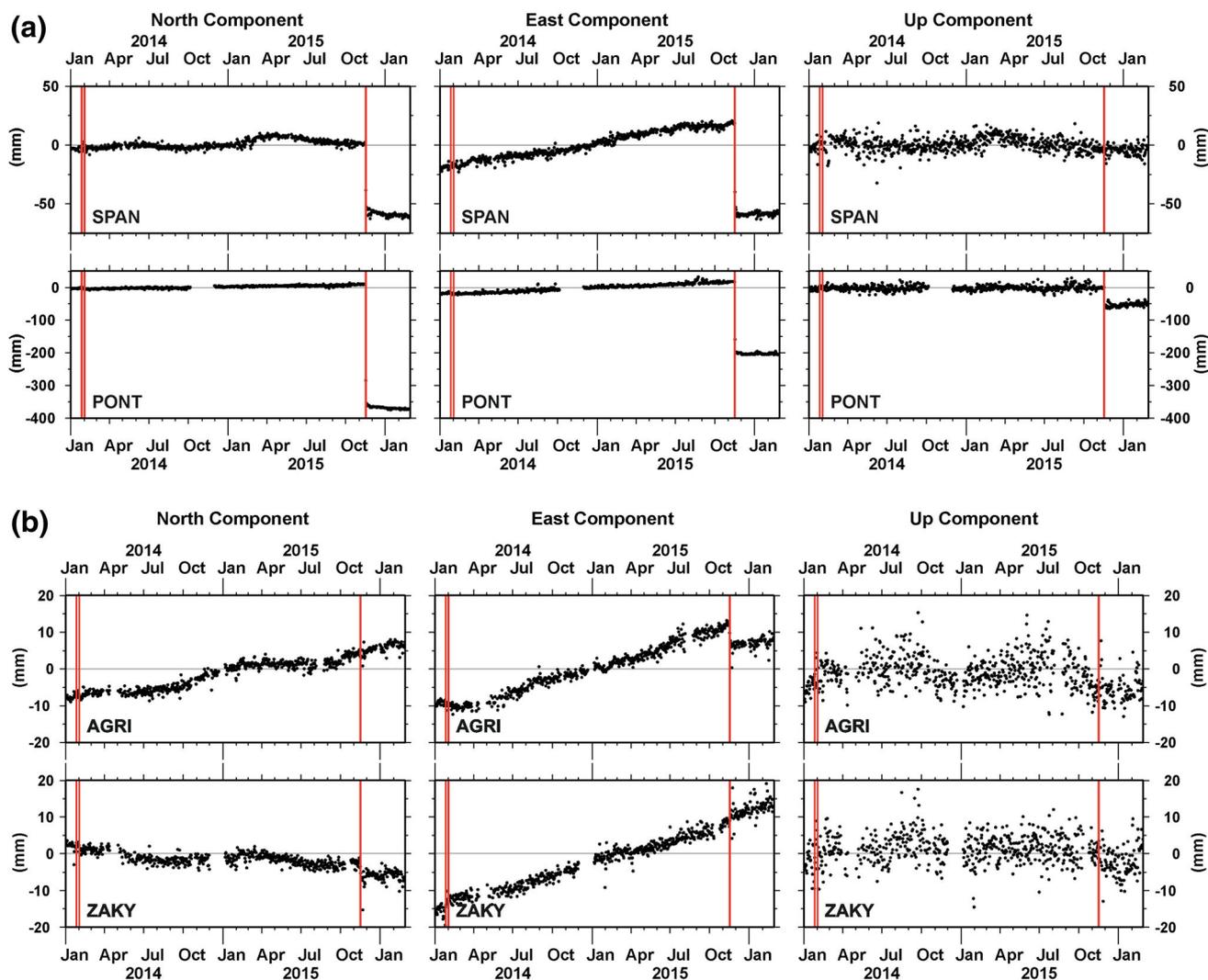


Fig. 4 Time-series of the components observed at the continuous GPS stations in **a** Lefkas in its northern (SPAN) and southern (PONT) parts, and **b** in central-western Greece (AGRI) and Zakynthos Island

(ZAKY) (Reference Frame: IGB08). *Red vertical lines* mark the occurrence of the main seismic events in the broader area (2014–2016)

stations on the island (Fig. 3). However, there were some discrepancies in the deformational pattern along the island. The southern part of Cephalonia exhibited an almost ENE displacement with small horizontal amplitudes (~ 11 mm) and almost parallel vectors for all stations located in the area. In the western part of Cephalonia, in Paliki Peninsula, a slight differentiation of the motion along a north–south axis seemed to appear, with stations to the north showing a more northward displacement compared to those to the south. This differentiation took place in the vicinity of station 16, where a boundary of displacement has already been observed and outlined in an earlier analysis (Sakkas and Lagios 2015).

In the northern peninsula of Cephalonia, the amplitude of the horizontal component was slightly increased compared to the rest of the island (~ 25 mm). However, station

12 shows a smaller amplitude of displacement (~ 12 mm), a behavior that has also been observed in the past (Lagios et al. 2012). The increased horizontal component observed at these stations may be attributed to the seismic activity that followed the 2014 Cephalonian earthquakes, and was concentrated in the sea just to the west along the northern peninsula, and which was in progress during and after the observational period of March 2014. A similar behavior in this part of the island was also noticed during the post-seismic activity after the 2003 Lefkas earthquake (Lagios et al. 2007). A noticeable element during this period was the discrepancies in amplitude of the horizontal displacement for some neighboring stations, where their proximity did not fully justify the significant differences that were observed. Stations 08 and 09, VLSM and 06, as well as stations 10 and 11 showed displacements of similar

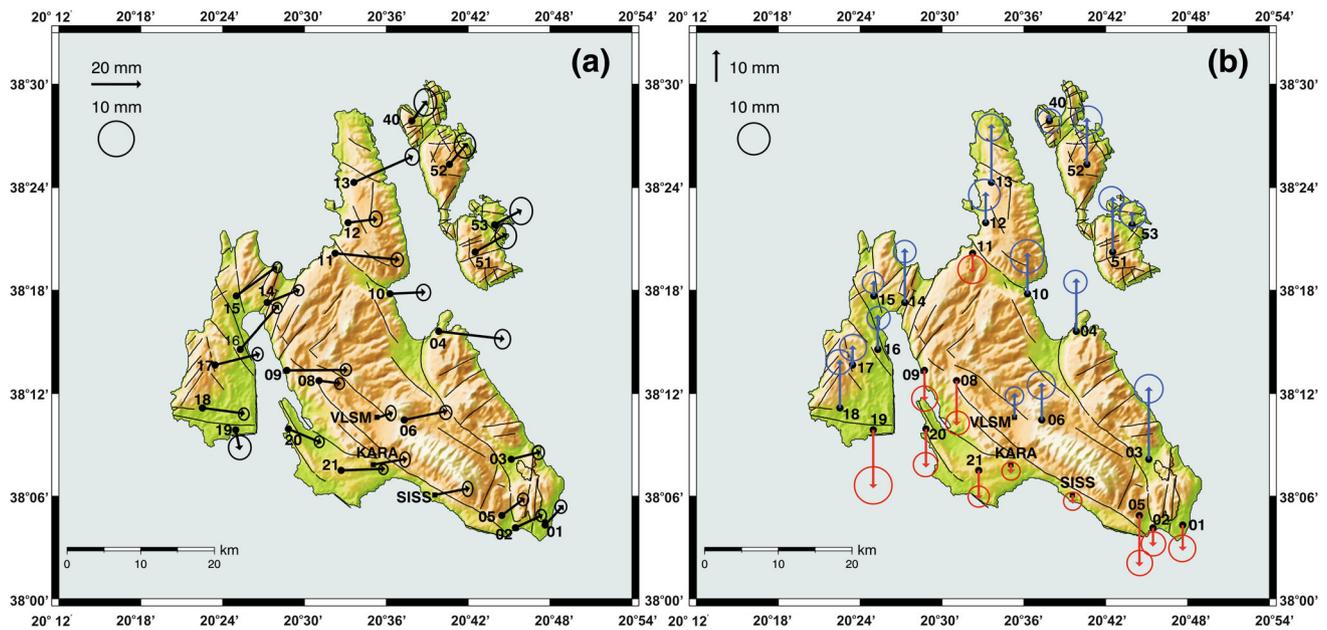


Fig. 5 **a** Horizontal, and **b** vertical displacements from the Cephalonia–Ithaca GPS network for the period March 2014 to March 2015. *Red* and *blue* arrows indicate subsidence and uplift, respectively (Reference Frame: IGB080)

Table 2 Component displacements of GPS and cGPS stations referred to IGB08 GPS period 2014–2015

| GPS station | N–S (mm) | N–S SD (mm) | E–W (mm) | E–W SD (mm) | Up (mm) | Up SD (mm) |
|-------------|----------|-------------|----------|-------------|---------|------------|
| 01 | 8.1 | 3.7 | 7.3 | 3.2 | −8.1 | 8.2 |
| 02 | 5.3 | 3.4 | 11.0 | 2.8 | −5.9 | 7.6 |
| 03 | 3.1 | 3.9 | 11.9 | 3.4 | 22.8 | 8.8 |
| 04 | −3.1 | 5.4 | 26.3 | 4.4 | 16.6 | 7.2 |
| 05 | 7.0 | 3.6 | 9.2 | 3.0 | −16.0 | 8.0 |
| 06 | 3.7 | 3.9 | 17.3 | 3.3 | 11.7 | 8.6 |
| 08 | −1.2 | 3.2 | 8.9 | 2.8 | −14.1 | 7.4 |
| 09 | 0.2 | 3.3 | 24.4 | 2.9 | −9.6 | 7.8 |
| 10 | 0.7 | 4.6 | 14.2 | 3.9 | 13.0 | 10.4 |
| 11 | −2.6 | 3.9 | 25.7 | 3.4 | −6.2 | 8.9 |
| 12 | 1.5 | 4.3 | 11.7 | 3.6 | 9.7 | 9.7 |
| 13 | 10.7 | 4.7 | 23.7 | 3.9 | 18.3 | 8.5 |
| 14 | 5.3 | 3.1 | 13.1 | 2.7 | 17.0 | 6.9 |
| 15 | 12.0 | 3.0 | 16.4 | 2.5 | 4.9 | 6.6 |
| 16 | 17.9 | 3.1 | 15.4 | 2.7 | 10.8 | 7.1 |
| 17 | 4.6 | 3.7 | 18.1 | 3.1 | 6.2 | 8.3 |
| 18 | −2.3 | 3.4 | 17.2 | 2.9 | 15.1 | 7.7 |
| 19 | −8.2 | 6.7 | 1.7 | 6.2 | −18.3 | 11.7 |
| 20 | −5.6 | 3.2 | 13.2 | 2.7 | −12.1 | 7.5 |
| 21 | 0.7 | 2.9 | 17.6 | 2.6 | −9.1 | 6.6 |
| 40 | 8.2 | 8.6 | 6.1 | 7.3 | 2.3 | 7.4 |
| 51 | 7.4 | 8.1 | 12.8 | 6.9 | 17.4 | 7.7 |
| 52 | 8.7 | 8.2 | 7.2 | 6.9 | 14.7 | 8.8 |
| 53 | 6.0 | 8.6 | 10.9 | 7.2 | 4.2 | 8.0 |
| KARA | 2.3 | 3.1 | 13.4 | 2.6 | −2.9 | 4.1 |
| SISS | 2.80 | 3.0 | 14.4 | 2.9 | −2.7 | 4.5 |
| VLSM | 1.9 | 2.9 | 5.7 | 2.4 | 7.7 | 3.7 |

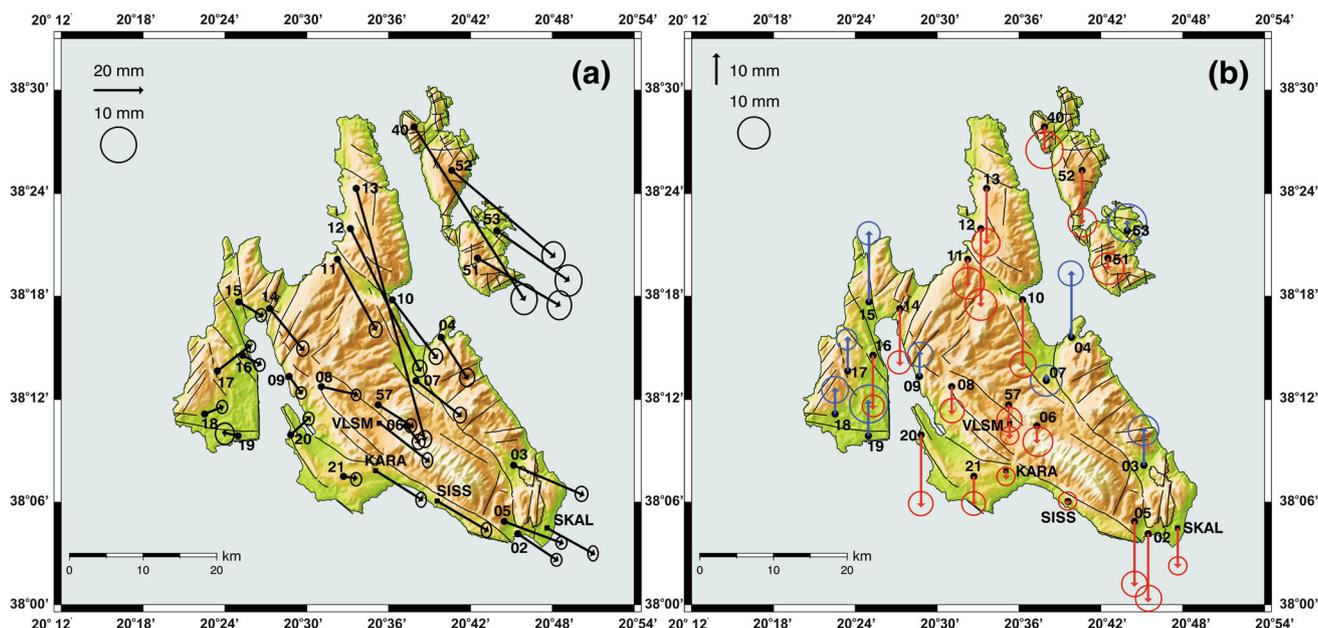


Fig. 6 **a** Horizontal, and **b** vertical displacements from the Cephalonia–Ithaca GPS Network for the period March 2015 to February 2016. *Red and blue arrows* indicate subsidence and uplift, respectively (Reference Frame: IGB08)

direction, but with high-amplitude deviations. These deviations are attributed to local phenomena possibly associated with the complex and extensive nature of faulting in the near vicinity.

The vertical component (Fig. 5b) revealed a more unified pattern of displacement. Subsidence reaching amplitude of up to -15 mm was observed along the southwestern part, an area characterized by a dominant NW–SE thrusting zone in this part of the island. For the rest of the stations, the vertical component revealed a moderate uplift. Taking into account the larger uncertainties involved in the calculation of the vertical component, it was assumed that the small discrepancies in that pattern (e.g., station 11) and in the amplitude may be attributed to local tectonic phenomena. For instance, the strong uplift (~ 20 mm) at the eastern stations 03 and 04 could be attributed to the presence of underlying evaporites of the Ionian Zone (Lekkas and Mavroulis 2016).

Finally, a non-consistent pattern of displacement for all measuring periods (2014–2016) both in the vertical and horizontal direction was noticed at station 19, in the southern part of Paliki. This irregularity may be attributed to predominately loose marine sediments in that area, probable fluctuation of the local water table, and lack of good satellite visibility that noticeably decreased the accuracy of the calculated coordinates.

The deformational pattern along Ithaca for this first remeasurement period depicted the previously described differential motion between its northern and southern parts (Lagios et al. 2012; Sakkas and Lagios 2015). The latter

becomes evident by the slightly different directions of the horizontal displacements between these parts. The amplitude of the horizontal displacements was about 11 mm at almost all stations, while the vertical components exhibited larger variations. It may be assumed that stations 52 and 51 better describe the kinematic state of the island, where almost the same amplitude of uplift (~ 15 mm) was measured.

Period March 2015–February 2016

The deformation that took place in the study area during this period was strongly affected by the 2015 Lefkas earthquake and the seismic activity that followed (Fig. 6; Table 3). Intense SSE displacement was observed in the northern, central and southern parts of Cephalonia, as well as in Ithaca. The amplitude of the horizontal displacements gradually decreased from about 100 mm in the northernmost part (station 13) to about 20 mm in the southern part of the island. These amplitudes were almost an order of magnitude higher in the north as compared to the previous period. In the central, western and southern parts, the horizontal displacement was almost two times larger than that in the period 2014–2015.

However, Paliki Peninsula, located in the western part of Cephalonia, revealed a significantly smaller horizontal displacement compared to the rest of the island, inferring a boundary that differentiated the motion in this part of the island from the other areas. Differentiations of the deformation pattern are also observed and along the stations

Table 3 Component displacements of GPS and cGPS stations referred to IGB08 GPS period 2015–2016

| GPS station | N–S (mm) | N–S SD (mm) | E–W (mm) | E–W SD (mm) | Up (mm) | Up SD (mm) |
|-------------|----------|-------------|----------|-------------|---------|------------|
| 02 | –11.1 | 3.6 | 16.3 | 3.0 | 3.6 | 8.1 |
| 03 | –12.2 | 4.3 | 28.3 | 3.6 | 4.2 | 9.4 |
| 04 | –17.1 | 5.1 | 10.8 | 4.2 | 5.0 | 7.6 |
| 05 | –9.1 | 3.6 | 23.8 | 2.9 | 3.6 | 7.9 |
| 06 | –7.1 | 4.3 | 5.4 | 3.6 | 4.2 | 9.2 |
| 07 | –14.6 | 4.4 | 18.3 | 3.6 | 4.3 | 9.8 |
| 08 | –3.4 | 3.3 | 14.9 | 2.7 | 3.2 | 7.5 |
| 09 | –7.4 | 3.3 | 5.3 | 2.9 | 3.3 | 7.6 |
| 10 | –23.8 | 4.4 | 18.0 | 3.7 | 4.4 | 8.0 |
| 11 | –28.9 | 4.3 | 15.6 | 3.6 | 4.3 | 9.7 |
| 12 | –57.7 | 4.8 | 28.0 | 3.9 | 4.8 | 9.7 |
| 13 | –101.9 | 4.8 | 27.5 | 3.9 | 4.8 | 8.6 |
| 14 | –17.0 | 4.2 | 13.8 | 3.4 | 4.2 | 7.1 |
| 15 | –6.0 | 3.7 | 10.0 | 2.9 | 3.6 | 7.0 |
| 16 | –4.0 | 3.5 | 7.8 | 2.9 | 3.5 | 6.8 |
| 17 | 10.7 | 3.3 | 13.6 | 2.8 | 3.3 | 6.4 |
| 18 | 3.1 | 3.6 | 7.8 | 3.0 | 3.6 | 8.2 |
| 19 | 1.5 | 6.0 | –6.5 | 5.2 | 6.0 | 11.8 |
| 20 | 7.5 | 3.6 | 8.0 | 3.0 | 3.6 | 7.5 |
| 21 | –0.9 | 3.6 | 5.7 | 3.1 | 3.6 | 7.2 |
| 40 | –70.2 | 8.6 | 44.2 | 7.2 | 8.6 | 10.5 |
| 51 | –19.4 | 8.1 | 33.2 | 6.8 | 8.1 | 9.8 |
| 52 | –35.2 | 7.6 | 41.3 | 6.4 | 7.6 | 8.7 |
| 57 | –9.1 | 3.9 | 14.3 | 3.2 | 3.9 | 7.7 |
| 53 | –20.5 | 8.8 | 29.1 | 7.4 | 8.8 | 10.8 |
| KARA | –12.0 | 3.5 | 18.8 | 2.6 | 4.2 | 4.0 |
| SISS | –12.5 | 4.2 | 20.7 | 2.8 | 4.2 | 4.5 |
| SKAL | –10.6 | 3.1 | 19.0 | 2.0 | 4.24 | 3.6 |

located in Palliki. Comparing the displacements between the southern and northern parts of Paliki, a kinematic boundary becomes evident just to the south of station 16, as deduced by the observed deformation at the GPS stations. This inferred boundary caused the southern part of Paliki to have an almost NE motion, while the northern part showed an SE motion that should have been affected mainly by the 2015 Lefkadian event. This kinematic discontinuity is also supported by our previous studies (Sakkas and Lagios 2015), where it was mentioned that a boundary should exist in the vicinity of station 16. Moreover, the small horizontal amplitudes at the stations in the northern Paliki area with respect to those at the northernmost part of Cephalonia clearly seem to mark a discontinuity between Paliki and the rest of the island. This argument is supported by the gradual decrease of the horizontal displacements from ~21 mm (station 14) to ~11 mm (station 15), and finally to ~8 mm (station 16).

Similar to the previous period (2014–2015), neighboring stations (08 and 09, VLMS and 06, KARA and 21) also

exhibited non-uniform patterns of deformation regardless of their proximity. This highlights the prevailing intense tectonism in the area, which together with the increased seismicity during the span of this study (2014–2016) seems to govern the local and regional motions.

The pattern of the vertical component (Fig. 6b) confirmed the differentiation between the mainly uplifting Paliki and the subsidence of the remaining portion of Cephalonia, while deviations from this main pattern are attributed to local tectonic and geological features. The pattern of the vertical component showed an intense subsidence in the northern part (–17 mm), close to the Lefkas epicentral area, decreasing gradually towards the central part (about –6 mm), and slightly increasing again in the southern part (about –15 mm), as it was similarly observed during the previous observational period (2014–2015). Stations 03 and 04 in the eastern and southeastern parts revealed uplift almost identical to the previous period, while two other stations in central Cephalonia (07 and 57), which have not been remeasured for several years,

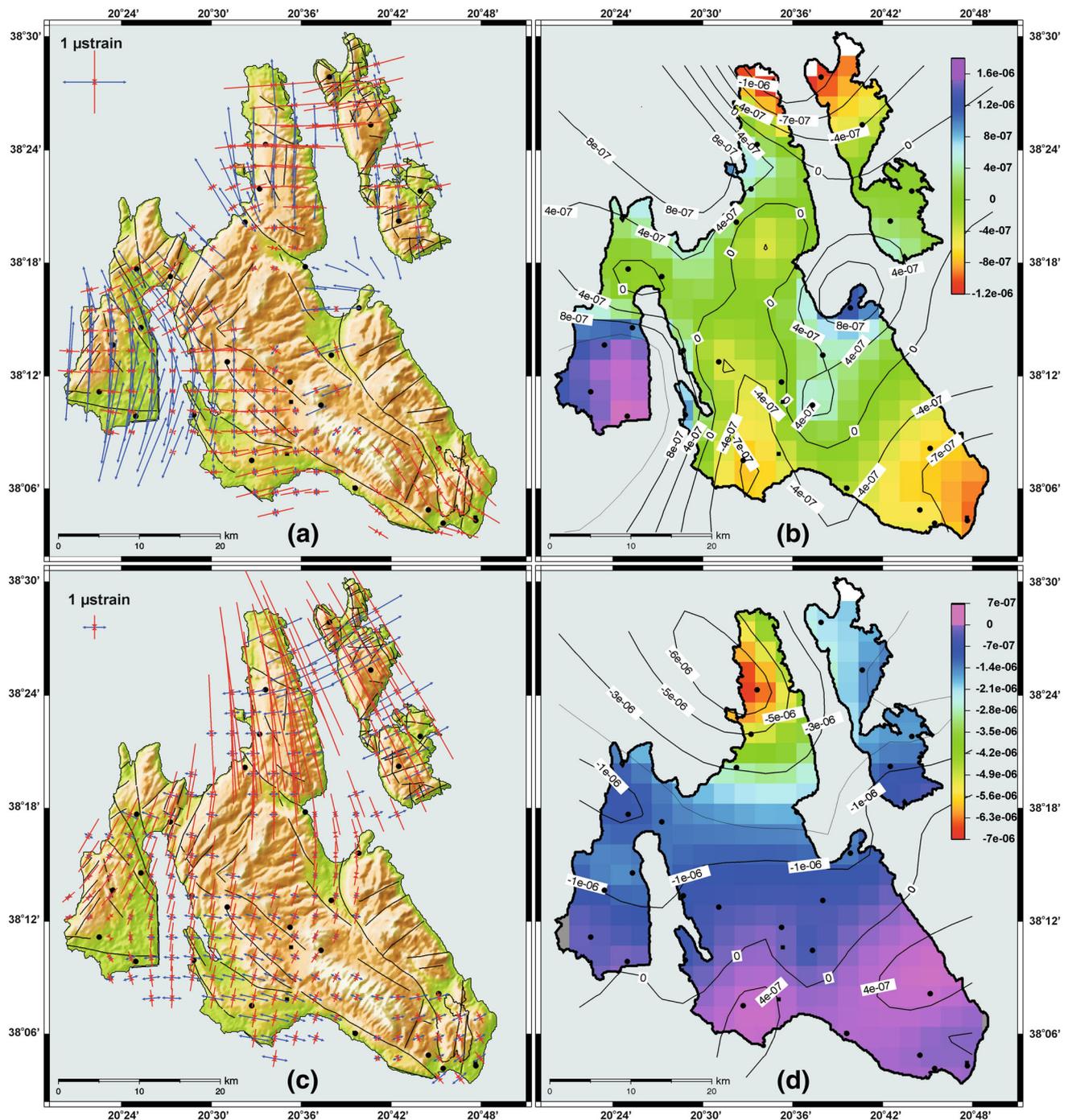


Fig. 7 **a** Horizontal strain tensor distribution for the period 2014–2015; **b** change field map for the period 2014–2015; **c** horizontal strain tensor distribution for the period 2015–2016; **d** change field

map for the period 2015–2016. *Red and blue arrows* indicate compression and extension, respectively. Note the different strain scales in **a** and **c**

exhibited an overall displacement in accordance with the other stations in the vicinity (06 and 04).

Itthaca was close to the epicentral area of the 2015 Lefkas seismic event and was shifted to the SE, preserving the differential type of displacement between its northern

and southern parts. The horizontal components of the northern stations (~ 80 mm) were significantly higher than the southern ones (~ 40 mm), while the northern part seemed to subside compared to the southern part whose vertical motion was found to be within the error limits.

Strain field

Calculation of the strain tensor based on the horizontal components of displacement in the two observational periods was performed to better define the deformation occurred in the area in the period 2014–2016, during the post-seismic period of the 2014 Cephalonian events and 2015 Lefkadian one, importing time discontinuities (earthquakes) in the calculations. The strain field tensors describe the changes of dimensions or shape along a deformed area and it can be used to better define the tectonic forces in an area of high tectonic activity. The strain field was computed based on the Teza et al. (2008) algorithm on a uniform (2 km × 2 km) grid. The adopted small grid size that is about half the average station spacing (~4 km) may reveal useful information about the deformation occurred in the local tectonic blocks of the area, as well as along the space discontinuities (faults). In each grid node, the eigenvectors were calculated with positive and negative eigenvalues describing extension and compression, respectively. The strain was considered representative of the compressional and/or tensional state of the area only at the nodes whose distance from the data points is smaller or equal to the scale factor used in the calculations (Teza et al. 2008). Based on the computation of the strain tensor, the corresponding trace for each grid point was calculated describing the change in the area (dilatation) that is equal to the sum of the diagonal elements of the tensor. The dilatation corresponds to the relative change of surface area and thus provides the possibility to identify areas of compression or extension, and indicate regions of thrusting or normal faulting (Hackl et al. 2009).

The spatial distribution of the strain during the first remeasurement period 2014–2015 (Fig. 7a) shows a significant attenuation of the extensional deformation from the western part of Paliki toward the eastern part of the island, which is characterized by compression. The area of Paliki with the highest extensional amplitude was obviously affected by the post-seismic visco-elastic relaxation and/or afterslip induced by the 2014 events, as well as by the following seismic sequence. The northern part of Cephalonia and Ithaca exhibited compression with an almost E–W direction of the principal compressive strain, almost perpendicular to the Ionian Thrust Zone (Fig. 1). However, extension occurred in between the southern part of Ithaca and the central-eastern part of Cephalonia. Finally, the southern and southeastern parts of Cephalonia exhibited compression, with its main axis perpendicular to the thrusting lineaments of the area, indicating strain accumulation. The tensors trace (dilatation) which represents areal deformation (Fig. 7b) highlights the

differentiation between the northern and southern part of Paliki, where extension occurred in the southern part. Compression is evident between the northern parts of Cephalonia and Ithaca, as well as in the southern and southeastern part of Cephalonia, differentiating these parts from the rest of the island. The localized dilatation pattern in Paliki may suggest that afterslip, and not a broad-scale visco-elastic relaxation, is the dominant post-seismic mechanism.

The strong Lefkadian event in November 2015 dominated the pattern and the amplitudes of the strain field vectors for the next period, 2015–2016 (Fig. 7c). However, the study of the strain during that period may generally provide insight into the co-seismic deformation, slip location, rupture characteristics, resolving the rupture plain ambiguity of the focal mechanism solution. As expected by the strike slip nature of the Lefkadian event, high compressive strain exhibited in the northern part of Cephalonia, as well as in Ithaca, with the main axis of compression being in an NW–SE direction. A significantly smaller E–W extensional mechanism is observed between northern Cephalonia and Ithaca. This alteration of the deformation mechanism, in our case from extension during interseismic period to compressional on a co-seismic period, satisfies the elastic rebound theory. The amplitude of the compression attenuates towards the southern part of Cephalonia, which has affected less from the Lefkadian event. The dilatation analysis (Fig. 7d) shows that the far field has a very good agreement with the observed focal mechanism of the November 17th, 2015 event (Fig. 2). The southern and southeastern parts of Cephalonia are differentiated from the rest of the island as also noticed in the period 2014–2015.

At an effort to estimate the strain without the co-seismic effect during the period 2015–2016, the fault model produced by Ganas et al. (2016) was used to calculate the co-seismic deformation in Cephalonia–Ithaca islands. The calculated displacement vectors which represent the deformation without the co-seismic effect (which were produced subtracting the model vectors from the observed ones) resulted to a pattern that is not consistent to the previous interseismic periods 2014–2015 and 2001–2010 (Lagios et al. 2007, 2012). It seems that the fault model used to describe the co-seismic displacement in Lefkas resulted in significantly higher displacements in northern Cephalonia and Ithaca compared to the observed ones. Taking into consideration that this event occurred on the Lefkadian part of the CTF, west of the Ionian thrust zone (Fig. 1), it seems that a kinematic boundary should exist in this area (east of CTF) that reduces the co-seismic displacement amplitudes in Cephalonia and Ithaca. This boundary may be related to the Ionian thrust. As a consequence, the strain tensor computed for the period

2015–2016 subtracting the co-seismic effect cannot be used for further analysis.

Discussion and conclusions

The intense seismic activity caused by three major events of $\sim M6$ that took place in the broader area of the central Ionian Islands during 2014–2015 is not unusual for such a highly tectonically active area. The ground deformation measured over almost two years (March 2014 to February 2016) from local GPS networks in Cephalonia and Ithaca, including cGPS stations, revealed displacements that describe the kinematic state of significant tectonic features at both local and regional scales. Moreover, these displacements enabled us to better understand the tectonism during the interseismic (2014–2015) and co-seismic (2015–2016) periods.

The broader area of Cephalonia and Lefkas is usually strongly affected by the Cephalonia Transform Fault; nevertheless, the 2014 earthquakes in Cephalonia (Paliki) have shown that there are additional faulting zones with the potential to generate strong events. The 2015 Lefkadian earthquake took place just south of an older event in 2003, and was located along its southern continuation (Ganas et al. 2016; Chousianitis et al. 2016). Both epicentral areas in Cephalonia and Lefkas generated earthquakes with similar focal mechanisms (Fig. 2), showing seismogenic faults of similar geometric characteristics. These two distinct events were followed by seismic activity that extended northward and southward, respectively (Fig. 2). This activity linked the two epicentral areas, thus indicating that the major CTF zone is a system that encompasses several smaller branches capable of generating intense seismicity. The time series of the cGPS stations have shown that the 2014 Paliki events ($M_w 6.1$ and $M_w 5.9$) caused displacements that were limited mainly to Cephalonia and did not affect Lefkas to the north or Zakynthos to the south (Figs. 3, 4). On the contrary, the occurrence of the Lefkadian event ($M_w 6.4$) affected a much broader area, including Cephalonia and Zakynthos to the south and extending also to the west (Agrinio in western Central Greece, Fig. 4). This supports the observation that the Lefkadian earthquake was generated on the regional CTF system, while the Cephalonian events were generated by the local faulting system in Paliki.

Inspecting the time series of the cGPS stations prior to the Lefkadian event and after the Cephalonian ones, it was found that there was no pre-seismic deformation in the area, regardless of the CTF's major regional characteristics. All stations showed the anticipated regional motion, with the exception of small dislocations that were recorded

during localized smaller earthquakes (e.g., KARA in November 2014, Fig. 3a).

The cGPS time series from the Cephalonian stations (VLSM, KARA, SKAL) showed that the co-seismic deformation due to the Lefkadian earthquake was about 10 mm in the central and southern parts of the island, having an almost linear type of motion in the period before the event. Comparing the amplitude of the deformation vectors for the periods 2015–2016 and 2014–2015, it is evident that the doubling of the deformation amplitude in the period 2015–2016, at least for the central and southern part of the island, was due to the Lefkadian earthquake. The estimated deformation prior to the Lefkadian earthquake may be considered of an almost linear type, at least for the central and southern parts of Cephalonia. For the rest of the island, the large deformation amplitudes observed on the northern peninsula, together with the offshore seismic activity west of the peninsula after the 2014 and 2015 earthquakes, do not provide any indication of the deformation type (linear or non-linear) prior to the Lefkadian event.

A fault zone just to the west of station 14 that extends at an almost N–S orientation along the eastern part of Paliki has already been modeled to explain the co-seismic deformation in the area during the 2014 earthquakes (Sakkas and Lagios 2015). This zone was also active during the two observational periods discussed in this paper (2014–2015 and 2015–2016), both in terms of horizontal and vertical components, although the patterns and amplitudes of displacement are different for each period. A continuation of this faulting zone to the offshore and north of Paliki may be assumed, taking into consideration the extension of the seismic activity to the north that occurred in this area after the 2014 events. This zone may be composed by multiple small fault segments, similar to the modeled faults of the 2014 events. This northward extending fault system may eventually be linked to the major CTF in the offshore, as is indicated by the seismicity in that area (Fig. 2). In addition, another NW–SE trending fault, just to the north of station 13, may also be truncated in the same marine area to the west that marks the northern and southern limits of the seismic activity during the 2014 Cephalonian and 2015 Lefkadian events, respectively (Fig. 2). This fault marks a clear deformation boundary in the northernmost peninsula of Cephalonia, based on earlier PSInSAR analysis (Fig. 6 in Lagios et al. 2012), and may be associated with offshore structures of similar direction as was revealed by seismological analysis in the area (Karakostas et al. 2015).

Another feature of interest is the NW–SE trending fault on Paliki, just to the south of station 16 (Fig. 1), which seems to represent an important boundary between the

southern and northern parts of Paliki. In our efforts to model the early 2014 earthquakes (Sakkas and Lagios 2015), it became necessary to assume the presence of two consecutive faults belonging to the same system, but with a gap between them, which generated the two events in Cephalonia. The lack of GPS observations before the second M_w 5.9 event, as well as the intense aftershock activity that occurred during the study period, did not confirm convincingly this assumption. However, the observed ground deformation between the southern and northern parts of Paliki identified in this paper provides more evidence for the existence of such a discontinuity. During the first period (Fig. 5), northern Paliki exhibited a displacement to the NE, while the southern part showed an eastward motion. During the second period, 2015–2016 (Fig. 6), southern Paliki showed NE displacement, while its northern half, being affected by the Lefkadian event, was subject of an SE dislocation. These results outline the intense tectonism that is taking place in Paliki, which is an area dominated by the major CTF to the west and an almost parallel faulting zone to the east that has been postulated based on fault plane modeling of the 2014 earthquakes. The different displacements of the northern and southern part of Paliki across the NW–SE fault that lies to the south of station 16 seem to constitute an essential local kinematic boundary, as this was also indicated in the change map of the area (Fig. 7b).

The effect of the NW–SE thrusting zone just north of stations 10 and 11 is clearly evident in the period 2015–2016, when comparing the displacements observed at stations 10 and 11 to those at the northernmost stations 12 and 13. The latter can also be seen in the change map (Fig. 7d). This zone marks the limits of the northern peninsula of Cephalonia that affect the local kinematics. This peninsula moves south-eastward with higher deformational amplitude than the central Cephalonia. As a result, force is exerted towards the southeastern part, in the vicinity of station 10, without any evidence that the coastline marks the end of this zone. This kinematic behavior along this zone may be reflected to the neighboring station 10 that shows smaller dislocation when compared to the surrounding stations throughout the measuring periods (Lagios et al. 2007, 2012; Sakkas and Lagios 2015), both for the vertical and horizontal components.

Local tectonic features in the central and southern parts of Cephalonia, which cause differentiated displacements for stations located across/along them, are observed between stations 08 and 09, VLSM and 06. Thrusting lineaments caused the observed compressional forces, as these have been depicted by the change map for the interseismic period (Fig. 7b). The influence of these local tectonic features becomes more evident in some areas after the 2015 Lefkadian event. Stations 20, 21 and KARA,

located close to the NW–SE major thrusting zone, revealed significant variations in both amplitude and direction of the deformational vector for the period 2015–2016 (Fig. 6). In combination with the occurrence of several earthquakes of $\sim M_w$ 4 during November 2014, this indicates the probable existence of other buried or unknown faults in the vicinity, where their displacement resulted in a non-uniform deformational pattern observed at the surrounding stations. Similar differential displacements along the major thrusting zones running along the southern part of Cephalonia were also observed by radar interferometric analyses (Lagios et al. 2012) during the previous years when the seismic activity was at regular levels.

The displacements that occurred between Cephalonia and Ithaca are another interesting aspect of this study. A thrust system separates northern Cephalonia (belonging to the Paxos zone) from Ithaca of the Ionian zone (Fig. 1). The enhanced eastward displacement during 2014–2015 (Fig. 5) at the northern Cephalonian stations compared to those in Ithaca, in combination with the strain map of the area (Fig. 7a, b), is indicative of the E–W compressional forces between the two islands during that interseismic period. Based on these observations, it can be concluded that the two islands are “approaching” each other. The compressional forces that seem to prevail in the broader area (Ganas et al. 2013; Chousianitis et al. 2015) may explain this kinematic behavior, taking also into account the regional motion in the broader area of western Greece (e.g., Vassilakis et al. 2011). The effect of the Lefkadian event in the co-seismic period 2015–2016 changed the overall strain map of the northern Cephalonia–Ithaca area, as it is expected by the elastic rebound theory (Fig. 7c). The strain tensors have rotated about 90° , with an almost N–S compression and a small E–W co-seismic extension. This extensional component between Ithaca and Cephalonia was also occurred during the 2014 Cephalonian events (Sakkas and Lagios 2015) causing the two islands to “move away” from each other during major seismic events.

In the Paliki Peninsula, a significant extensional component observed along an almost N–S axis during the 2014–2015 measuring period may be attributed to the effect of the seismic sequence of the 2014 Cephalonian events, along the local faulting system that are running along Paliki, suggesting that afterslip is most likely the main post-seismic mechanism in the area.

Another feature of the observed ground deformation in Cephalonia and Ithaca is the differential displacement between southern Ithaca and central-eastern Cephalonia. Although Ithaca and this part of Cephalonia belong to the same geological unit of the Ionian Zone (Fig. 1), the two areas show considerable differences in the displacement vectors (direction and amplitude) for both periods. The

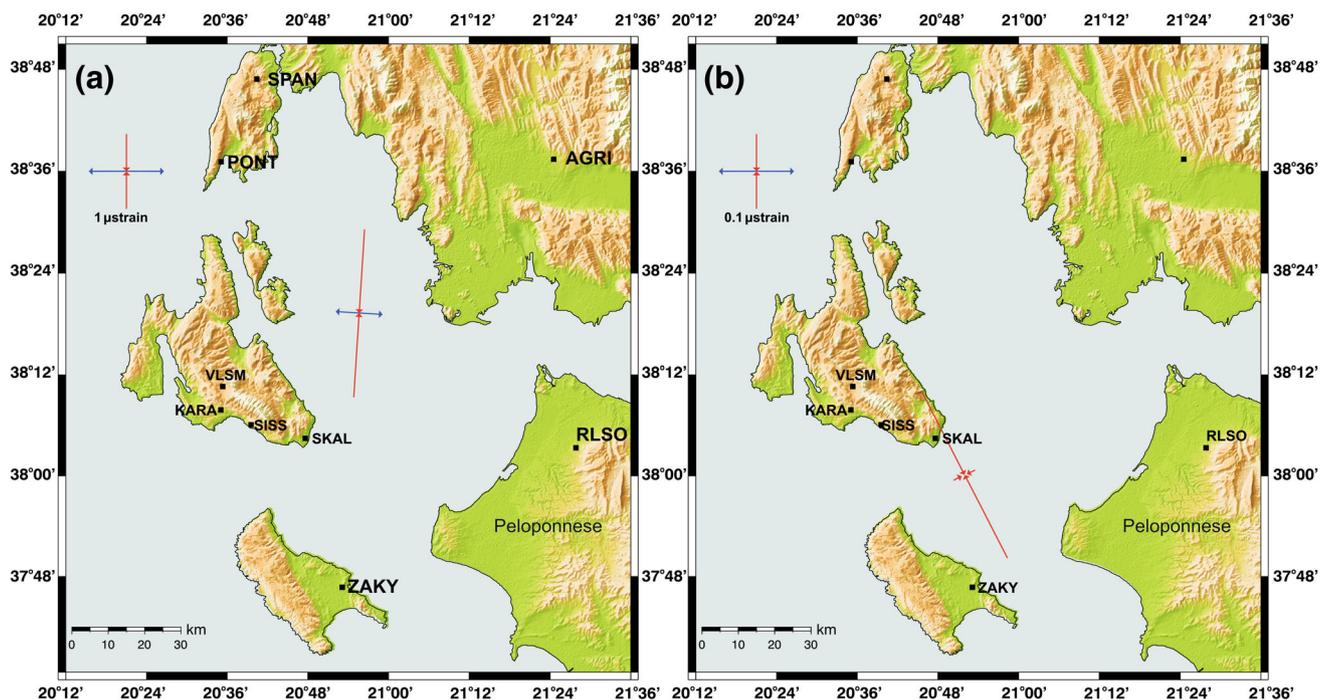


Fig. 8 Horizontal strain tensor in the broader area of the central Ionian Islands based on the recorded deformation at the cGPS stations during the Lefkas event: **a** all available stations (9 stations) were taken into consideration. **b** Only stations in the southern Ionian Sea (6

stations) were included in the calculations. *Red* and *blue* arrows indicate compression and extension, respectively. Note the different strain scales in **a** and **b**

Ionian Thrust that traverses this area may explain the differences in northern Cephalonia and Ithaca, but cannot account for the difference between southern Ithaca and central-eastern Cephalonia, especially taking into consideration the high dilatation observed in the vicinity of station 04 in the period 2014–2015 (Fig. 7b). The differential displacement between the study areas may be attributed to the presence of an offshore E–W faulting system that crosscuts the Ionian Thrust and is almost parallel to the southern coastline of Ithaca, as well as the coastline to the north of station 04 in Cephalonia. The morphology of these two coastlines is consistent with a possible ENE–WSW fault offshore, especially since many faults in both islands (southern Ithaca and eastern Cephalonia) have a similar strike direction. However, the high dilatation observed locally in the area of station 04 may be caused by a small creeping segment, and may explain also the observed differential motions.

The compression in northern Cephalonia (Fig. 7b) resulting from the strong horizontal displacements due to the $M_w 6.4$ Lefkadian event implies that the area of the main island is under a significant stress. Cephalonia–Ithaca contains several deformation zones that experience motions related to local and regional tectonic features. The regional features control the observed deformation field, as suggested by the motion of the main tectonic fragments. However, internal localized faults are clearly affecting the

movement of smaller units, which is associated with the locally recorded seismic activity.

Motions along the major CTF structure are the main source of strong seismic activity in the area, as well as regional increase of stress. The localized seismic events of the 2014 and the regional 2015 event on the Lefkadian part of CTF may be linked to stress transfer to nearby fault zones in Cephalonia. Strain calculations based only on cGPS data in the broader area (i.e., two stations in Lefkas, four in Cephalonia, one in Zakynthos, one in NW Peloponnese, and one in west-central Greece) during the Lefkadian event showed the strong compression occurred in the center of the study area (Fig. 8a). Compression is also observed in the sea area between Cephalonia and Zakynthos, when taking into consideration just the southern part of the Ionian Sea (Fig. 8b). According to these observations, and taking into consideration the compression observed in southern Cephalonia during the interseismic periods (Fig. 7a), it may be argued that strain should be accumulating in the southern part of Cephalonia. Previously published work combining GPS and Permanent Scatterer Interferometric data has shown that an intense non-linear uplift took place in Paliki, and another one later along the major NW–SE trending thrusting zone at the southern part of Cephalonia, when compared to the previous decade, which was interpreted as a possible regional dilatancy (Lagios et al. 2012) that could potentially lead to

strong seismic events. Already two major earthquakes have occurred in Paliki in 2014, where the maximum increasing rate of non-linear uplift was observed, compatible with that interpretation. The second, therefore, area, potentially relating to the possible occurrence of future earthquakes, should be taken into consideration.

As a final point, in southern Cephalonia, the seismic activity is relatively low, strain is accumulating, and the deformation showed a pattern similar to the observed one in the already affected area of Paliki. Based on these remarks, it may be assumed that in the future the seismic activity may be shifted toward the southern part of Cephalonia (onshore or even offshore), even though it seems that the seismic events in 2014–2015 have partially released the strain that was locally accumulated in the previous years, and the geodetic data from a number of cGPS stations in the south since 2014 do not currently indicate any change of the deformation pattern. Moreover, a regional offshore source of activity may not be ruled out, especially since the northern part of Zakynthos has also shown a non-linear uplift (Sakkas et al. 2014) and the offshore area southwest of the 2014 Cephalonian events has generated even stronger earthquakes in the past ($M_s \sim 7.0$ in 1983; Scordilis et al. 1985; Papadimitriou 1993). In addition, the possibility of a new activity in-between Cephalonia and Ithaca is also possible due to the presence of major tectonic features in the area.

Acknowledgements The authors would like to thank Dr. M. Keiding, an anonymous reviewer and the editor Prof. E. Papadimitriou for their constructive comments. Data from the cGPS stations SPAN, PONT, VLSM, ZAKY AGRI were provided by the Geodynamic Institute of National Observatory of Athens (NOA), Greece. GPS receivers were provided by the Hellenic Forestry Commission, Athens. Figures were created with Generic Mapping Tool s/w (Wessel and Smith 1998). Ithaca GPS benchmarks were installed by Dr. Brian Damiata (UCLA).

References

- Briole P, Elias P, Parcharidis I, Bignami C, Benekos G, Samsonov G, Kyriakopoulos C, Stramondo S, Chamot-Rooke N, Drakatos ML, Drakatos G (2015) The seismic sequence of January–February 2014 at Cephalonia Island (Greece): constraints from SAR interferometry and GPS. *Geophys J Int* 203(3):1528–1540. doi:10.1093/gji/ggv353
- Chousianitis K, Ganas A, Evangelidis CP (2015) Strain and rotation rate patterns of mainland Greece from continuous GPS data and comparison between seismic and geodetic moment release. *J Geophys Res B* 120(5):3909–3931. doi:10.1002/2014JB011762
- Chousianitis K, Konca AO, Tselentis GA, Papadopoulos GA, Gianniu M (2016) Slip model of the 17 November 2015 $M = 6.5$ Lefkada earthquake from the joint inversion of geodetic and seismic data. *Geophys Res Lett* 43(15):7973–7981. doi:10.1002/2016GL069764
- Dach R, Lutz S, Walser P, Fridez P (2015) Bernese GNSS Software version 5.2., available from <http://www.berne.unibe.ch/>. Accessed November 2015. doi:10.7892/boris.72297
- Ganas A, Marinou A, Anastasiou D, Paradissis D, Papazissi K, Tzavaras P, Drakatos G (2013) GPS-derived estimates of crustal deformation in the central and north Ionian Sea, Greece: 3-year results from NOANET continuous network data. *J Geodyn* 67:62–71. doi:10.1016/j.jog.2012.05.010
- Ganas A, Elias P, Bozionelos G, Papathanassiou G, Avallone A, Papastergios A, Valkaniotis S, Parcharidis I, Briole P (2016) Coseismic deformation, field observations and seismic fault of the 17 November 2015 $M = 6.5$, Lefkada Island Greece earthquake. *Tectonophysics* 687:210–222. doi:10.1016/j.tecto.2016.08.012
- Hackl M, Malsevrisi R, Wdowski S (2009) Strain rate patterns from dense GPS networks. *Nat Hazards Earth Syst Sci* 9:1177–1187. doi:10.5194/nhess-9-1177-2009
- Karakostas VG, Papadimitriou EE, Papazachos CB (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, Papadimitriou EE, Mesimeri M, Gkarlaoui C, Paradisopoulou P (2015) The 2014 Kefalonia Doublet ($M_w 6.1$ and $M_w 6.0$), central Ionian Islands, Greece: seismotectonic implications along the Kefalonia transform fault zone. *Acta Geophys* 63(1):1–16. doi:10.2478/s11600-014-0227-4
- Karastathis VK, Mouzakiotis E, Ganas A, Papadopoulos GA (2015) High-precision relocation of seismic sequences above a dipping Moho: the case of the January–February 2014 seismic sequence on Cephalonia island (Greece). *Solid Earth* 6(1):173–184. doi:10.5194/se-6-173-2015
- Lagios E, Sakkas V, Papadimitriou P, Damiata BN, Parcharidis I, Chousianitis K, Vassilopoulou S (2007) Crustal deformation in the Central Ionian Islands (Greece): results from DGPS and DInSAR analyses (1995–2006). *Tectonophysics* 444:119–145. doi:10.1016/j.tecto.2007.08.018
- Lagios E, Papadimitriou P, Novali F, Sakkas V, Fumagalli F, Vlachou K, Del Conte S (2012) Combined seismicity pattern analysis, DGPS and PSInSAR studies in the broader area of Cephalonia (Greece). *Tectonophysics* 524(525):43–58. doi:10.1016/j.tecto.2011.12.015
- Le Pichon X, Chamot Rooke N, Lallemand S, Noomen R, Veis G (1995) Geodetic determination of the kinematics of central greece with respect to Europe: implications for eastern mediterranean tectonics. *J Geophys Res* 100(B7):12675–12690. doi:10.1029/95JB00317
- Lekkas EL, Mavroulis SD (2015) Earthquake environmental effects and ESI 2007 seismic intensities of the early 2014 Cephalonia (Ionian Sea, western Greece) earthquakes (January 26 and February 3, M_w 6.0). *Nat Hazards* 78(3):1517–1544. doi:10.1007/s11069-015-1791-x
- Lekkas EL, Mavroulis SD (2016) Fault zones ruptured during the early 2014 Cephalonia Island (Ionian Sea, Western Greece) earthquakes (January 26 and February 3, M_w 6.0) based on the associated co-seismic surface ruptures. *J Seismolog* 20(1):63–78. doi:10.1007/s10950-015-9510-3
- Papadimitriou EE (1993) Focal mechanism along the convex side of the Hellenic Arc and its tectonic significance. *Boll Geof Teor Appl* 35(140):401–426
- Papadimitriou P, Kaviris G, Makropoulos K (2006) The $M_w = 6.3$ 2003 Lefkada Earthquake (Greece) and induced stress transfer changes. *Tectonophysics* 423:73–82. doi:10.1016/j.tecto.2006.03.003
- Papadopoulos GA, Karastathis V, Koukouvelas I et al (2015) The Cephalonia, Ionian Sea (Greece), sequence of strong earthquakes

- of January-February (2014): a first report. *Res Geophys* 4:5441. doi:[10.4081/rg.2014.5441](https://doi.org/10.4081/rg.2014.5441)
- Sakkas V, Lagios E (2015) Fault modelling of the early-2014 ~M6 Earthquakes in Cephalonia Island (W. Greece) based on GPS measurements. *Tectonophysics* 644:184–196. doi:[10.1016/j.tecto.2015.01.010](https://doi.org/10.1016/j.tecto.2015.01.010)
- Sakkas V, Novali F, Lagios E, Vassilopoulou S, Damiata BN, Fumagalli A (2014) Ground deformation of Zakynthos island (Western Greece) observed by PSI and DGPS. *Int Geosci Remote Sens Symp*. doi:[10.1109/IGARSS.2014.6947566](https://doi.org/10.1109/IGARSS.2014.6947566)(**art. no. 6947566**)
- Saltogianni V, Stiros S (2015) A two-fault model of the 2003 Leucas (Aegean Arc) earthquake based on topological inversion of GPS data. *Bull Seismol Soc Am* 105(5):2510–2520. doi:[10.1785/0120140355](https://doi.org/10.1785/0120140355)
- Scordilis EM, Karakaisis GF, Karakostas BG, Panagiotopoulos DG, Comninakis PE, Papazachos BC (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](https://doi.org/10.1007/BF00880738)
- Teza G, Pesci A, Galgaro A (2008) Grid_strain and grid_strain3: software packages for strain field computation in 2D and 3D environments. *Comput Geosci* 34(9):1142–1153. doi:[10.1016/j.cageo.2007.07.006](https://doi.org/10.1016/j.cageo.2007.07.006)
- van Hinsbergen DJJ, van der Meer DG, Zachariasse WJ, Meulenkaamp JE (2006) Deformation of western Greece during Neogene clockwise rotation and collision with Apulia. *Int J Earth Sci* 95:463–490. doi:[10.1007/s00531-005-0047-5](https://doi.org/10.1007/s00531-005-0047-5)
- Vassilakis E, Royden L, Papanikolaou D (2011) Kinematic links between subduction along the Hellenic trench and extension in the Gulf of Corinth, Greece: a multidisciplinary analysis. *Earth Planet Sci Lett* 303(1–2):108–120. doi:[10.1016/j.epsl.2010.12.054](https://doi.org/10.1016/j.epsl.2010.12.054)
- Wessel P, Smith WHF (1998) New, improved version of generic mapping tools released. *EOS Trans AGU* 79:579. doi:[10.1029/98EO00426](https://doi.org/10.1029/98EO00426)

Application of electrical resistivity imaging (ERI) for the assessment of peat properties: a case study of the Całowanie Fen, Central Poland

Sebastian Kowalczyk¹  · Kornelia Anna Żukowska¹ · Maciej Jan Mendecki² · Dominik Łukasiak¹

Received: 3 February 2017 / Accepted: 7 February 2017 / Published online: 1 March 2017
© The Author(s) 2017. This article is published with open access at Springerlink.com

Abstract Complex studies were carried out to recognize the fen structure and peat properties in the Całowanie Fen area, belonging to the Natura 2000 network. The studies were conducted in two study areas that differ significantly in terms of peat thickness. Electrical resistivity imaging (ERI) was used to identify the properties of the peat and its substrate, such as thickness and electrical resistivity. Comparison of the field studies with the laboratory tests has shown that the ash content rises electrical resistivity in peat. In addition, the study has shown that the application of non-invasive geophysical methods in protected areas is justified. The fen, as a medium containing mostly water, was a proper test area for the ERI measurements.

Keywords Peat · Electrical resistivity imaging (ERI) · Physical properties · Całowanie Fen

Introduction

Incomplete decomposition of plant debris under the influence of strong hydration results in peat formation. The process of peat accumulation is called sedimentation and it is related to the deposition of the material in the place of origin (Ingram 1978). Generally, in Poland dominate shallow fens that are fed by surface water and/or groundwater. The Całowanie peatland belongs to such fens; it is also sub-classified as soligenous being fed only by the

unconfined aquifer system. Furthermore, the Całowanie Fen is located in the Vistula River Valley in close proximity to a postglacial upland and it is an interesting object of study because of its geomorphological conditions. However, the fen location within the Mazowiecki Landscape Park and its inclusion within the Natura 2000 Network forced us to use non-invasive research methods that would not affect the environment significantly. In this case, the selected geophysical methods, beside geological drillings, are an excellent tool to study areas sensitive to human interference. Variation in the physical parameters of the peat and its substrate allow to apply several geophysical methods in the investigation of the vertical succession and recognition of hydrogeological conditions. Peatlands are investigated mostly using the ground penetrating radar (GPR) method. Only a few researchers (e.g., Slater and Reeve 2002; Comas et al. 2015; Walter et al. 2016) have applied induced polarization imaging, electromagnetic terrain conductivity and electrical resistivity imaging, apart from the GPR method, as alternative tools to determine the peatland thickness.

The primary objective of the presented investigations was to test the potential of electrical resistivity imaging (ERI) to analyze the variability of peat layer thickness and estimate the physical parameters of peat. Moreover, the studies were focused on assessing the measurement and interpretation methodology of ERI in such a difficult terrain as peatlands. Our intention was also to assess the geophysical data relation to geological information obtained from the drilling such as physical properties of the peat. The studies were conducted in two areas, located within one large peatland, that differ significantly in thickness and macroscopic properties of organic soils. The paper also showed that the resistivity imaging provides valuable information about the stratigraphic contacts. The

✉ Sebastian Kowalczyk
s.kowalczyk@uw.edu.pl

¹ Faculty of Geology, University of Warsaw, Warsaw, Poland

² Faculty of Earth Sciences, University of Silesia in Katowice, Sosnowiec, Poland

data obtained from the geophysical and geological studies have been correlated to provide a wider picture of sediments building the peatland and its substrate.

Peatlands are characterized by variability in thickness of the deposit on a very small area. Additionally, the legal protection which they are subject to very often hampers or precludes peat investigation using standard geological methods. Therefore, application of non-invasive geophysical methods is particularly useful in the study of peat. The resistivity imaging, *inter alia*, provides information about the thickness of the deposit along the test profile. They also provide information on the geological structure of the peatland substrate which is valuable data in determining the geological conditions and, moreover, in conjunction with hydrogeological studies, makes it possible to design a hydrodynamic model of the analyzed area.

Presently, the geoengineering investments are designed very often in areas where organic soils are present; therefore, the identification of the hydrogeological conditions as well as knowledge about the physical–mechanical properties of peats and its substrate are crucial issues, and the information obtained using geophysical methods can become more practical.

Study area

The Całowanie Fen, covering an area of 12 km², is one of the largest peatlands in the Masovia region. It is located to the south-east of Warsaw and lies almost entirely within the limits of the Masovian Landscape Park (MLP). It extends in an almost 15 km long and 3 km wide belt (Fig. 1).

The fen developed in the paleochannel of the Vistula River within the lower upper terrace called the Praga terrace. Beside peatlands, the study area contains a system of dunes within organic fen deposits. The dunes originated from the accumulation of fluvial deposits such as out-washes (e.g., Góra Niedźwiedzia) or shoal islands representing mineral ‘islands’ within the peatland (e.g., Kobyła Góra, Góra Pękatka), which are in fact initial dunes, 2–3 m high and 20–30 m wide, with a length up to several tens of meters (Baraniecka 1982).

The peat thickness within the Całowanie Fen is variable and depends on the morphology of the mineral substrate, generally decreasing to the north; thus, the largest thicknesses around 3–4 m are noted in the southern part of the peatland. In the northern part, the peat thickness does not exceed 1.3 m. The organic deposits are generally underlain by sand of the Vistula River lower upper terrace. Additionally, in several places the peat is underlain by gyttia or thin mud layers. The levelled surface of the peatland is

Fig. 1 Location of the study areas: **a** on a sketch-map of the Masovian Landscape Park; **b** on a fragment of the Detailed Geological Map of Poland in the scale of 1:50,000, Otwock sheet (Baraniecka 1975) and Osieck sheet (Sarnacka 1964); **c**, **d** on a fragment of a topographic map

characterized by a western dip, from 2.3‰ (in the south) to 4.4‰ (in the north).

Two survey areas characterized by different peat thickness were selected for the geophysical study. The first was in the Kobyła Góra area (Fig. 1c) where the peat originated on a mineral island and attained a small thickness on its rim. The second was in the Gole area (Fig. 1d); this site has the largest peat thickness in the entire Całowanie Fen. Kobyła Góra represents an incipient dune characterized by the height of 2.5–3 m and a meridional orientation reaching a maximum length of 255 m. The fen within the Kobyła Góra area is mostly overgrown by sedges, reeds (occurring in several places) and shrub vegetation growing along the drainage ditches. In contrast, the Gole area is flat and covered mainly by grass.

The peatland is generally supplied by groundwater from the first aquifer. Seasonal oscillations occur in the water table of the first aquifer, depending on the dominant water–soil and atmospheric conditions. Within the study areas, the depth of the water table varies from above the surface (western part of Kobyła Góra) to below the surface (Gole and the eastern part of Kobyła Góra).

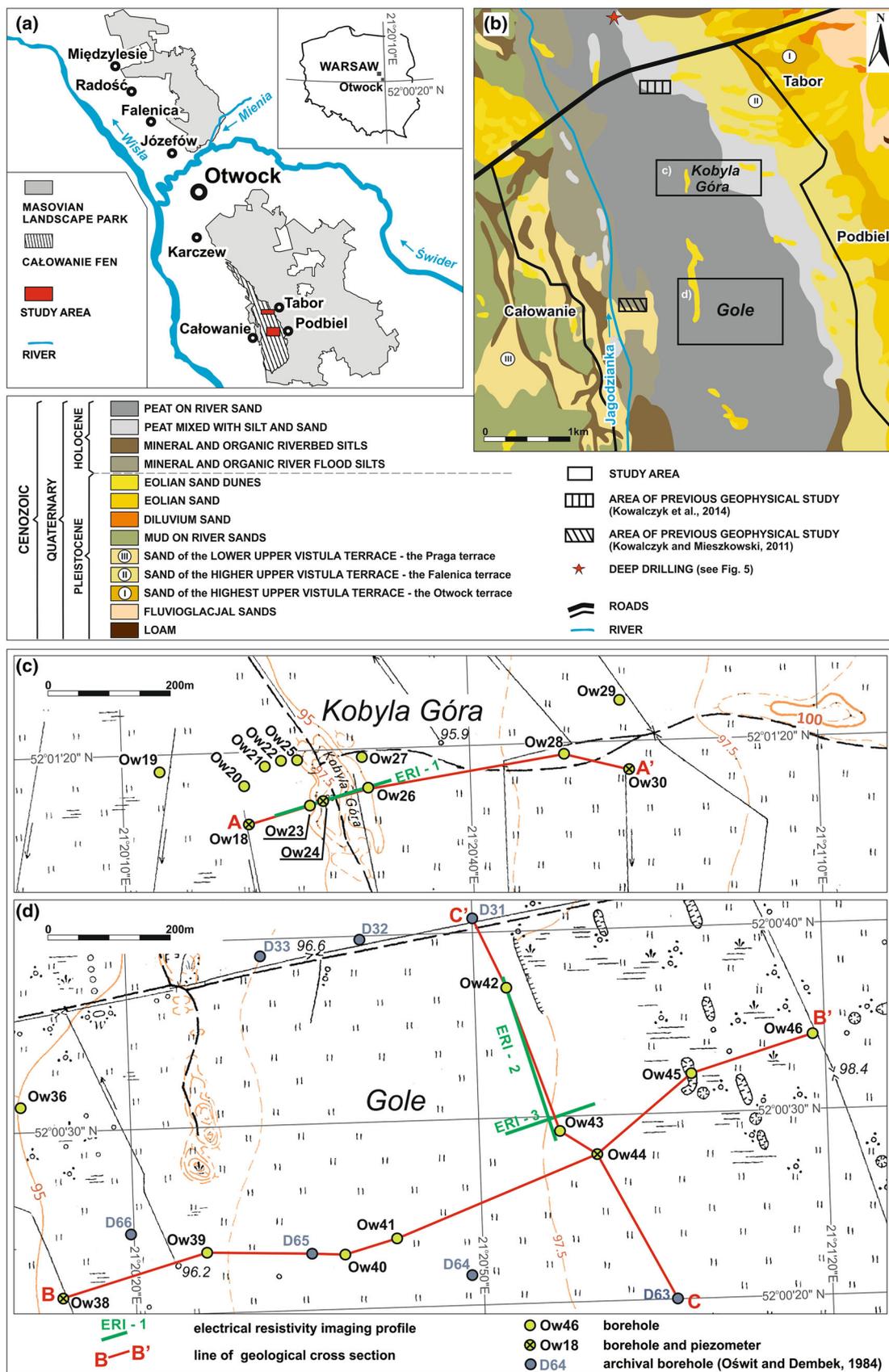
Methods

Geological survey

Fieldworks, including general observations, were focused on the analysis of the geological environment of the peatland and they covered following methods: drillings to show the vertical peat succession, macroscopic descriptions with particular attention drawn on the degree of peat decomposition according to the von Post scale (von Post 1922), groundwater level measurements, grain size composition of the coarse-grained soils, and studies of the physical and chemical parameters of water.

The fieldworks methodology was focused on identifying peat properties that depend on: the botanical composition of the peat-forming areas, the degree of peat decay, and hydrogeological conditions (peatland recharge, groundwater table depth, physical and chemical properties of water).

Within the Gole study area, 13 drillings were made using the Instorf probe; 60 samples of peat were collected from all drillings. In the Kobyła Góra study area, 15 drillings were made and 16 samples of peat were collected.



The collected samples were used for macroscopic determination of the peat type and the degree of decomposition based on the 10-point von Post scale (von Post 1922); these parameters were identified by hand squeezing the peat sample and observing the amount, color and volume of the extruded water, and the nature of plant remains in the hand. The peat samples were tested for: natural moisture content, ash content and pH value. The ash content value was determined by igniting the samples at 440 °C (ASTM D2974 2000; Myślińska 2001). The value of pH was tested in a suspension of peat and distilled water with a ratio of the solid phase to distilled water of 1:2.5 (Myślińska 2001). The pH measurements were carried out using the CX-401 multi-device connected to a pH-metric electrode.

The five piezometers were installed within the following study areas: three Kobyla Góra and two in the Gole area (one of them (Ow44) was filtered on three depths: 1.5, 3.0 and 4.2 m). Piezometers were used to measure the depth of the water table and to estimate the electrical conductivity of water (ECW), water pH, and to collect water samples for chemical analysis.

Electrical resistivity imaging

The resistivity method is one of the oldest geophysical survey techniques (Schlumberger et al. 1933; Archie 1942; Sumner 1976; Telford et al. 1990; Loke 2011) and is commonly applied in fields such as engineering and environmental investigations (e.g., Zogała et al. 2009; Pacanowski et al. 2014; Amini and Ramazi 2016), geological (e.g., Shemang and Molwalefhe 2009; Barski and Mieszkowski 2014; Mieszkowski et al. 2014) and hydrogeological (e.g., Asfahani and Abou Zakhem 2013; Giang et al. 2013) studies. The purpose of electrical surveys is to determine the subsurface resistivity distribution by performing measurements on the ground surface. From these measurements, true resistivity can be estimated by an appropriate inversion routine. Ground resistivity is related to various geological medium and it is clearly dependent on the moisture content, water content and ways of its distribution in the medium (porosity, degree of saturation, degree of cementing, fracturing), as well as mineral composition, structure, and texture of the soil or rock. The relationships between these factors for coarse-grained soils were described by Archie (1942). Electrical properties of cohesive soils are more complex due to the amount and type of clay minerals and are related to ionic, molar and electric conductivity, as well as presence of the electric double layer in clay particles (Waxman and Smith 1968), the degree of saturation, structure, and soil consolidation (Fukue et al. 1999). By analogy, peat electrical resistivity depends on water conductivity in the peat pores, moisture content and electric conductivity of the peat matrix that

depends on the organic and mineral content, mineral composition, degree of peat decomposition, structure and porosity.

Organic (peat), non-cohesive (river sands) and consolidated cohesive soils (loams and Pliocene clays) occur beneath the surface of the study. In consequence, the geology of the study area could be identified more precisely using the method of electrical resistivity imaging (ERI). Electrical resistivity imaging is also known as electrical resistivity tomography (ERT) (Loke 2011; Loke et al. 2013). The measurements were conducted using the Dipole–Dipole array with application of the Terrameter LS apparatus of the Swedish company ABEM, connected to 4 cables, each containing 21 electrodes, spaced every 2 m. When the length of the measuring line was longer than 160 m, the roll-along technique was applied.

Field data obtained from ERI measurements were interpreted using Res2DInv software (Loke 2001; Loke 2012; Loke and Barker 1996; Loke et al. 2003). Processing involves the inversion routine in which the most probable geo-electrical model is obtained. Raw data were converted to a Res2DInv format using the LS Terrameter Utilities Software. Later, the bad points were removed. The topography correction was included in the profile from Kobyla Góra. The best model was selected by successive calculations and subsequent comparisons (iterations) of the model and the data. The parameters of the processed data are presented in Table 1. Low values of the RMS error indicate a very good fit and high correlation between the result of apparent resistivity measurements and the estimated model of a “true” resistivity model. According to the theoretical assumptions of the ERI method, the resistivity values are assigned to an individual depth of the block centres forming a plane-parallel block grid. The inversion result can be presented as a block model scheme or an interpolated contour map with a color scale. Data interpolation allowed to obtain a quasi-continuous resistivity distribution. The geological interpretation was performed based on a block model scheme of resistivity distribution.

Results

Geological survey

In the Kobyla Góra study area, the peat thickness is up to 1.5 m. The top of the peat deposit is dominated by a medium decomposed (H5–H6) reed peat and sedge peat. Below is situated highly decomposed (H7) reed peat. In the peat base, wooden peat characterized by a small thickness (~0.25 m) is observed in several places (Fig. 2). The depth of the water table in the analyzed area ranges from

Table 1 Parameters of the electrical resistivity imaging survey

| Study area | Array | Electrode spacing (m) | Length of profile (m) | Total number of data points | Number of data points after removed bad points | Minimum – maximum apparent resistivity values | Iteration | RMS error (%) |
|-------------|---------------|-----------------------|-----------------------|-----------------------------|--|---|-----------|---------------|
| Kobyła Góra | Dipole–Dipole | 2.0 | 200 | 1794 | 1631 | 23 ÷ 13,757 | 7 | 3.3 |
| Gole (S–N) | Dipole–Dipole | 2.0 | 280 | 2999 | 2993 | 22 ÷ 103 | 7 | 0.7 |
| Gole (W–E) | Dipole–Dipole | 2.0 | 160 | 1467 | 1432 | 28 ÷ 107 | 7 | 0.6 |

0.2 m (measurements from 2013-10-01) below the ground surface in the eastern part (Ow30 borehole) to 0.15 m (2013-10-01) above the surface in the western part (Ow18 borehole). Due to the level of the groundwater table of the first aquifer and its small fluctuations, a moorsh layer is absent or insignificant in the western part of the peatland.

The southern part of the Całowanie Fen (Gole study area) is characterized by the largest thickness of the peat deposit up to the maximum value of 4.65 m (Ow43 borehole) and the taxonomic diversity of the peat profile is very low. The peat profile composed mainly of sedge peat is characterized by medium decomposition (H4–H5). Moreover, thin sand interbeds (about 0.05 m thick) were noted in several places within the sedge peat (Fig. 2). Highly decomposed (H7) moss peat is present in the base of the peat profile. Sedge-reed peat was drilled in the eastern part of the Gole study area, whereas sedge-moss peat with calcareous gyttia occurred (D63 borehole) in the southern part. In the Gole study area, the depth of the groundwater table varied from 0.22 m (Ow46 borehole) to 0.28 m (Ow44 borehole) (measured on 2013-10-01). Below the entire Całowanie Fen, the substrate consisted of medium- and fine-grained sands of the Praga terrace—the lower upper terrace of the Vistula River.

Based on the laboratory tests, domination of weakly acidic peat ($5 \leq \text{pH} < 7$) was noted. The analyzed peat samples showed slight pH variation. Within the Gole study area, the highest pH was obtained for mossy peat, and the lowest—for reed peat. The pH of sedge peat and reed peat from Kobyła Góra was slightly higher compared to the pH of the same peat type in Gole (Table 2).

In the analyzed area of the Całowanie Fen, reed peat had the highest ash content in Kobyła Góra; this resulted from the mineral material blown off the sand dune. Within the Gole study area, a significant increase in the mineral content was noted (Table 2) due to the presence of sand interbeds (Fig. 2; Ow44 and Ow42 boreholes) and calcareous gyttia (Fig. 2; Ow46 borehole) in the sedge peat.

Determinations of the natural moisture content have indicated that the peat samples taken from Kobyła Góra

have a significantly lower moisture content than the samples from the Gole study area (Table 2).

The collected data of electrical conductivity of water indicate significant diversity of water physical properties between these two studies areas. ECW in Kobyła Góra is more than two times lower than in the Gole area. Also, tests have shown temporal/seasonal changes of pH and conductivity of water (Table 3).

Resistivity results on the fen

Resistivity measurements carried out in the Kobyła Góra study area, where the peat sedented on a mineral island and its thickness ranges from 0 to 1.5 m, allowed a generalized investigation of this peat layer. Resistivity of this layer ranges within 50–70 Ωm , but in the nearest vicinity of the dune and on the western side of the fen the peat was characterized by higher values ($\sim 70 \Omega\text{m}$). This may be caused by the higher ash content of the peat that originated from the blown-off mineral material. On the eastern side of the dune, in the direction of the main axis of the fen, the peat had a resistivity value in the range of 40–50 Ωm (Fig. 3).

The small thickness of peat in Kobyła Góra and the relatively large electrode spacing (2 m) as compared to the thickness of the peat are the reasons why the ERI studies failed to identify the vertical variation of the physical properties within the peat profile.

In Gole, the peat usually had resistivity values in the range of 30–60 Ωm . This area has the greatest peat thickness in the entire Całowanie Fen. A much greater thickness of peat compared to electrode spacing allowed to identify the diversity of physical properties within the peat layer both in horizontal and vertical directions. Several zones with low resistivity values (15–30 Ωm) were identified within the peat layer—in the S–N cross-section between: 82 and 86 m, 202 and 216 m, and 230 and 232 m of the measurement profile, and in a W–E cross-section between: 68 and 74 m, and 104 and 106 m of the measurement profile (Fig. 4). In the W–E cross-section, a high resistivity

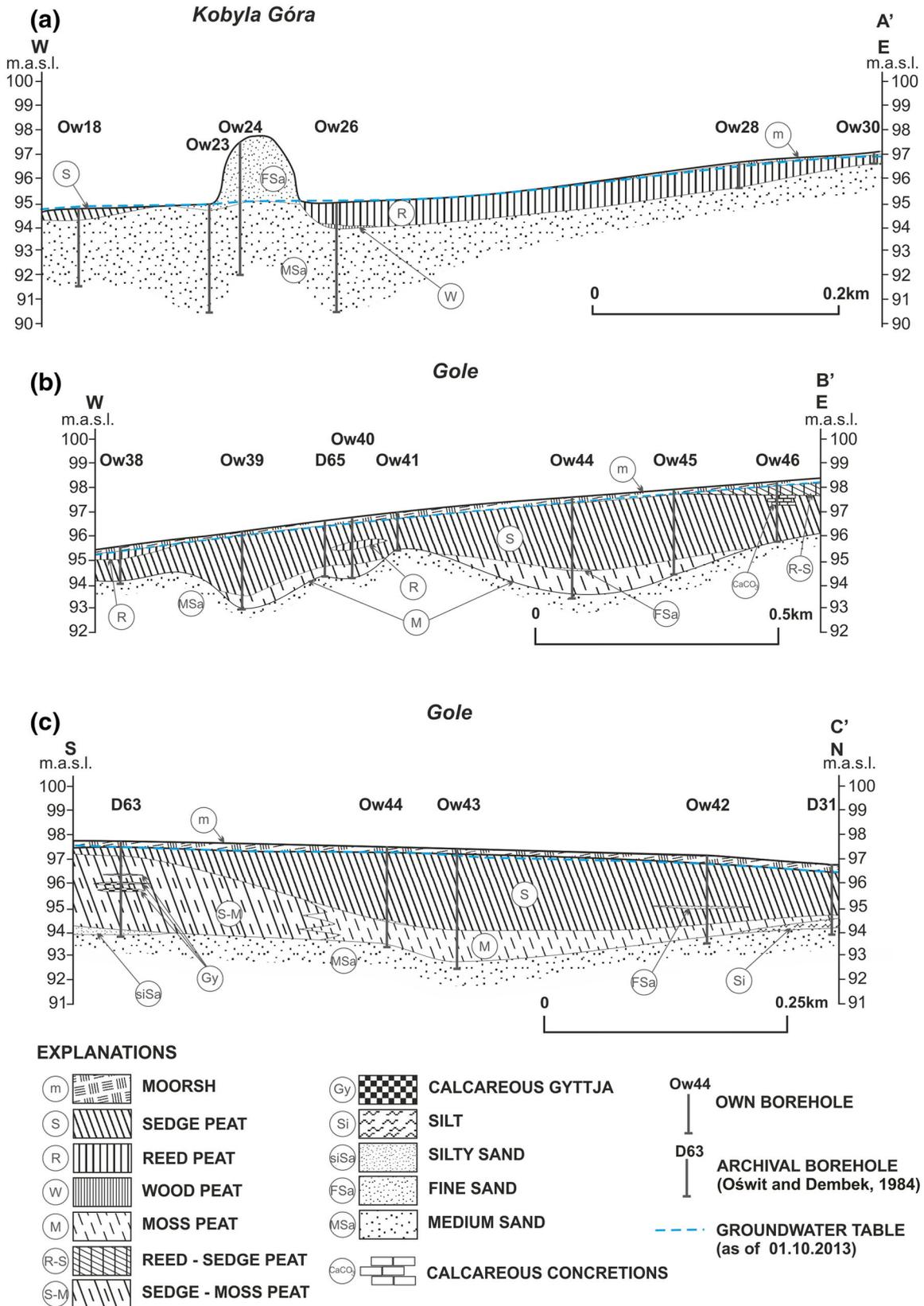


Fig. 2 Geological cross-sections (for location see Fig. 1)

Table 2 Results of laboratory tests: natural moisture content, ash content and pH of peats for the Kobyla Góra and Gole study areas

| Study area | Peat type | | Natural moisture content (%) | Ash content (%) | pH (–) |
|-------------|------------|---------------------------------|------------------------------|-----------------|-----------|
| Kobyla Góra | Reed peat | Average value | 457.12 | 47.16 | 5.62 |
| | | Range | 209.14–623.55 | 18.22–72.96 | 5.54–5.67 |
| | | Estimator of standard deviation | 218.91 | 21.66 | 0.05 |
| | | Number of measurements | 3 | 10 | 5 |
| | Sedge peat | Average value | 347.68 | 36.46 | 5.66 |
| | | Range | 230.32–441.66 | 22.71–47.07 | 5.54–5.81 |
| | | Estimator of standard deviation | 92.27 | 12.47 | 0.11 |
| | | Number of measurements | 5 | 3 | 6 |
| Gole | Moorsh | Average value | 502.87 | 15.15 | 5.43 |
| | | Range | 285.38–701.16 | 8.60–26.74 | 5.16–5.70 |
| | | Estimator of standard deviation | 131.24 | 6.95 | 0.20 |
| | | Number of measurements | 10 | 10 | 9 |
| | Sedge peat | Average value | 745.21 | 21.12 | 5.51 |
| | | Range | 529.80–738.03 | 7.23–80.93 | 4.90–6.60 |
| | | Estimator of standard deviation | 134.97 | 20.06 | 0.43 |
| | | Number of measurements | 33 | 33 | 36 |
| | Reed peat | Average value | 684.94 | 20.52 | 5.33 |
| | | Range | 499.11–906.93 | 7.74–85.46 | 4.99–5.60 |
| | | Estimator of standard deviation | 131.80 | 22.30 | 0.20 |
| | | Number of measurements | 9 | 9 | 10 |
| | Moss peat | Average value | 796.66 | 40.87 | 5.57 |
| | | Range | 389.78–1383.26 | 10.62–86.99 | 5.00–6.30 |
| | | Estimator of standard deviation | 343.27 | 26.62 | 0.46 |
| | | Number of measurements | 6 | 6 | 8 |

Table 3 Results of electrical conductivity and pH of water measured in the piezometers in both study areas

| Study area | Borehole no. | Filter depth (m) | pH (–) | | | ECW ($\mu\text{S}/\text{cm}$) | | |
|-------------|--------------|------------------|-----------------|----------------|----------------|---------------------------------|----------------|----------------|
| | | | 01. 10. 2013 | 08.03. 2014 | 12.04. 2014 | 01.10. 2013 | 08.03. 2014 | 12.04. 2014 |
| Gole | Ow38 | 1.5 | 7.42 | 6.64 | 6.26 | 739.0 | 606.8 | 588.0 |
| | Ow44 | 1.5 | 7.59 | 6.79 | 6.17 | 408.0 | 321.7 | 483.0 |
| | Ow44 | 3.0 | 7.39 | 6.71 | 6.01 | 623.0 | 554.1 | 559.0 |
| | Ow44 | 4.2 | 7.24 | 6.59 | 6.31 | 616.0 | 508.8 | 574.0 |
| Kobyla Góra | Ow18 | 3.0 | 6.83 | 6.38 | 6.51 | 219.0 | 218.3 | 190.5 |
| | Ow24 | 5.0 | 6.86 | 6.51 | 6.55 | 297.0 | 278.9 | 291.0 |
| | Ow30 | 1.0 | 7.11 | 6.11 | 6.37 | 214.0 | 175.1 | 157.7 |

anomaly was also identified between the 106 and 110 m directly near the surface.

Moreover, river sands and Pliocene clays were distinguished based on the obtained resistivity imaging (Figs. 3, 4). A loam layer (glacial till) probably occurs between these deposits. Lenses with higher resistivity values occurring within the river sands mainly consisted of a saturated deposit characterized by lower resistivity. These lenses are probably composed of gravel.

Discussion

The general outline of the geological structure in the study areas is known from the reports of Baraniecka (1976, 1982) and Sarnacka (1968, 1987). The thickness, type and physical properties of the peat from the Całowanie Fen are known from drilling surveys performed by Borówko-Dłużakowa (1961) (1 borehole), Oświt and Dembek (1984) (90 boreholes), Kowalczyk et al. (2014) (7 boreholes), and

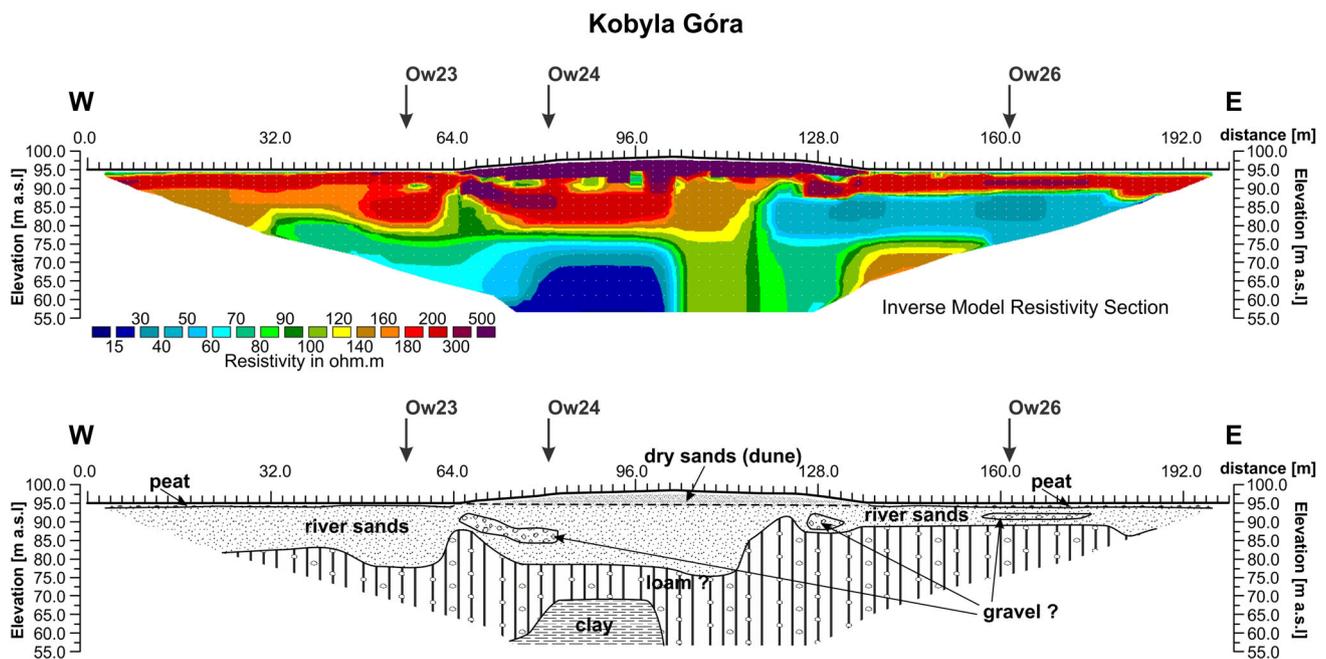


Fig. 3 Resistivity results for the Całowanie Fen in the Kobyla Góra study area: W–E pseudo cross-section

those presented in this paper. In the Całowanie Fen area, several geophysical surveys were conducted previously; they were focused on determining the basal surface of the organic deposits (Kowalczyk and Mieszkowski 2011; Kowalczyk et al. 2014). Application of geophysical methods such as ground penetrating radar, vertical electrical sounding, and refraction seismic have allowed for precise recognition of the basal surface of the organic soil and its lateral relationships with the surrounding layers (Kowalczyk and Mieszkowski 2011). However, these surveys were conducted on the margin of the Całowanie Fen and peat did not occur in the geological succession of the investigated areas; in turn, organic river flood silts covered by sands were noted. Thus, for a major part of the fen with an area of 12 km², there were no geophysical surveys characterizing the physical properties (such as electrical resistivity) of the peat. Kowalczyk et al. (2014) showed a successful application of GPR in peatland investigations and concluded that the method provides an excellent alternative to drilling and mapping of the stratigraphy and peat thickness. This is confirmed herein by a GPR survey in the central and eastern part of the Całowanie Fen. This survey allowed to: determine the boundary between the different peat types and distinguish the boundaries within one peat type characterized by variable physical properties resulting from the degree of decomposition. However, a clear impact of the botanical composition on the value of electrical resistivity was not recognized in the fen.

As it was mentioned in the Geological survey results section, peat in the Całowanie Fen is characterized by a

high taxonomic diversity and variation in the basic physical parameters such as ash content (Table 2) and degree of decomposition. The obtained values of ash content are higher than those indicated by Ilnicki (2002) as characteristic (5–20%) of fens. A high mineral content could be also caused by a sandier basal part of the peat profile due to strong geomorphological processes occurring during peat sedimentation (Domińczak and Okupny 2010; Rydelek 2013). Geological studies of peatlands indicate that peats often lie on a thin layer of organic sands of the early phase of the lacustrine sedimentation (Falkowska 2001). Usually, the basal zone of peatlands is strongly silted which is connected with the fact that this zone marks the beginning of peat accumulation on the mineral substrate and peat is strongly enriched in mineral content of substrate (Rydelek 2011; Rydelek et al. 2015).

The degree of peat decomposition affects its physical properties. An increase in the degree of peat decomposition causes a decrease in porosity, which results in increase in the bulk density and reduction of matrix density as well as natural moisture content (Hobbs 1986; Myślińska 1999, 2001). Laboratory resistivity investigations of Asadi and Huat (2009) have shown that the resistivity of peat decreased with the increasing degree of decomposition and moisture content and that peat resistivity increased with increasing organic content. These laboratory tests have not been explicitly confirmed by the field surveys presented in this paper. This results probably from the spatial variability of the peatland and influence of individual factors (such as moisture content, ash

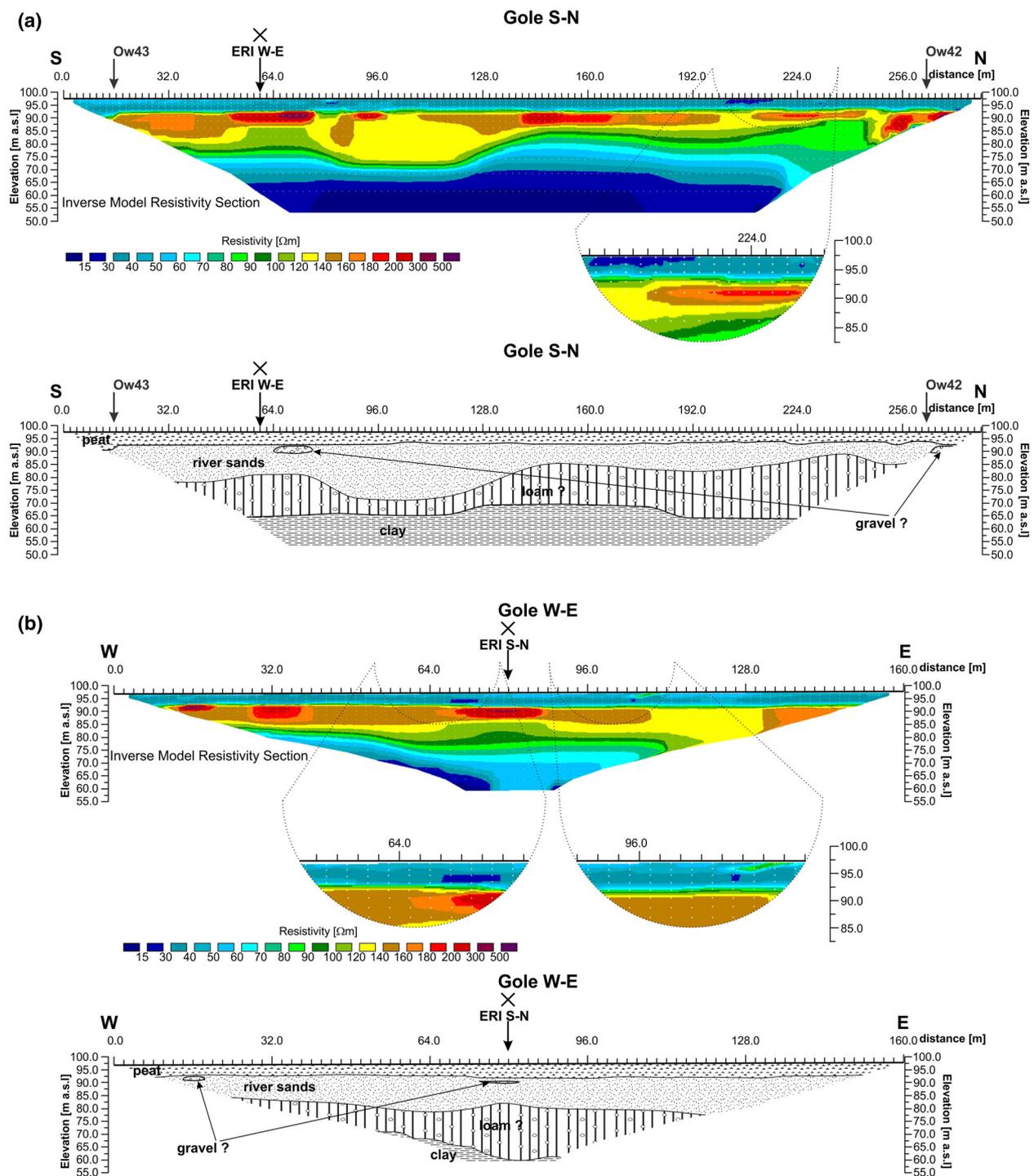


Fig. 4 Resistivity results for the Catowanie Fen in the Gole study area: **a** S–N pseudo cross-section; **b** W–E pseudo cross-section

content, degree of decomposition) on the final resistivity image obtained in the ERI survey.

The ERI method may be useful for determining the stratigraphic contact between the peat layer and the mineral substrate (Slater and Reeve 2002; Comas et al. 2015; Walter et al. 2016), which is also confirmed by this report.

The performed electrical resistivity survey proved its usefulness for mapping peat layers, regardless of their thickness. It is obvious that the resistivity imaging results are better where the peat thicknesses are larger. Moreover, differences in the physical properties within the peat layers can be also indicated. However, when the peat layer is

characterized by a small thickness, electrode spacing must be selected carefully, so that the peat layer will not be hidden or generalized. It is commonly known that the Dipole–Dipole array does not work properly during investigation of horizontal structures. However, the application of this array has yielded good results in the researches carried out in the Caribou Bog by Comas et al. (2004). Results of electrical resistivity imaging presented in this paper, obtained with a Dipole–Dipole array, are also promising.

The ERI survey in Gole has indicated the vertical and horizontal variation of the physical properties of peat. The diversity of physical properties in vertical profile is also visible in the water samples collected from three depths several times during the year from the filtered piezometer of the OW44 drilling site. The research has shown (Table 3) that when the measurement would have been made only in April, 2014, it would be possible to conclude that water conductivity increased with depth. However, previous investigations have disrupted this trend because the measurements of the water samples from the middle depth of the filtered piezometer (3.0 m) had the highest conductivity values in comparison to the others. ECW spatial variability was also noticed within the Kobyla Góra study area; as manifested in higher values of water conductivity in the dune (Ow24 borehole) than in the fen (Ow18, Ow30 boreholes) (Table 3). The reasons for this phenomenon probably lie in the high sorption properties of peat. Soil colloids are electronegative; therefore, cations such as Na^+ , K^+ , Ca^{2+} are subjected to sorption processes. Chemical analysis of the groundwater samples in Kobyla Góra seems to confirm this observation. Plant root system assimilation of some solutes (especially during spring and summer) and, to a smaller extent, rainfall infiltration in the dune may also have an impact on the decrease of groundwater conductivity in the peatland.

The variability of the properties in the peatland can manifest itself, as in the western side of the Kobyla Góra dune, by a higher resistivity associated with the presence of blown-off mineral material. ERI method did not provide accurate information on the substrate morphology and the thickness of peat contrary to the GPR survey. Detailed information from the GPR survey (Kowalczyk et al. 2014) would be very difficult to obtain using data solely from the drillings. The potential of GPR and ERI for estimating peat thickness was confirmed by Comas et al. (2015). However, these researchers indicate the limitations of geophysical methods in determining the thickness of the peat. Therefore, the studies require refinement in other test areas to develop an appropriate, non-invasive and inexpensive methodology.

A geological interpretation of the ERI measurements is hampered by the lack of a deep borehole within the

Całowanie Fen. However, a few deep drillings were made close to the fen area. One of such deep drillings was conducted near the village of Tabor, as shown in Fig. 5. Lithological data from the drill log were helpful in the interpretation of the resistivity model and increased the confidence of the generalized interpretation of soil types based on the resistivity values. The spatial identification of the geological structure can be used in further studies on the formation of the palaeo-riverbed and referring these analyses to contemporary processes influencing the morphology of the modern Middle Vistula riverbed (Falkowski 2007a). Geomorphological features of the terrace surface may be used in the engineering-geological assessment for the correct location of regulating structures and their protection. These features can be regarded as indicative in searching for protrusions of the alluvial substrate composed of deposits resistant to erosion in river valleys of lowland areas with a postglacial relief (Falkowski 2007b).

Conclusion

Application of the ERI method allowed to obtain 2D quasi-continuous images of the geological structure, which is essential for a correct interpretation of the geological phenomena occurring currently and in the past. An example is the spatial identification of the basal surface of the river sands (Figs. 3, 4) which provides information on the surface water flow during the formation of the valley before peat sedentation. In protected areas, where low invasiveness of research methods is necessary, application of geophysical methods is the best way to obtain information about the ground variability. Not only did the presented investigations allow to obtain data on the distribution of peat and its properties but also allowed for indirect understanding of the deeper geological structure in which the discrepancy in the thickness estimation based on geophysical methods increases with the depth of the prospecting.

The most accurate method to show the spatial distribution of peat basal surface occurring below the ground surface is the GPR method. Both the GPR and ERI methods enable to obtain quasi-continuous information on the variability of the physical properties of peat.

The performed pieces of research on the Całowanie Fen demonstrate significant heterogeneity of the peat in terms of physical properties such as moisture content, ash content, degree of decomposition and electrical resistivity.

The studies presented in the paper have both a cognitive and practical aspect. Presently, engineering investments are more commonly designed in areas with organic soils. The presented distribution of resistivity in relation to the physical properties of peat may become a useful tool in the

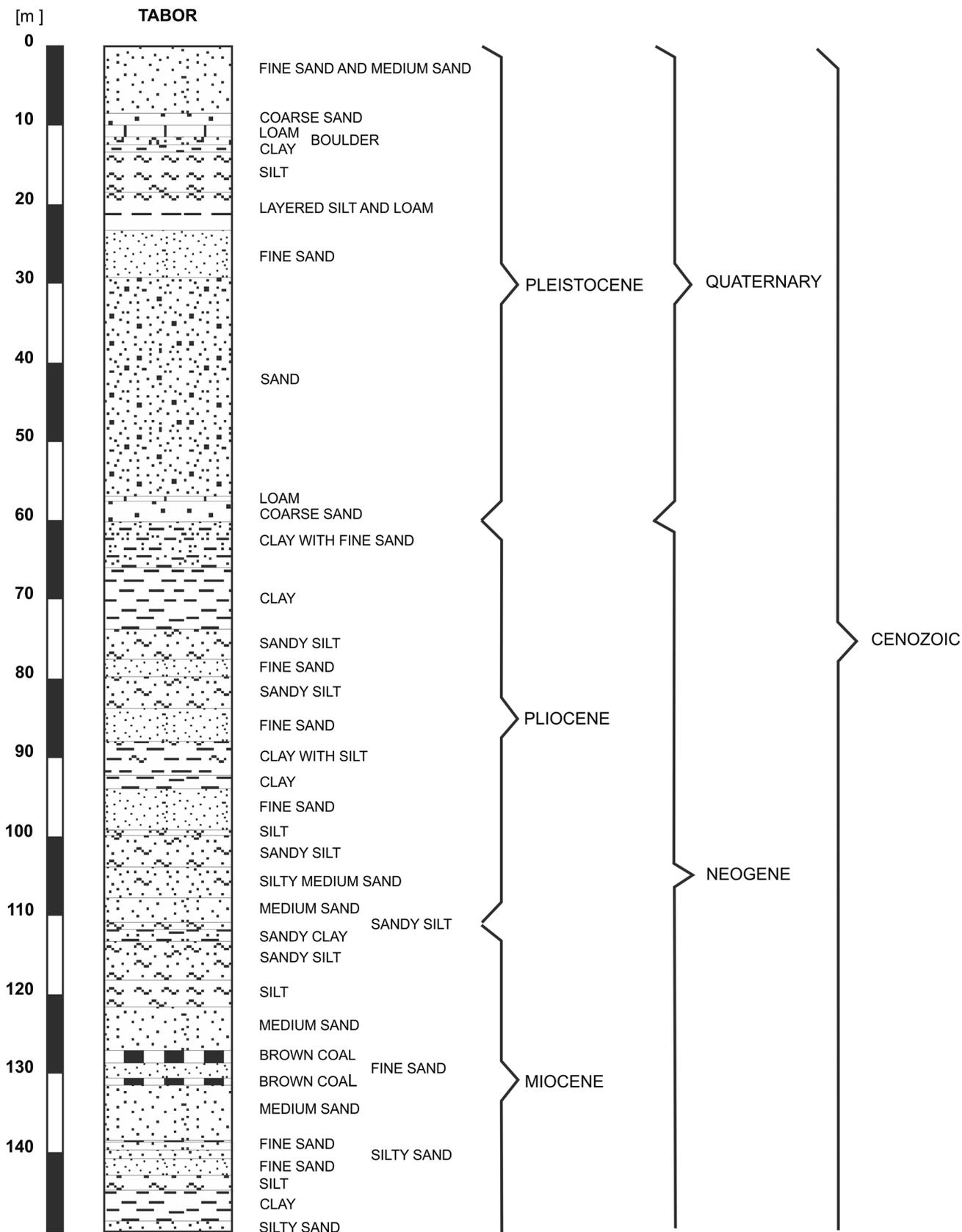


Fig. 5 Lithological column of the archival deep drilling near the area of Tabor (Baraniecka 1972)

precise recognition of soils with low bearing capacity in other areas, which in turn may influence the geological engineering assessment of the ground.

Acknowledgements The geological and electrical resistivity imaging surveys were supported financially by the scientific projects DSM-102930 and DSM-105528 of the Faculty of Geology, University of Warsaw. Terrameter LS apparatus, produced by Swedish company ABEM, was purchased as a part of the Regional Operational Program project entitled: “Modernization and equipment of the laboratories of the Faculty of Geology, University of Warsaw, to conduct research essential for Mazovia in the field of environment geoengineering - STAGE I, no. RPMA.01.01.00- 14-011/10”. The authors would like to thank the Reviewers as well as the Editor for their valuable comments and suggestions.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Amini A, Ramazi H (2016) Application of electrical resistivity imaging for engineering site investigation. A case study on prospective hospital site, Varamin, Iran. *Acta Geophys* 64:2200–2213. doi:[10.1515/acgeo-2016-0100](https://doi.org/10.1515/acgeo-2016-0100)
- Archie GE (1942) The electrical resistivity log as an aid in determining some reservoir characteristics. *Trans AIME* 146:54–62
- Asadi A, Huat BBK (2009) Electrical resistivity of tropical peat. *Electron J Geotech Eng* 14:1–9
- Asfahani J, Abou Zakhem B (2013) Geoelectrical and hydrochemical investigations for characterizing the salt water intrusion in the Khanasser valley, northern Syria. *Acta Geophys* 61:422–444. doi:[10.2478/s11600-012-0071-3](https://doi.org/10.2478/s11600-012-0071-3)
- ASTM D2974–00 (2000) Standard test methods for moisture, ash, and organic matter of peat and other organic soils. doi:[10.1520/D2974](https://doi.org/10.1520/D2974)
- Baraniecka MD (1972) Karta otworu wiertniczego Tabor-otwór nr 6 CAG PIG Warszawa, nr w archiwum 116889, nr CBDG 5265
- Baraniecka MD (1975) Szczegółowa Mapa Geologiczna Polski w skali 1:50,000; arkusz Otwock (561). Wydawnictwa Geologiczne, Warszawa **(in Polish)**
- Baraniecka MD (1976) Objasnienia do Szczegółowej Mapy Geologicznej Polski w skali 1:50 000 arkusz Otwock (561). Wydawnictwa Geologiczne, p 102 **(in Polish)**
- Baraniecka MD (1982) Dunes of the Otwock region against a geologic structure. *Biuletyn Instytutu Geologicznego* 337:5–32 **(in Polish with English summary)**
- Barski M, Mieszkowski R (2014) Upper Jurassic large-scale debris flow deposits in interbiohermal basins of the sponge megafacies in Poland—new insights. *Neues Jahrbuch für Geologie und Paläontologie Abhandlungen* 272:47–59. doi:[10.1127/0077-7749/2014/0396](https://doi.org/10.1127/0077-7749/2014/0396)
- Borówko-Dłużakowa Z (1961) Przyczynę do znajomości torfowisk holocenkich na tarasach Wisły pod Warszawą. *Arch. Instytut Geologiczny, Warszawa* **(in Polish)**
- Comas X, Slater L, Reeve A (2004) Geophysical evidence for peat basin morphology and stratigraphic controls on vegetation observed in a Northern Peatland. *J Hydrol* 295:173–184
- Comas X, Terry N, Slater L, Warren M, Kolka R, Kristiyono A, Sudiana N, Nurjaman D, Darusman T (2015) Imaging tropical peat lands in Indonesia using ground-penetrating radar (GPR) and electrical resistivity imaging (ERI): implications for carbon stock estimates and peat soil characterization. *Biogeosciences* 12:2995–3007. doi:[10.5194/bg-12-2995-2015](https://doi.org/10.5194/bg-12-2995-2015)
- Dominićzak P, Okupny D (2010) Spatial variability of selected physicochemical properties of biogenic sediments in the Kopanicha peatland near Skierniewice. *Prace Geograficzne* 123:99–110 **(in Polish with English summary)**
- Falkowska E (2001) Regularities in the occurrence of protection zones in polygenetic river valleys from the eastern part of the Polish Lowlands. *Acta Geol Pol* 51:163–192
- Falkowski T (2007a) The application of geomorphological analysis of the Vistula River, Poland in the evaluation of the safety of regulation structures. *Acta Geologica Polonica* 57:377–390
- Falkowski T (2007b) Geomorphologic analysis of the river terrace as an element in the geological engineering prognosis of the regulation structures safety. *Geologos* 11:163–172 **(in Polish with English summary)**
- Fukue M, Minatoa T, Horibe H, Taya N (1999) The micro-structure of clay given by resistivity measurements. *Eng Geol* 54:43–53
- Giang NV, Duan NB, Thanh L, Hida N (2013) Geophysical techniques to aquifer locating and monitoring for industrial zones in North Hanoi, Vietnam. *Acta Geophys* 61:1573–1597. doi:[10.2478/s11600-013-0147-8](https://doi.org/10.2478/s11600-013-0147-8)
- Hobbs NB (1986) Mire morphology and the properties and behaviour of some foreign peats. *Q J Eng Geol Hydrogeol* 19:7–80. doi:[10.1144/GSL.QJEG.1986.019.01.02](https://doi.org/10.1144/GSL.QJEG.1986.019.01.02)
- Ilnicki P (2002) Peatlands and peat. *Wyd. AR, Poznań*, p 606 **(in Polish with English summary)**
- Ingram HAP (1978) Soil layers in mires: function and terminology. *J Soil Sci* 29:224–227
- Kowalczyk S, Mieszkowski R (2011) Determination of a bottom layer of organic soil using geophysical methods at two sites on the Polish Lowland. *Biuletyn Państwowego Instytutu Geologicznego* 446:191–198 **(in Polish with English summary)**
- Kowalczyk S, Łukasiak D, Żukowska K (2014) Ground penetrating radar survey in the central and eastern part of the Całowanie Fen, Central Poland. In: *Proceedings of 15th international conference on ground penetrating radar, Brussels, Belgium, June 30–July 4 2014*, pp 1041–1046. doi:[10.1109/ICGPR.2014.6970579](https://doi.org/10.1109/ICGPR.2014.6970579)
- Loke MH (2001) Electrical imaging surveys for environmental and engineering studies, a practical guide to 2-D and 3-D surveys: RES2DINV manual, pp 1–65
- Loke MH (2011) Electrical resistivity surveys and data interpretation. In: Gupta H (ed) *Solid earth geophysics encyclopaedia*, 2nd ed, “Electrical & Electromagnetic”. Springer, New York, pp 276–283
- Loke MH (2012) Rapid 2-D Resistivity & IP inversion using the least-squares method (RES2DINV ver. 4.01 for Windows XP/Vista/7, manual), Geotomo Software, pp 1–169
- Loke MH, Barker RD (1996) Rapid least squares inversion of apparent resistivity pseudosections by a quasi-Newton method. *Geophys Prospect* 44:131–152
- Loke MH, Acworth I, Dahlin T (2003) A comparison of smooth and blocky inversion methods in 2D electrical imaging surveys. *Explor Geophys* 34:182–187. doi:[10.1071/EG03182](https://doi.org/10.1071/EG03182)
- Loke MH, Chambers JE, Rucker DF, Kuras O, Wilkinson PB (2013) Recent developments in the direct-current geoelectrical imaging method. *J Appl Geophys* 95:135–156. doi:[10.1016/j.jappgeoe.2013.02.017](https://doi.org/10.1016/j.jappgeoe.2013.02.017)
- Mieszkowski R, Dzierżek J, Stańczuk D (2014) Application of geoelectrical profiling in the delineation of shallow periglacial structures on the Drohiczyn Plateau. *Studia Quaternaria* 31:73–81. doi:[10.2478/squa-2014-0007](https://doi.org/10.2478/squa-2014-0007)

- Myślińska E (1999) Parametry fizyczne torfów i ocena metod ich oznaczania [Physical properties of peats and evaluation of methods of their determination]. *Przegląd Geologiczny* 47:676–682 (in Polish with English summary)
- Myślińska E (2001) Grunty organiczne i laboratoryjne metody ich badania, Wydawnictwo naukowe PWN, Warszawa, p 208 (in Polish)
- Oświt J, Dembek W (1984) Ekspertyza przyrodniczo - łąkarska obiektu Całowanie - Podbiel. IMUZ, Zakład Przyrodniczych Podstaw Melioracji, Falenty (in Polish)
- Pacanowski G, Czarniak P, Bąkowska A, Mieszkowski R, Welc F (2014) The role of geophysical ERT method to evaluate the leakproofness of diaphragm wall of deep foundation trenches on the example of the construction of retail and office complex in Lublin, Poland. *Studia Quaternaria* 31:91–99. doi:10.2478/squa-2014-0009
- Rydelek P (2011) Peatlands of Lubartów upland as the potential natural geological barriers. *Biuletyn Państwowego Instytutu Geologicznego* 446:407–416 (in Polish with English summary)
- Rydelek P (2013) Origin and composition of mineral constituents of Fen Peats from Eastern Poland. *J Plant Nutr* 36:911–928
- Rydelek P, Bąkowska A, Zawrzykraj P (2015) Variability of horizontal hydraulic conductivity of fen peats from Eastern Poland in relation to function of peatlands as a natural geological barriers. *Geol Quart* 59:426–432. doi:10.7306/gq.1180
- Sarnacka Z (1964) Szczegółowa Mapa Geologiczna Polski w skali 1:50,000; arkusz Osieck (598). Wydawnictwa Geologiczne, Warszawa (in Polish)
- Sarnacka Z (1968) Objasnienia do Szczegółowej Mapy Geologicznej Polski 1: 50,000. Arkusz Osieck (598), Wydawnictwa Geologiczne (in Polish)
- Sarnacka Z (1987) Evolution of the Vistula valley between the outlets of Radomka and Świder rivers in the Late Glacial and Holocene. *Geogr Stud Special Issue no. 4*:131–150
- Schlumberger C, Schlumberger M, Leonardon EG (1933) A New contribution to subsurface studies by means of electrical measurements in drill holes. *Am Inst Min Eng Tech Pub* 503:1–18
- Shemang EM, Molwalefhe LN (2009) DC resistivity and seismic refraction survey across the SE margin of Lake Ngami, NW Botswana. *Acta Geophys* 57:728–742. doi:10.2478/s11600-009-0017-6
- Slater LD, Reeve A (2002) Investigating peatland stratigraphy and hydrogeology using integrated electrical geophysics. *Geophysics* 67:365–378. doi:10.1190/1.1468597
- Sumner JS (1976) Principles of induced polarization for geophysical exploration. Elsevier, Amsterdam
- Telford WM, Geldart LP, Sheriff RE (1990) Applied geophysics, 2nd edn. Cambridge University Press, Cambridge, p 792
- Von Post L (1922) Sveriges Geologiska Undersöknings torvinentering och nogra av dess hittills vunna resultat (SGU peat inventory and some preliminary results). *Svenska Mosskulturforeningens Tidsskrft* 36:1–37
- Walter J, Hamann G, Lück E, Kligenfuss C, Zeitz J (2016) Stratigraphy and soil properties of fens: geophysical case studies from northeastern Germany. *Catena* 142:112–125. doi:10.1016/j.catena.2016.02.028
- Waxman MH, Smith LJM (1968) Electrical conductivities in oil-bearing shaly sand. *Soc Pet J* 8:107–122
- Zogała B, Dubiel R, Zuberek WM, Rusin-Zogala M, Steininger M (2009) Geoelectrical investigation of oil contaminated soils in former underground fuel base: Borne Sulnowo, NW Poland. *Environ Geol* 58:1–9. doi:10.1007/s00254-008-1458-y

A different method for interpretation of magnetic anomalies due to 2-D dipping dikes

İbrahim Kara¹ · Oya Tarhan Bal¹ · Anisya B. Tekkeli¹ · Gökhan Karcioğlu¹

Received: 8 February 2017 / Accepted: 14 February 2017 / Published online: 1 March 2017
© Institute of Geophysics, Polish Academy of Sciences & Polish Academy of Sciences 2017

Abstract In this study a new method is presented to determine model parameters from magnetic anomalies caused by dipping dikes. The proposed method is applied by employing only the even component of the anomaly. First, the maximum of the even component is divided to its value at any distance x in order to obtain $S1$. Then, theoretical even component values are computed for the minimal depth (h) and half-width (b) values. $S2$ is obtained by dividing their maximum to the value computed for the same distance x . A set of $S2$ values is calculated by slowly increasing the half-width, and h and b for the $S2$ closest to $S1$ are determined. The same procedure is repeated by increasing the depth. The determined b values are plotted against the corresponding values of h . After repeating the process and plotting curves for different distances, it is possible to determine the actual depth and half-width values.

Keywords Magnetic interpretation · Dikes · Even component

Introduction

Interpretation of magnetic anomalies of dikes is often used in mining and oil exploration geophysics. Several researchers interpret these anomalies by separating the anomalies into even and odd components which are origin-symmetric.

Among these researchers, Hutchinson (1958) used logarithmic curve fitting, Bhimasankaram et al. (1978) used Fourier transform method, and Kara et al. (1996) and Kara (1997) used correlation factors and integration nomograms on the even and odd components. Rao et al. (1973) developed two different methods, applied using the horizontal derivative of the anomaly. Atchuta Rao and Ram Babu (1981) used nomograms, depending whether the anomaly is due to a dike or a fault, for interpretation using the maximum and minimum values of the anomaly.

Additionally, Murthy (1985) applied the midpoint method, and Mohan et al. (1982) and Ram Babu and Atchuta Rao (1991) applied the Hilbert transform method on magnetic anomalies. Atchuta Rao et al. (1981) used complex gradient method, and Ram Babu et al. (1982) derived relation diagrams. Abdelrahman and Essa (2005) estimated depth and shape factors for simple geological models from magnetic data using least squares method. Abdelrahman et al. (2007) developed a least squares approach to determine the depth and width of a thick dipping dike using filters of successive window lengths. Abdelrahman and Essa (2015) developed a method to recover depth and shape properties of simple geological models from second derivative anomalies. Abo-Ezz and Essa (2016) linearized magnetic anomaly formula for simple geological models and estimated model parameters using least squares method. Essa and Elhussein (2017) estimated all the model parameters of a dipping dike by calculating second horizontal gradient anomalies using successive window lengths and minimizing the misfit between observed and predicted data. Essa (2007) obtained the shape factors and depth values for spheres, horizontal cylinders, and vertical infinite cylinders using a procedure similar to the method proposed in this paper.

In this study, a set of curves is obtained using the even component, each by dividing the maximum of the even

Prof. İbrahim Kara: Retired from Istanbul University.

✉ Gökhan Karcioğlu
gkarci@istanbul.edu.tr

¹ Istanbul University, Istanbul, Turkey

component to its value for a different distance x . Depth and half-width values of the dike are determined from the intersection point of these curves. The proposed method is tested on a synthetic model and thereafter applied to field data.

Method

For a Cartesian coordinate system with the Y axis showing the strike of the anomalous structure and X axis showing the profile direction, the magnetic anomaly (ΔF) of a 2D infinite vertical dipping dike (Fig. 1) at any point $P(x)$ is given by the following formula (Parker Gay 1963):

$$\Delta F = C \left\{ \cos Q \left[\tan^{-1} \left(\frac{x+b}{h} \right) - \tan^{-1} \left(\frac{x-b}{h} \right) \right] + \frac{1}{2} \sin Q \ln \left[\frac{(x+b)^2 + h^2}{(x-b)^2 - h^2} \right] \right\} \tag{1}$$

where x is the distance between the observation point and the origin, h is the depth to the top of the dike, b is the half-width, θ is the dipping angle, k is the susceptibility contrast, and T is the normalized total field strength. C and

Q denote the amplitude coefficient and index parameter, respectively, and their values for total, vertical, and horizontal field are given in Table 1. $I = \tan^{-1}(\tan i / \cos \alpha)$, where i is the geomagnetic inclination in the study area, and α is the azimuth of the profile according to the magnetic north (Figs. 2, 3).

Since Eq. 1 may be given as a summation of even and odd functions, it can be re-written as

Table 1 Amplitude coefficient and index parameter values for different magnetic field components

| Anomaly in | Amplitude coefficient, C | Index Parameter, Q |
|------------------|--|--------------------------|
| Total field | $2kT \sin \theta (1 - \cos^2 i - \sin^2 \alpha)$ | $2I - \theta - 90^\circ$ |
| Vertical field | $2kT \sin \theta (1 - \cos^2 i \sin^2 \alpha)^{1/2}$ | $I - \theta$ |
| Horizontal field | $2kT \sin \theta \sin \alpha (1 - \cos^2 i \sin^2 \alpha)^{1/2}$ | $I - \theta - 90^\circ$ |

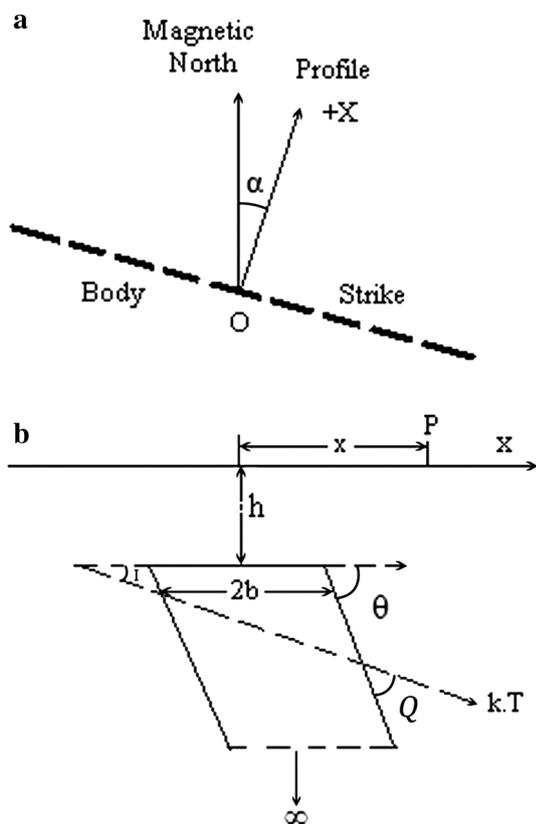


Fig. 1 a Plan view of a 2D body showing the direction of the profile and magnetic north. b Cross-sectional view of a 2D dike

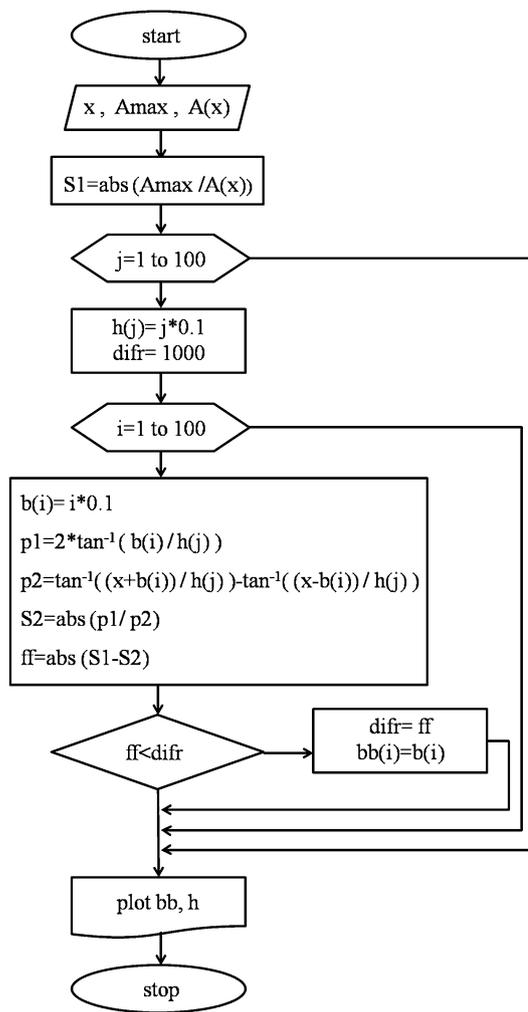


Fig. 2 Flowchart of the proposed method for a single curve. The process should be repeated for different values of x , in order to obtain a set of curves

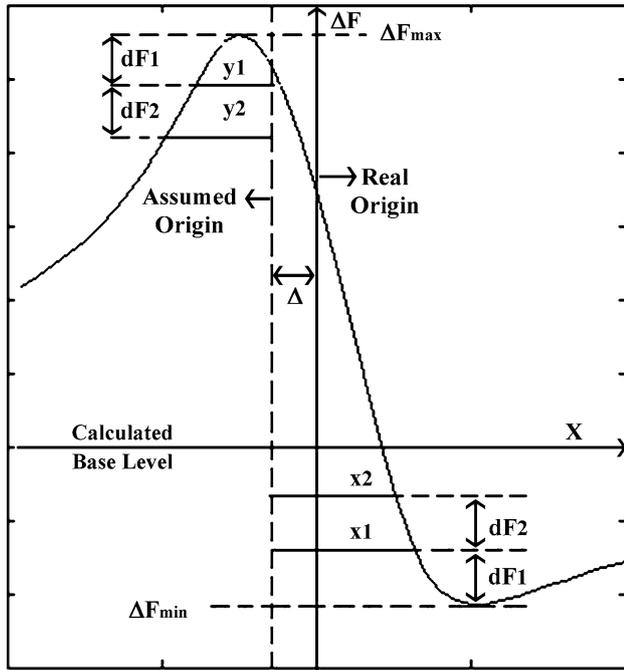


Fig. 3 Locating origin of a magnetic anomaly

$$\Delta F(x) = A(x) + B(x), \tag{2}$$

where

$$A(x) = C \cdot \cos Q \left[\tan^{-1} \left(\frac{x+b}{h} \right) - \tan^{-1} \left(\frac{x-b}{h} \right) \right] \text{ (even)} \tag{3a}$$

$$B(x) = \frac{1}{2} C \cdot \sin Q \cdot \ln \left[\frac{(x+b)^2 + h^2}{(x-b)^2 + h^2} \right] \text{ (odd)}, \tag{3b}$$

Anomalies of $\Delta F(x)$, $A(x)$, and $B(x)$ may be seen in Fig. 4 for the synthetic example.

$A(x)$ takes its maximal value for $x = 0$ and can be expressed as

$$A(0) = A_{\max} = 2 \cdot C \cdot \cos Q \cdot \tan^{-1} \left(\frac{b}{h} \right) \tag{4}$$

After dividing the maximal value of the even component by its value for any x , left and right sides of the equation may be given separately as

$$S1 = \frac{A(0)}{A(x)} \tag{5a}$$

and

$$S2 = \frac{2 \cdot \tan^{-1} \left(\frac{b}{h} \right)}{\tan^{-1} \left(\frac{x+b}{h} \right) - \tan^{-1} \left(\frac{x-b}{h} \right)} \tag{5b}$$

Variable $S1$ is obtained from the field data, by dividing the maximum of the even component to its value at any x . The first set of $S2$ values is calculated by assigning a minimal

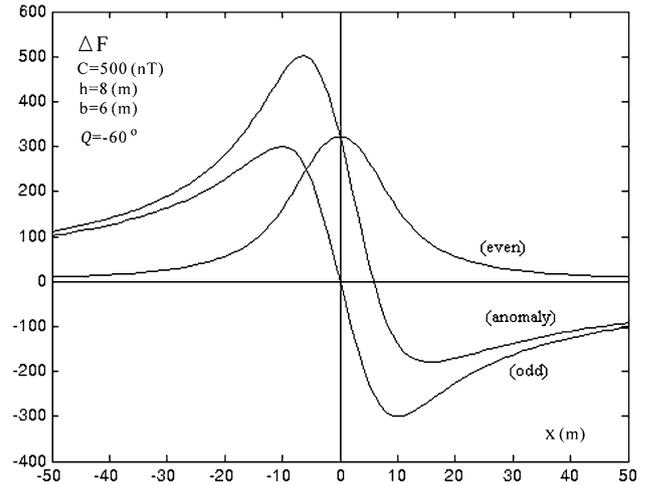


Fig. 4 Magnetic anomaly and its even and odd components calculated for the theoretical example

value for the depth (h), and slowly increasing the half-width (b), starting from a minimal value until it reaches its maximum possible value. Different sets of $S2$ are calculated by slowly increasing the value of h to a maximal value. For each set of $S2$, values of h and b yield to the $S2$ closest to $S1$ are obtained. These values of h and b are plotted as a curve by assigning h to the horizontal axis and b to the vertical axis. The flowchart of the proposed method is given in Fig. 2. Even though the obtained curve depicts a non-uniqueness, after repeating the process for different distances, and plotting resulting curves (each depicting a non-uniqueness) on the same graph, the intersection point of these curves yields to the actual values of h and b .

Location of the origin and the base level

To obtain even and odd components of an anomaly due to a dike, the location of the origin and the base level should be determined in prior. Powell (1967) proposed a method aiming at this need. Accordingly, if the origin is in its actual location, there applies the following equation:

$$x_1 \cdot y_1 = x_2 \cdot y_2$$

and if it is not, the equation becomes (Fig. 3),

$$(x_1 + \Delta) \cdot (y_1 - \Delta) = (x_2 + \Delta) \cdot (y_2 - \Delta)$$

When this equation is solved for Δ ,

$$\Delta = \frac{x_2 \cdot y_2 - x_1 \cdot y_1}{x_2 - x_1 - y_2 + y_1} \tag{6}$$

is obtained. The base level is obtained from the equation (Koulomzine et al. 1970)

$$\Delta F(0) = \Delta F_{\max} + \Delta F_{\min} \tag{7}$$

where $\Delta F(0)$ is the value of the anomaly at the origin.

Following the determination of the origin and the base level, the even and odd components are calculated using the equations

$$A(x) = \frac{1}{2} [\Delta F(x) + \Delta F(-x)] \text{ (even)} \tag{8a}$$

and

$$B(x) = \frac{1}{2} [\Delta F(x) - \Delta F(x)] \text{ (odd)} \tag{8b}$$

Once the proposed method is implemented on the even component, and the depth (h) and half-width (b) parameters are obtained, the index parameter (Q) should be calculated for determining the dipping angle.

For any $x \neq 0$, it is possible to solve Q by using $B(x)/A(x)$ as below:

$$Q = \tan^{-1} \left\{ \frac{B(x) \left[\tan^{-1} \left(\frac{x+b}{h} \right) - \tan^{-1} \left(\frac{x-b}{h} \right) \right]}{A(x) \left[\frac{1}{2} \ln \left[\frac{(x+b)^2 + h^2}{(x-b)^2 + h^2} \right] \right]} \right\} \tag{9}$$

Since Q is obtained between 0 and 90 (degrees) from Eq. 9, the following criterions should be applied to calculate Q correctly;

- Major positive anomaly towards the positive x -axis: $Q = Q_n$;
- Major positive anomaly towards the negative x -axis: $Q = -Q_n$;
- Major negative anomaly towards the positive x -axis: $Q = Q_n - 180^\circ$;
- Major negative anomaly towards the negative x -axis: $Q = -(Q_n + 180^\circ)$,

where Q_n is the value of Q obtained using Eq. 9 (Atchuta Rao and Ram Babu 1981).

Since all of the parameters except amplitude coefficient (C) are derived, the value of C can be calculated using Eq. 4.

Theoretical example

The parameters of the theoretical model are as follows: $C = 500$ nT, $h = 8$ m, $b = 6$ m, and $Q = -60^\circ$ (degrees). The calculated anomaly and the even and odd components for the described model are given in Fig. 4. The curves given in Fig. 5a are obtained by implementing the proposed method for $x = 3, 6, 9,$ and 12 . According to the method, the intersection point of these curves should correspond to the depth (h) and half-width (b) of the dike. In the figure, it is easy to determine that the intersection point corresponds to $h = 8$ m, and $b = 6$ m. The above-mentioned parameters obtained are the same as those of the theoretical model, showing the validity of the proposed method.

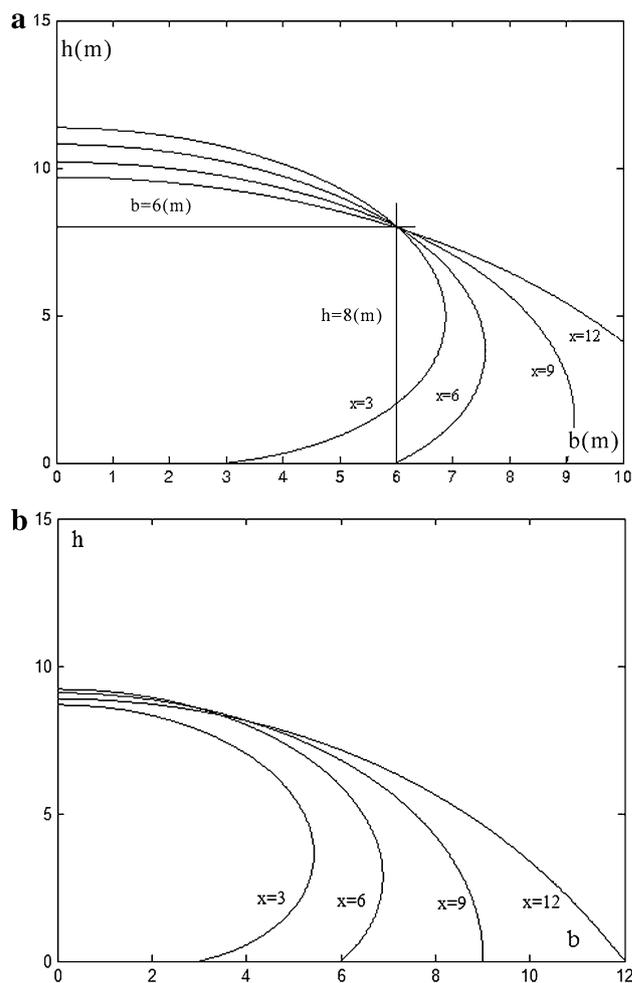


Fig. 5 **a** The curves obtained by implementing the proposed method to the anomaly calculated for the theoretical model (Fig. 4); **b** the same curves in the presence of an erroneously assessed origin. In this case, the curves are not intersecting at a common point as they do in Fig. 5a, and the point of intersection is not leading to the actual values of h and b

Before the implementation of the method, regional effects, and density irregularities must be eliminated. Besides, the zero line and particularly the origin must be determined correctly; otherwise, the calculated curves may not intersect at a single point. To show this, the origin shown in Fig. 4 is shifted by 1 unit the right and the curves are re-calculated, as shown in Fig. 5b. One should note that there is no common intersection point for all curves. Besides, for the point the curves for $x = 6, 9,$ and 12 intersect, the values of h and b are different from their actual values.

Field example

For the field example, the vertical field magnetic anomaly data over Marcona district in Peru (Fig. 11 of Parker Gay 1963) is sampled with 50 m intervals and is shown in

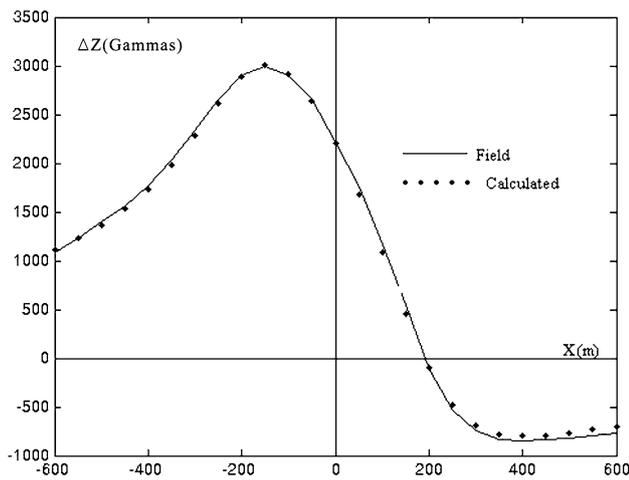


Fig. 6 The vertical field magnetic anomaly over Marcona district, Peru (Parker Gay 1963) and the calculated anomaly for the model parameters obtained using the proposed method

Fig. 6. The even and odd components calculated for this anomaly are shown in Fig. 7.

The proposed method is applied on the even component for $x = 0, 100, 200, 300, 400$ m, and $b = 188$ m and $h = 154$ m are obtained (Fig. 8). After using Eq. 9, $Q = -50.3^\circ$ is obtained and since all model parameters except the amplitude coefficient (C) are derived, the value of C is determined by using Eq. 4. The anomaly calculated for these model parameters is shown in Fig. 6. It is easy to note the similarity between the observed and the calculated anomalies. Previous interpretations of the same anomaly by

Fig. 7 The even and odd components of the field data given in Fig. 6

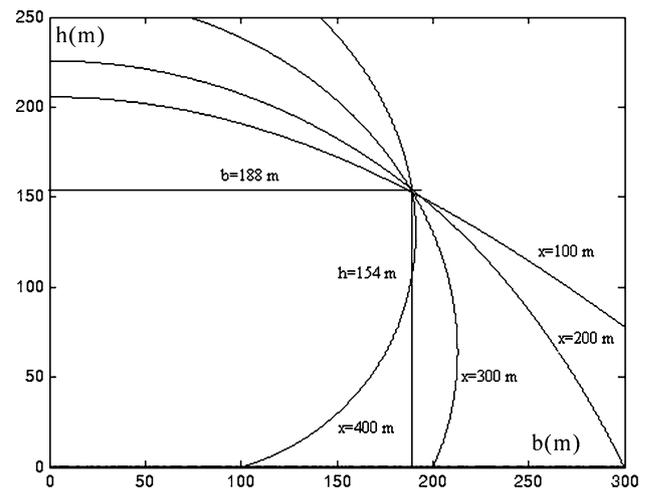
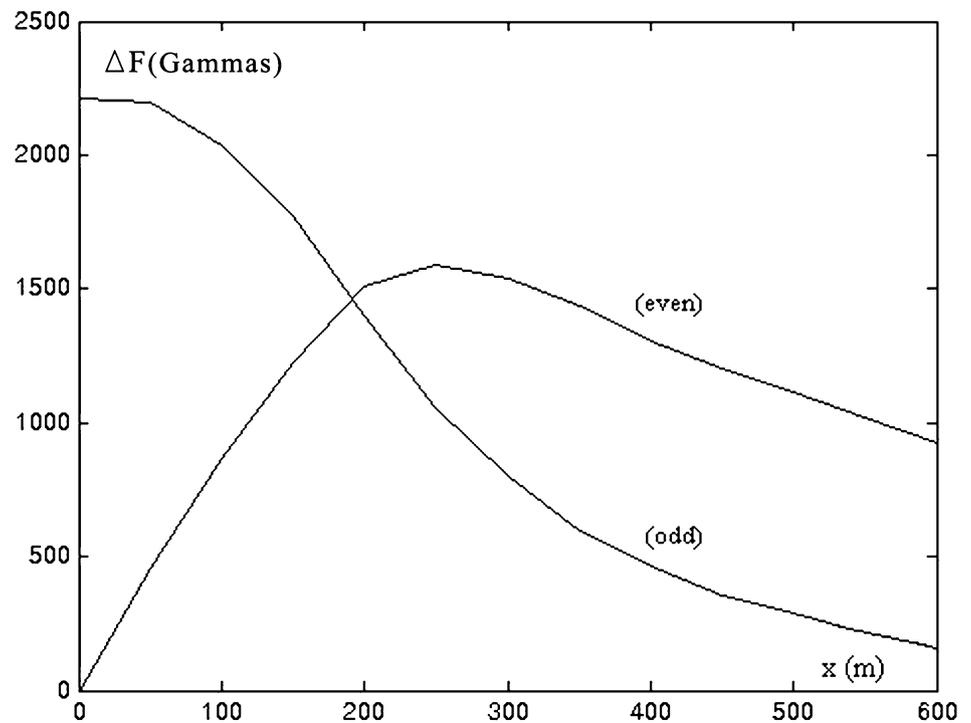


Fig. 8 The curves obtained by the implementation of the proposed method to the field data

several authors are compared to the results of this study in Table 2.

Conclusions

In this study, a method incorporating the even component of magnetic anomalies is presented to obtain depth and half-width of dikes. In the method, a curve is obtained by dividing the maximum of the even component to its value at any distance x . After several curves for different x values

Table 2 Analysis results of the vertical magnetic anomaly shown in Fig. 6

| Parameter | Parker Gay (1963) | Koulomzine et al. (1970) | Atchuta Rao and Ram Babu (1981) | Kara et al. (1996) | Presented method |
|-------------------------|----------------------|-----------------------------|------------------------------------|-----------------------|---------------------|
| Depth (h) in meters | 124 | 126.7 135.5 | 140.5 140.3 | 134 | 154 |
| Width (2b) in meters | 372 | 411.9 405.5 | 421.5 420.9 | 424 | 376 |
| Q in degrees | -50 | -50.13 -50.53 | -49.5 | -53.6 | -50.3 |

are plotted on a graph, the top-depth and half-width of the dike are delineated from the intersection point of these curves. Using this method, the anomaly may be interpreted using the maximum of the even component and its value at two different distances. The previously proposed methods were not applicable for $Q = 0^\circ$ or 90° , whereas the method proposed in this study allows interpretation for $Q = 0^\circ$. Even though the method has such advantages, the curves would not lead to the actual values in the presence of noise, erroneous assessment of the zero line or the origin.

References

- Abdelrahman EM, Essa KS (2005) Magnetic interpretation using a least-squares, depth-shape curves method. *Geophysics* 70(3):L23–L30. doi:[10.1190/1.1926575](https://doi.org/10.1190/1.1926575)
- Abdelrahman EM, Essa KS (2015) A new method for depth and shape determinations from magnetic data. *Pure appl Geophys* 172(2):439–460. doi:[10.1007/s00024-014-0885-9](https://doi.org/10.1007/s00024-014-0885-9)
- Abdelrahman EM, Abo-Ezz ER, Soliman KS, El-Araby TM, Essa KS (2007) A least-squares window curves method for interpretation of magnetic anomalies caused by dipping dikes. *Pure appl Geophys* 164(5):1027–1044. doi:[10.1007/s00024-007-0205-8](https://doi.org/10.1007/s00024-007-0205-8)
- Abo-Ezz ER, Essa KS (2016) A least-squares minimization approach for model parameters estimate by using a new magnetic anomaly formula. *Pure appl Geophys* 173(4):1265–1278. doi:[10.1007/s00024-015-1168-9](https://doi.org/10.1007/s00024-015-1168-9)
- Atchuta Rao D, Ram Babu HV (1981) Nomograms for rapid evaluation of magnetic anomalies over long tabular bodies. *Pure appl Geophys* 119(5):1037–1050. doi:[10.1007/BF00878968](https://doi.org/10.1007/BF00878968)
- Atchuta Rao D, Ram Babu HV, Sanker Narayan PV (1981) Interpretation of magnetic anomalies due to dikes: the complex gradient method. *Geophysics* 46(11):1572–1578. doi:[10.1190/1.1441164](https://doi.org/10.1190/1.1441164)
- Bhimasankaram VLS, Mohan NL, Seshagiri Rao SV (1978) Interpretation of magnetic anomalies of dikes using Fourier transforms. *Geoexploration* 16(4):259–266. doi:[10.1016/0016-7142\(78\)90015-7](https://doi.org/10.1016/0016-7142(78)90015-7)
- Essa KS (2007) A simple formula for shape and depth determination from residual gravity anomalies. *Acta Geophys* 55(2):182–190. doi:[10.2478/s11600-007-0003-9](https://doi.org/10.2478/s11600-007-0003-9)
- Essa KS, Elhoussein M (2017) A new approach for the interpretation of magnetic data by a 2-D dipping dike. *J Appl Geophys* 136:431–433. doi:[10.1016/j.jappgeo.2016.11.022](https://doi.org/10.1016/j.jappgeo.2016.11.022)
- Hutchison RD (1958) Magnetic analysis by logarithmic curves. *Geophysics* 23(4):749–769. doi:[10.1190/1.1438525](https://doi.org/10.1190/1.1438525)
- Kara İ (1997) Magnetic interpretation of two-dimensional dikes using integration- nomograms. *J Appl Geophys* 36(4):175–180. doi:[10.1016/S0926-9851\(96\)00054-7](https://doi.org/10.1016/S0926-9851(96)00054-7)
- Kara İ, Özdemir M, Yüksel FA (1996) Interpretation of magnetic anomalies of dikes using correlation factors. *Pure appl Geophys* 147(4):777–788. doi:[10.1007/BF01089702](https://doi.org/10.1007/BF01089702)
- Koulomzine T, Lamontagne Y, Nadeau A (1970) New methods for the direct interpretation of magnetic anomalies caused by inclined dikes of infinite length. *Geophysics* 35(5):812–830. doi:[10.1190/1.1440131](https://doi.org/10.1190/1.1440131)
- Mohan NL, Sundararajan N, Seshagiri Rao SV (1982) Interpretation of some two-dimensional magnetic bodies using Hilbert transforms. *Geophysics* 47(3):376–387. doi:[10.1190/1.1441342](https://doi.org/10.1190/1.1441342)
- Murthy IVR (1985) The midpoint method: magnetic interpretation of dikes and faults. *Geophysics* 50(5):834–839. doi:[10.1190/1.1441958](https://doi.org/10.1190/1.1441958)
- Parker Gay S (1963) Standard curves for interpretation of magnetic anomalies over long tabular bodies. *Geophysics* 28(2):161–200. doi:[10.1190/1.1439164](https://doi.org/10.1190/1.1439164)
- Powell DW (1967) Fitting observed profiles to a magnetized dike or fault-step model. *Geophys Prospect* 15(2):208–220. doi:[10.1111/j.1365-2478.1967.tb01784.x](https://doi.org/10.1111/j.1365-2478.1967.tb01784.x)
- Ram Babu HV, Atchuta Rao D (1991) Application of the Hilbert transform for gravity and magnetic interpretation. *Pure appl Geophys* 135(4):589–599. doi:[10.1007/BF01772408](https://doi.org/10.1007/BF01772408)
- Ram Babu HV, Subrahmanyam AS, Atchuta Rao D (1982) A comparative study of the relation figures of magnetic anomalies due to two-dimensional dike and vertical step models. *Geophysics* 47(6):926–931. doi:[10.1190/1.1441359](https://doi.org/10.1190/1.1441359)
- Rao BSR, Murthy IVR, Visweswara Rao C (1973) Two methods for computer interpretation of magnetic anomalies of dikes. *Geophysics* 38(4):710–718. doi:[10.1190/1.1440370](https://doi.org/10.1190/1.1440370)

A physics-based earthquake simulator and its application to seismic hazard assessment in Calabria (Southern Italy) region

Rodolfo Console^{1,2} · Anna Nardi² · Roberto Carluccio² · Maura Murru² · Giuseppe Falcone² · Tom Parsons³

Received: 10 February 2017 / Accepted: 16 February 2017 / Published online: 28 February 2017
© Institute of Geophysics, Polish Academy of Sciences & Polish Academy of Sciences 2017

Abstract The use of a newly developed earthquake simulator has allowed the production of catalogs lasting 100 kyr and containing more than 100,000 events of magnitudes ≥ 4.5 . The model of the fault system upon which we applied the simulator code was obtained from the DISS 3.2.0 database, selecting all the faults that are recognized on the Calabria region, for a total of 22 fault segments. The application of our simulation algorithm provides typical features in time, space and magnitude behavior of the seismicity, which can be compared with those of the real observations. The results of the physics-based simulator algorithm were compared with those obtained by an alternative method using a slip-rate balanced technique. Finally, as an example of a possible use of synthetic catalogs, an attenuation law has been applied to all the events reported in the synthetic catalog for the production of maps showing the exceedance probability of given values of PGA on the territory under investigation.

Keywords Earthquake simulator · Seismic catalogs · Fault systems · PGA · Calabria area

Introduction

The characteristic earthquake hypothesis is not strongly supported by observational data because of the relatively short duration of historical and even paleoseismological records. For instance, for the Calabria (southern Italy) region, historical information on strong earthquakes exists for at least two thousand years, but they can be considered complete for $M > 6.0$ only for the latest few centuries. As a consequence, characteristic earthquakes are seldom reported for individual fault segments, and hazard assessment is not reliably estimated by means of only minor seismicity reported in the historical and instrumental catalogs.

The algorithm applied in this study was developed upon the concepts introduced for earthquakes simulators in California (Tullis 2012) such as the constraint for the long-term slip rate on fault segments, and the adherence to a physically based model of rupture growth, without making use of time-dependent rheological parameters on the fault. Because of its limited sophistication, our algorithm is suitable for the production of synthetic catalogs resembling the long-term seismic activity of relatively simple fault systems, including hundreds of thousands earthquakes of small and moderate magnitude, even using quite limited computing resources.

Even if they cannot substitute the information contained in a good historical catalog, physics-based earthquake simulators have become popular in the recent literature, and their application has been justified by a number of reasons. In particular, earthquake simulators can provide interesting information on which renewal models can better describe the recurrence statistics, and how this is affected by features such as local fault geometry and kinematics.

✉ Rodolfo Console
rodolfo.console@ingv.it

¹ Center of Integrated Geomorphology for the Mediterranean Area (CGIAM), Potenza, Italy

² Istituto Nazionale di Geofisica e Vulcanologia (INGV), Rome, Italy

³ United States Geological Survey, USGS, Menlo Park, CA, USA

Algorithm of the simulator code

The algorithm of our earthquake simulator is based on the principle that the accumulated tectonic loading is released by co-seismic slip on fault patches that can rupture together during an earthquake. No physical constitutive laws (such as state and rate friction, pore fluid migration or viscosity) are adopted, except for the theory of elasticity and very few empirical rules establishing the value of the tectonic loading necessary for initiating and expanding a rupture.

A detailed outline of a previous version of the computer code is provided by Console et al. (2015). In the present version of the code, described by the flow chart of Fig. 1 and graphically shown in Fig. 2, the seismogenic system is modeled by rectangular fault segments, each of which is suitably located and oriented in a 3D space, and is composed by many rectangular cells with assigned dimensions in along-strike and down-dip directions. Each cell is

randomly assigned an initial slip budget (i.e., slip deficit) within a given interval around an arbitrary average value. This is to accommodate our lack of knowledge about the initial status of stress on each point of the fault. It was noted that the size and the distribution of the initial slip budget does not affect the self-organized fractal space–time–magnitude features of the system after a suitably long running of the simulation.

Each cell is loaded by a tectonic stress expressed in terms of backslip according to the slip rate available from the literature concerning each individual fault segment. A cell nucleates a rupture if the accumulated slip budget reaches a value exceeding a slip threshold that is initially fixed, but decreases after the rupture is initiated. After its rupture, the slip budget in the cell is decreased proportionally to the square root of the area of the expanding rupture. A rupture of a single cell causes a redistribution of stress on all the other cells determined through the Coulomb Failure Function and expressed in terms of additional (positive or negative) slip budget.

After nucleation, all the cells surrounding the ruptured area are searched for a possible propagation of the rupture. In this process, as a proxy for dynamic weakening, the effective slip threshold on the cells in the surrounding search area is decreased by a value proportional to the square root of the ruptured area, through the “strength reduction” (S-R) coefficient, encouraging the expansion of the rupture. The dynamic weakening is limited if the length of the search area exceeds a given number of times the width of the rupturing segment, through the “aspect ratio” (A-R) coefficient, so discouraging a further expansion of the rupture. A rupture stops if there are no cells in the search area that have a slip budget that exceeds the effective slip threshold (Fig. 2). A cell can rupture more than once in the same event.

A rupture can propagate through the borders between two segments, if their distance is shorter than a given limit length.

After the rupture is stopped, its area and its average slip are stored in the synthetic catalog. The seismic moment is computed by its definition:

$$M_0 = \mu \bar{S} A \tag{1}$$

where μ is the rigidity of the medium ($=30\text{GPa}$), \bar{S} is the average slip and A is the area of the rupture.

The magnitude is estimated from the seismic moment (Hanks and Kanamori 1979):

$$M = (\log M_0 - 9.1)/1.5 \tag{2}$$

The simulation algorithm, in its simplicity, provides preference for new ruptures to nucleate at the points of the fault where the stress budget is higher, i.e., where the time elapsed since the last event is longer. Once it is nucleated,

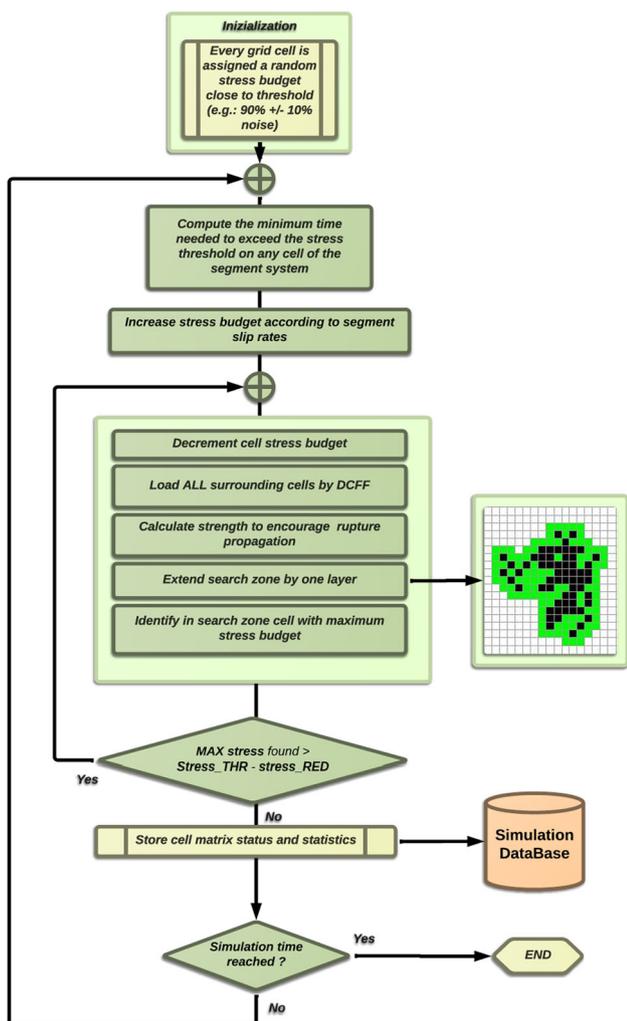


Fig. 1 Simplified flow chart of the physics-based algorithm for space–time simulation of seismic activity

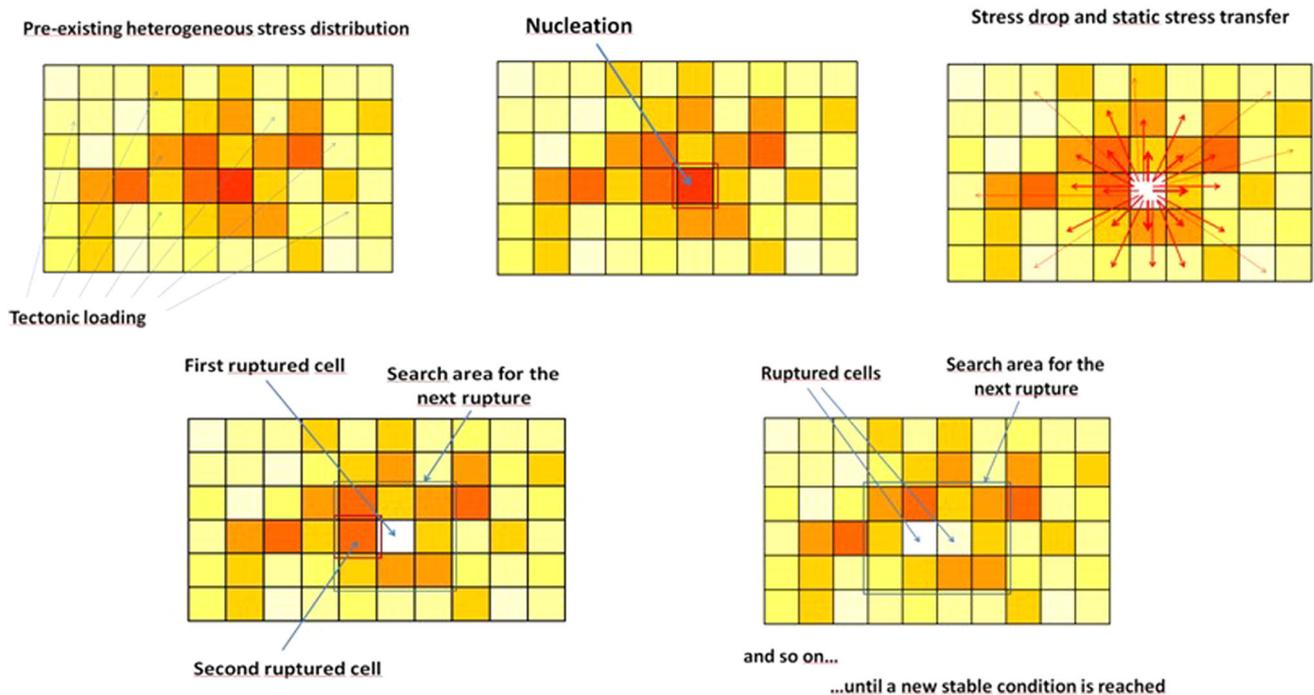


Fig. 2 Overview of the model adopted in our earthquake simulator. The seismogenic system is modeled by rectangular fault segments, each of which is composed by many rectangular cells of the same size

the rupture expands in the directions where the stress budget is still higher, thus simulating a preference for filling preexisting gaps and epicenter migration.

It should be stressed that the magnitude distribution of the simulated catalog is in no way pre-defined. It comes from the behavior of the self-organized physical process of rupture nucleation and propagation, through the choice of the two above-mentioned free parameters S-R and A-R.

Seismotectonic features of the Calabrian Arc

The Calabrian Arc is the seismic region upon which we tested the above introduced simulator code. It represents a portion of the complex plate boundary between the Eurasia and Africa plates formed as a consequence of the subduction of the Ionian oceanic crust under the southern margin of the European plate. It developed between the Tyrrhenian Sea (to the W) and the Ionian Sea (to the E) and stretches for about 300 km between the Southern Apennines (to the N) and Sicily (to the SW).

Since the Middle Pleistocene, this region has been affected by normal faulting, related to uniform ESE–WNW extension that is responsible for the crustal seismicity occurring in this area. Along the different NE–SW trending normal fault segments (Serre, Cittanova and Reggio Calabria faults) extension is accommodated by dip-slip movements, whereas slightly right-lateral components of

motion characterize the N–S trending faults (Crati Valley fault system). Large left-lateral components of motion occur along the ENE–WSW trending fault segments (e.g., S. Eufemia fault) in the southern part of the arc, along which the extension jumps westwards from the Serre-Aspromonte mountain to the Strait of Messina area. Morphological features of fault escarpments suggest long-term slip rates of 0.8–1.1 mm/year for the last 700 kyr and values of 0.7–0.9 mm/year for the last 120 kyr (Tortorici et al. 1995).

Several disastrous earthquakes of magnitude $6.6 \leq M \leq 7.1$ hit this region during the last four centuries, as reported in CPTI11 (Catalogo Parametrico dei Terremoti Italiani) (Rovida et al. 2011) (Table 1). Among the strongest historical events, two earthquake sequences, characterized by multiple mainshocks in March–June 1638 and February–March 1783 caused extensive destructions in the areas of Sant’Eufemia-Lamezia Terme and Messina-Gioia Tauro, respectively. More recently, two other catastrophic earthquakes destroyed the Sant’Eufemia gulf and the Messina straits areas in 1905 and 1908, respectively. Since it is hard to find out in which period data are missing and in which period there are definitely no other earthquakes except some particular events reported for the previous centuries, this makes the statistical evaluation of their recurrence time based on historical data difficult, but still possible, resulting in conclusions with big uncertainty. However, based on paleoseismological data, Galli and Bosi

Table 1 Major earthquakes of the Calabria fault system since 1000 AD (from CPTI11)

| Year | Month | Day | Latitude | Longitude | Mw |
|------|-------|-----|----------|-----------|------|
| 1184 | 05 | 24 | 39.40 | 16.19 | 6.74 |
| 1626 | 04 | 04 | 38.85 | 16.46 | 6.03 |
| 1638 | 03 | 27 | 39.05 | 16.29 | 7.03 |
| 1638 | 06 | 08 | 39.28 | 16.81 | 6.89 |
| 1659 | 11 | 05 | 38.69 | 16.25 | 6.55 |
| 1767 | 07 | 14 | 39.38 | 16.28 | 5.98 |
| 1783 | 02 | 05 | 38.30 | 15.97 | 7.02 |
| 1783 | 02 | 07 | 38.58 | 16.20 | 6.62 |
| 1783 | 03 | 28 | 38.79 | 16.46 | 6.98 |
| 1786 | 03 | 10 | 38.10 | 15.02 | 6.15 |
| 1791 | 10 | 13 | 38.64 | 16.27 | 6.03 |
| 1832 | 03 | 08 | 39.08 | 16.92 | 6.59 |
| 1836 | 04 | 25 | 39.57 | 16.74 | 6.20 |
| 1854 | 02 | 12 | 39.26 | 16.20 | 6.21 |
| 1870 | 10 | 04 | 39.22 | 16.33 | 6.10 |
| 1894 | 11 | 16 | 38.29 | 15.87 | 6.07 |
| 1905 | 09 | 08 | 38.82 | 15.94 | 6.94 |
| 1908 | 12 | 28 | 38.15 | 15.69 | 7.06 |

(2004) observed a rough elapsed time interval of 800–1000 years between consecutive events rupturing the Cecita Lake fault, which they consider the source of the June 1638 earthquake.

Although it is generally recognized that the major Calabria earthquakes were caused by normal faults rupturing the upper crust of the southern Calabria-Peloritani area (Boccaletti et al. 1984; Tortorici et al. 1995; Mattei et al. 1997; Serpelloni et al. 2010; D’Agostino et al. 2011; Presti et al. 2013), no consensus has been reached yet on seismogenic source location and orientation of these faults. For instance, the source of the Messina 1908 earthquake has been predominantly interpreted as a blind low-angle normal fault dipping toward the SE and located below the Messina Straits (Capuano et al. 1988; Boschi and Pantosti 1989; De Natale and Pingue 1991; Valensise and Pantosti 1992; Amoruso et al. 2002; Valensise et al. 2008; DISS Working Group 2015; De Natale and Pino 2014). Conversely, another set of studies support a NW dipping, high-angle normal faults on mainland Calabria as causative sources of this earthquake (Schick 1977; Mulargia and Boschi 1983; Ghisetti 1984, 1992; Bottari et al. 1986; Westaway 1992; Tortorici et al. 1995; Bottari 2008; Aloisi et al. 2013, 2014; Minelli et al. 2015).

The source model adopted in this study upon which we applied the simulator code, was obtained from the DISS 3.2.0 database (DISS Working Group 2015), maintained by the Istituto di Geofisica e Vulcanologia (INGV), Rome,

Italy, which is the only complete and homogeneous dataset available for the whole Italian territory. Table 2 reports the list and the parameters of the 13 Composite sources (ITCS) recognized in the Calabrian Arc, besides 2 Debated sources (ITDS), the Catena Costiera and Cecita Lakes faults (Fig. 3). The Composite Seismogenic Source is a three-dimensional representation of a crustal fault which is based on geological and geophysical data and characterized by geometric (strike, dip, width, depth) and kinematic (rake) parameters. The length of the “characteristic” rupture, however, is poorly defined or unknown, thus the source spans an unspecified number of Individual Sources (Basili et al. 2008). The Debated Seismogenic Source is an active fault that has been proposed in the literature as potential seismogenic source but was not considered reliable enough or its parameters could not be constrained sufficiently in order for it to be included in any of the other categories. We parameterize the two Debated Seismogenic Sources to increase information in the central Calabrian region where seismically active structures are located.

The faults that have a curved shape (e.g., ITCS019, ITCS104, and ITDS053) were approximated by two or three adjacent rectangular segments with appropriate strikes, giving a total of 22 segments.

For some sources, based on available information obtained from GPS measurements (Serpelloni et al. 2010; D’Agostino et al. 2011), we refined the slip rates reported by DISS 3.2.0 substituting them with values characterized by smaller uncertainties.

Application to the seismicity of Calabria region

In the application of the simulator, the rectangular fault segments representing the Calabria region fault system (CRFS) reported in Table 2 were discretized in two ways: (a) using 10,968 cells of 1.0 km × 1.0 km and (b) using a larger number of 175,488 cells of 0.25 km × 0.25 km. We chose for the synthetic catalogs in the first case a minimum magnitude of 4.5, which is produced approximately by the rupture of a patch of 4 km², and for the second case a minimum magnitude of 3.0, which is produced by the rupture of a patch of 0.12 km². The duration of all the synthetic catalogs was 100 kyr for the first case and 10 kyr for the second.

Main results

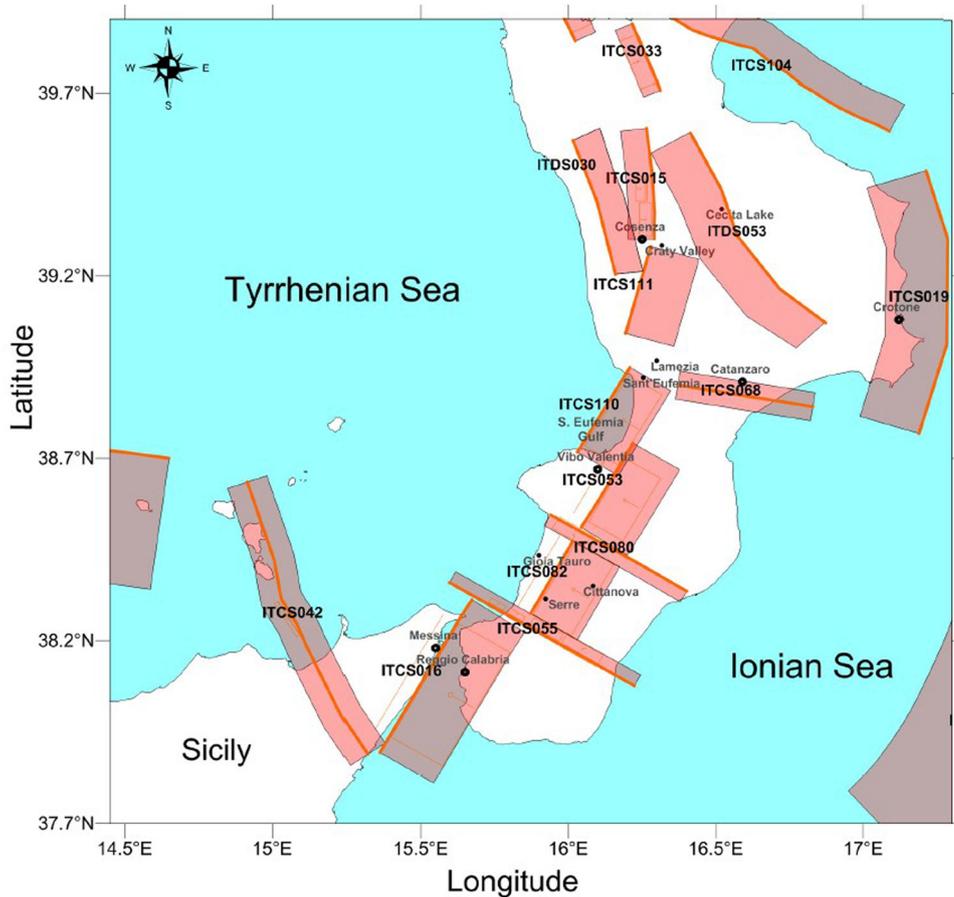
As an example of the output catalogs, we show in Fig. 4 the epicentral map of a synthetic catalog lasting 1000 years, obtained by the 1.0 km × 1.0 km discretization and choosing the values 0.7 and 2 for the S-R and A-R coefficients, respectively. The main features of a 100 kyr

Table 2 Parameters of the Calabria region fault system

| DISS code | Name | Latitude (°N) | Longitude (°E) | Depth (km) | Strike (°) | Dip (°) | Length (km) | Width (km) | Slip rate (mm/year) |
|-----------|-------------------------|---------------|----------------|------------|------------|---------|-------------|------------|---------------------|
| ITCS015 | Crati Valley | 39.60 | 16.27 | 1.0 | 155 | 50 | 22 | 18 | 1.20 |
| ITCS016 | Aspromonte-Peloritani | 37.88 | 15.36 | 1.0 | 20 | 32.5 | 48 | 20 | 1.45 |
| ITCS019A | Crotone-Rossano | 39.49 | 17.21 | 3.0 | 160 | 30 | 21 | 18 | 0.30 |
| ITCS019B | Crotone-Rossano | 39.30 | 17.29 | 3.0 | 175 | 30 | 32 | 18 | 0.30 |
| ITCS019C | Crotone-Rossano | 39.49 | 17.21 | 3.0 | 190 | 30 | 28 | 18 | 0.30 |
| ITDS030 | Catena Costiera | 39.46 | 16.12 | 1.5 | 345 | 50 | 28 | 10 | 0.22 |
| ITCS033 | Pollino | 39.992 | 16.23 | 1.0 | 160 | 60 | 22 | 10 | 1.25 |
| ITCS042A | Patti-Eolie | 38.64 | 14.92 | 1.0 | 155 | 80 | 37 | 24 | 1.0 |
| ITCS042B | Patti-Eolie | 38.31 | 15.03 | 1.0 | 155 | 80 | 53 | 24 | 1.0 |
| ITCS053 | Serre | 38.55 | 16.03 | 3.0 | 30 | 30 | 27 | 16 | 0.55 |
| ITDS053A | Cecita Lakes | 39.58 | 16.19 | 1.0 | 155 | 50 | 22 | 18 | 0.22 |
| ITDS053B | Cecita Lakes | 39.40 | 16.55 | 1.0 | 155 | 50 | 12 | 18 | 0.22 |
| ITDS053C | Cecita Lakes | 39.29 | 16.58 | 1.0 | 155 | 50 | 35 | 18 | 0.22 |
| ITCS055 | Bagnara-Bovalino | 38.08 | 16.22 | 3.0 | 300 | 70 | 63 | 10 | 0.30 |
| ITCS068 | Caraffa-Squillace | 38.91 | 16.37 | 3.0 | 100 | 80 | 40 | 12 | 0.30 |
| ITCS080 | Nicotera-Rocella Ionica | 38.55 | 15.94 | 3.0 | 120 | 70 | 46 | 10 | 0.30 |
| ITCS082 | Gioia Tauro Plain | 38.29 | 15.84 | 3.0 | 30 | 30 | 25 | 15 | 0.55 |
| ITCS104A | Amendolara | 39.87 | 16.43 | 1.0 | 290 | 45 | 11 | 13 | 0.50 |
| ITCS104B | Amendolara | 39.82 | 16.62 | 1.0 | 290 | 45 | 17 | 13 | 0.50 |
| ITCS104C | Amendolara | 39.60 | 17.08 | 1.0 | 290 | 45 | 46 | 13 | 0.50 |
| ITCS110 | Sant'Eufemia | 38.72 | 16.03 | 0.50 | 30 | 40 | 30 | 14 | 0.55 |
| ITCS111 | Savuto Valley | 39.04 | 16.19 | 1.0 | 20 | 40 | 28 | 19 | 1.25 |

| DISS code | Name | Latitude (°N) | Longitude (°E) | Depth (km) | Strike (°) | Dip (°) | Rake (°) | Length (km) | Width (km) | Slip rate (mm/yr) |
|-----------|-------------------------|---------------|----------------|------------|------------|---------|----------|-------------|------------|-------------------|
| ITCS015 | Crati Valley | 39.60 | 16.26 | 1.0 | 175 | 60 | 270 | 33.5 | 10 | 1.20 |
| ITCS016 | Aspromonte-Peloritani | 37.88 | 15.36 | 1.0 | 30 | 32.5 | 265 | 50 | 20 | 1.45 |
| ITCS019A | Crotone-Rossano | 39.49 | 17.21 | 3.0 | 160 | 30 | 90 | 21 | 18 | 0.30 |
| ITCS019B | Crotone-Rossano | 39.30 | 17.29 | 3.0 | 175 | 30 | 90 | 32 | 18 | 0.30 |
| ITCS019C | Crotone-Rossano | 39.01 | 17.28 | 3.0 | 190 | 30 | 90 | 28 | 18 | 0.30 |
| ITDS030 | Catena Costiera | 39.46 | 16.12 | 1.5 | 345 | 50 | 270 | 28 | 10 | 0.22 |
| ITCS033 | Pollino | 39.89 | 16.22 | 1.0 | 160 | 60 | 270 | 22 | 10 | 1.25 |
| ITCS042A | Patti-Eolie | 38.64 | 14.92 | 1.0 | 165 | 80 | 200 | 37 | 24 | 1.0 |
| ITCS042B | Patti-Eolie | 38.31 | 15.03 | 1.0 | 155 | 80 | 200 | 53 | 24 | 1.0 |
| ITCS053 | Serre | 38.51 | 16.05 | 3.0 | 30 | 30 | 270 | 27 | 16 | 0.55 |
| ITDS053A | Cecita Lakes | 39.64 | 16.41 | 1.0 | 155 | 50 | 250 | 22 | 18 | 0.22 |
| ITDS053B | Cecita Lakes | 39.47 | 16.52 | 1.0 | 170 | 50 | 250 | 12 | 18 | 0.22 |
| ITDS053C | Cecita Lakes | 39.35 | 16.58 | 1.0 | 140 | 50 | 250 | 35 | 18 | 0.22 |
| ITCS055 | Bagnara-Bovalino | 38.08 | 16.22 | 3.0 | 300 | 70 | 225 | 63 | 10 | 0.30 |
| ITCS068 | Caraffa-Squillace | 38.90 | 16.37 | 3.0 | 100 | 80 | 350 | 40 | 12 | 0.30 |
| ITCS080 | Nicotera-Rocella Ionica | 38.55 | 15.94 | 3.0 | 120 | 70 | 315 | 46 | 10 | 0.30 |
| ITCS082 | Gioia Tauro Plain | 38.27 | 15.88 | 3.0 | 30 | 30 | 270 | 25 | 15 | 0.55 |
| ITCS104A | Amendolara | 39.87 | 16.43 | 1.0 | 300 | 45 | 77.5 | 11 | 13 | 0.50 |
| ITCS104B | Amendolara | 39.82 | 16.62 | 1.0 | 285 | 45 | 77.5 | 17 | 13 | 0.50 |
| ITCS104C | Amendolara | 39.60 | 17.08 | 1.0 | 300 | 45 | 77.5 | 46 | 13 | 0.50 |
| ITCS110 | Sant'Eufemia | 38.72 | 16.03 | 0.50 | 30 | 40 | 270 | 30 | 14 | 0.55 |
| ITCS111 | Savuto Valley | 39.04 | 16.19 | 1.0 | 20 | 40 | 270 | 28 | 19 | 1.25 |

Fig. 3 Seismotectonic setting of the Calabria region. The colored boxes show the composite seismogenic sources adopted in this study from DISS 3.2.0 (published in 2015). The light brown rectangles show the ruptures of the latest historical earthquakes associated to the seismogenic sources



simulation using the same input parameters are listed in Table 3.

An event of given magnitude produced by the simulation algorithm could have ruptured only part of a single segment for moderate magnitudes, or have propagated to more than one segment for the larger magnitudes. It should be noted that the number of 1.0 km^2 cells of the smallest segment in the fault system (ITCS104A, Amendolara, northern segment) is 143 (given by a width of 13 km and length of 11 km) (equivalent to a magnitude 6.0 earthquake), and that of the largest one (ITCS042B, Patti-Eolie, southern segment) is 1272 (given by a width of 24 km and length of 53 km) (equivalent to a magnitude 6.9 earthquake).

To carry out a statistical analysis on the number of ruptured segments contributing to a single earthquake, we based our analysis upon some arbitrary criteria to assign a specific earthquake to one or more segments. The rules adopted for this purpose are:

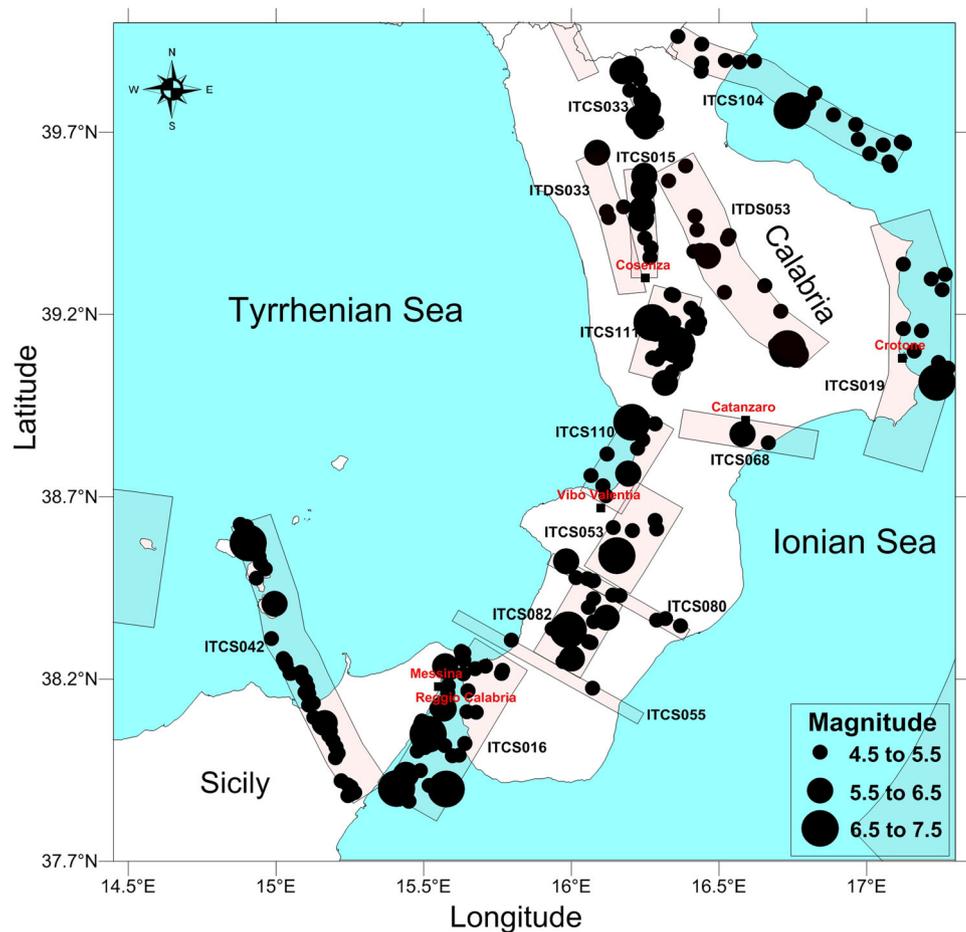
1. The earthquake must have a minimum magnitude of 6.0;
2. The earthquake is initially assigned to the segment that contains the largest number of fractured cells;

3. The same earthquake can be also assigned to other segments if the number of 1.0 km^2 cells ruptured by the event exceeds the number of 150 or at least 60% of the total number of cells for the specific segment.

In this way, we counted through the whole 100 kyr synthetic catalog the number of times that a given fault segment was present in any $M \geq 6.0$ earthquake alone or jointly with other segments. The results are shown in Fig. 5, where the light brown boxes contain the numbers of single-segment versus multi-segment ruptures. It is possible to see that some segments, in particular the large segments that are not very close to other segments, exhibit a preference to single segment rupturing, while the opposite preference is shown by smaller segments located very close each other.

To explore the effect of the arbitrary choice of the free parameters used in the algorithm we carried out a series of simulations, each using different combinations of these parameters. The effect of this arbitrary choice on the magnitude-frequency distribution, for the discretization of $1 \text{ km} \times 1 \text{ km}$ is shown in Fig. 6a, b. Note that the total seismic moment released by all the earthquakes in the catalogs is independent of the choice of the algorithm free

Fig. 4 Epicentral map ($M \geq 4.5$) of the first 1000 years of a synthetic catalog obtained by the $1.0 \text{ km} \times 1.0 \text{ km}$ discretization of the main seismogenic structures recognized by DISS 3.2.0 in Calabria region. The size of the black dots shows the magnitude class



parameters, but depends only on the constraints given by the slip rate on each fault segment.

Figure 6a, b show a clear deviation of the magnitude distribution obtained by the simulation algorithm from a plain linear Gutenberg–Richter distribution. The magnitude distribution exhibits a sharp “bump” around a value of magnitude 6.5 that could be defined as “characteristic”. This behavior can be explained by the circumstance that the fault system adopted in the simulation (Table 2) contains many isolated fault segments of similar length and width, whose maximum magnitude is typically around 6.5, and cannot link with others to generate larger magnitude events. We could recall also that the model of fault segments adopted in this study does not contain any smaller size isolated segments that could generate a magnitude-frequency distribution more similar to that of the Gutenberg–Richter (G–R) law.

Considering the case of the discretization of the fault segments in cells of smaller size, after having carried out a series of trials with different choices of the free parameters as in the previous case, we chose the combination of $A-R = 5$ and $S-R = 0.4$ to obtain a good match between the synthetic and the real catalogs in the intermediate

magnitude range. The results are shown in the cumulative magnitude-frequency plot of Fig. 7 and compared with the plots of the historical CPTI11 catalog (Rovida et al. 2011) and the instrumental ISIDE (Italian Seismological Instrumental and Parametric Data-Base) catalog (Mele et al. 2010).

The comparison reported in Fig. 7 shows a substantial similarity between the synthetic and the two real catalogs in the magnitude range $4.0 \leq M \leq 6.5$. A larger number of $M < 4.0$ events and a higher b -value (of the order of 2) are notable in the simulation results, with respect to the instrumental ISIDE data, which can be considered complete for magnitudes even smaller than 3.0. Note that the b -value of the synthetic catalog in the low magnitudes range is strongly dependent on the choice of the coefficient S-R in the simulation algorithm. The b -value of both the synthetic and the CPTI11 catalogs is smaller than 1 in the magnitude range $4.0 \leq M \leq 6.5$. This low value can be explained by the lack of smaller seismogenic structures generating intermediate magnitude earthquakes in the model of 22 seismogenic segments adopted in this study from DISS 3.2.0, but we do not know if this can be applied to the real seismicity.

Table 3 Main features of a synthetic catalog obtained by the simulator for seismogenic sources of the CRFS

| DISS Code | Name | Total number of events $M \geq 4.5$ | Number of events $M \geq 6.0$ | Maximum magnitude |
|-----------|-------------------------|-------------------------------------|-------------------------------|-------------------|
| ITCS015 | Crati Valley | 1350 | 343 | 6.74 |
| ITCS016 | Aspromonte-Peloritani | 5327 | 438 | 7.17 |
| ITCS019A | Crotone-Rossano | 461 | 43 | 6.98 |
| ITCS019B | Crotone-Rossano | 442 | 52 | 6.98 |
| ITCS019C | Crotone-Rossano | 375 | 48 | 6.98 |
| ITDS030 | Catena Costiera | 1207 | 339 | 6.73 |
| ITCS033 | Pollino | 2598 | 146 | 7.17 |
| ITCS042A | Patti-Eolie | 2857 | 201 | 7.17 |
| ITCS042B | Patti-Eolie | 830 | 153 | 6.92 |
| ITCS053 | Serre | 372 | 41 | 7.03 |
| ITDS053A | Cecita Lakes | 345 | 106 | 6.73 |
| ITDS053B | Cecita Lakes | 551 | 127 | 6.73 |
| ITDS053C | Cecita Lakes | 461 | 43 | 6.98 |
| ITCS055 | Bagnara-Bovalino | 2094 | 292 | 7.03 |
| ITCS068 | Caraffa-Squillace | 1350 | 343 | 6.74 |
| ITCS080 | Nicotera-Rocella Ionica | 437 | 74 | 6.88 |
| ITCS082 | Gioia Tauro Plain | 233 | 33 | 7.03 |
| ITCS104A | Amendolara | 5327 | 438 | 7.17 |
| ITCS104B | Amendolara | 2598 | 146 | 7.17 |
| ITCS104C | Amendolara | 2857 | 201 | 7.17 |
| ITCS110 | Sant'Eufemia | 392 | 109 | 7.16 |
| ITCS111 | Savuto Valley | 997 | 140 | 7.15 |

Some discrepancy can be observed between the synthetic and the CPTI11 catalogs in the high magnitude range ($M \geq 6.2$). Indeed, even if both catalogs are characterized by a rate of one $M \geq 6.2$ earthquake per century, the CPTI11 catalog reports a questionable distribution of no events in the magnitude range $6.2 \leq M \leq 6.8$, and four earthquakes of magnitude $6.9 \leq M \leq 7.1$, while the simulation shows a smooth decrease of the rate from magnitude 6.2–7.2. The limited number of events and the large uncertainty in magnitude estimations of the historical catalog does not allow a robust comparison with the simulator's results.

In the following, we shall consider only the results of the simulations obtained with the choice of the free parameters $A-R = 2$ and $S-R = 0.7$ for the discretization of $1 \text{ km} \times 1 \text{ km}$.

To test the potential use of the simulation algorithm for drawing conclusions about the predictability of strong earthquakes, it should be useful to make a comparison between the synthetic and the real catalog. However, this is not a trivial task, first of all because the three-century period covered by the CPTI11 catalog is not long enough to encompass the recurrence time of characteristic earthquakes on single segments of the CRFS.

Moreover, the two kinds of catalogs are conceptually different. The list of 18 earthquakes of $M \geq 6.0$, reported in Table 1 for the latest ten centuries assumes that each event occurred on a single segment. Nevertheless, as remarked in “Main results” section, a single earthquake could have ruptured only a part of a single segment for moderate magnitudes, or have propagated to more than one segment for the larger magnitudes, in agreement with the assumption made in our simulation algorithm. A discussion on the issue of recurrence intervals on seismogenic sources was made by Field (2015) addressing the “recurrence of what?” question.

To assess whether the earthquake occurrence time on single segments in the synthetic catalog behaves as a Poisson process or not, we carried out a statistical analysis of the inter-event times for the entire 100 kyr simulation. In this respect, Fig. 8 shows the inter-event time distribution of the characteristic earthquakes defined by the empirical rules defined in “Main results” section for three selected fault segments. Table 4 displays the mean inter-event time T_r , the standard deviation σ and the coefficient of variation C_v for each segment of the CRFS. The relatively short average inter-event times of the simulations can be justified by the circumstance that often two or more segments rupture simultaneously in a single earthquake.

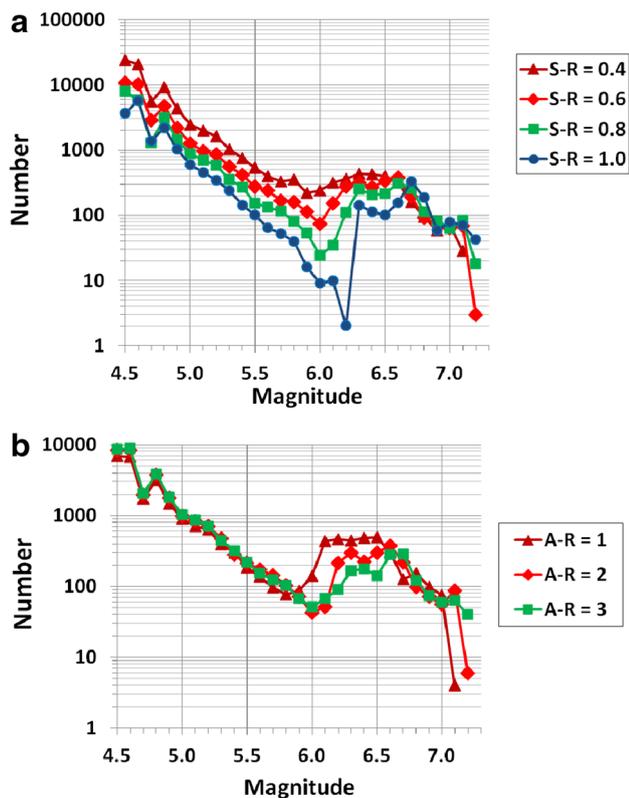


Fig. 6 **a** Magnitude-frequency distribution of the earthquakes in the synthetic catalog obtained from the simulation algorithm described in the text using a discretization of $1 \text{ km} \times 1 \text{ km}$ and an aspect ratio (A-R) coefficient equal to 2 and various values of the stress reduction (S-R) coefficient. **b** Magnitude-frequency distribution of the earthquakes in the synthetic catalog obtained from the simulation algorithm described in the text using a discretization of $1 \text{ km} \times 1 \text{ km}$ and various values of the aspect ratio (A-R) coefficient and a stress reduction (S-R) coefficient equal to 0.7

fault system, and the long-term fault slip rates. For the Calabria case we discretize the faults into 2 km by 2 km patches. Hypocenter locations are assigned to patches initially at random, and their magnitudes are also drawn at random from an input magnitude–frequency distribution. Ruptures grow patch-by-patch along the fault system, each absorbing some of the slip rate budget (annual rate multiplied by 100 kyr). We use the generalized magnitude–area relation of Wells and Coppersmith (1994) because there is a wide variety of observed fault rakes in the Calabria fault system. If there is not room within the fault system at the initial location, then hypocenters are attempted throughout the fault system until their full ruptures can be accommodated. In this way, the highest magnitude events tend to end up on the longest faults, or linked faults. If there is no place within the system and within the available slip budget for an assigned magnitude, then it is adjusted downward to fit into the largest remaining space. Thus, the output magnitude–frequency distribution will differ from the input as a result of the observed fault

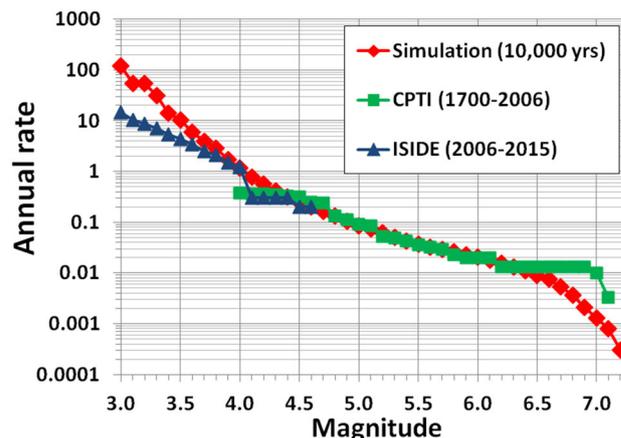


Fig. 7 Cumulative magnitude-frequency distribution of the earthquakes in Calabria region. *Red diamonds* synthetic catalog obtained from the simulation algorithm described in the text using a discretization of $0.25 \text{ km} \times 0.25 \text{ km}$, an aspect ratio (A-R) coefficient equal to 5 and a stress reduction (S-R) coefficient equal to 0.4; *green squares* the Italian historical catalog CPTI11 for the latest three centuries; *blue triangles* the Italian instrumental catalog ISIDE for the latest 20 years. All the three series are displayed in the same units (events per year)

geometry and slip rates. Each patch is given the average slip (\bar{S}) for its magnitude according to the definition of seismic moment (1) and the moment–magnitude relation (2):

$$\bar{S} = \frac{10^{(1.5M+9.05)}}{\mu A} \quad (3)$$

The primary advantage of the slip-rate method is that it relies almost entirely on well-known information in the form of fault locations, long-term slip rates, and the regional earthquake rate, which can be extrapolated from instrumental observations or pulled from historical catalogs. It therefore has very few parameters, and produces moment balanced long-term earthquake rate solutions. The primary disadvantage is that it cannot capture temporal clustering, irregular earthquake slip distributions, or stress interactions.

We conduct simulations over 100 kyr periods using two different regional magnitude–frequency distributions as inputs, a G–R distribution with b -value equal to 1.0, and a fully characteristic distribution with b -value equal to 0 (Fig. 10). In these simulations, we adopted the same source parameters as in Table 2, with the exception that for some of them the slip rates are those (uncorrected for the GPS observations) reported in Table 5. We allow an earthquake budget of $0.1 M \geq 4.5$ events per year. We find that the system wide long-term slip rates can be fit very well with either input distribution, with the Gutenberg–Richter distribution having an overall RMS misfit of 0.375 mm/year , and a mean misfit of -0.215 mm/year . The characteristic input distribution fits better with an RMS of 0.074 mm/

Fig. 8 Inter-event time distribution from a simulation of 100 kyr of seismic activity across the Calabria fault system for three selected fault segments: ITCS016, Aspromonte-Peloritani; ITCS110, Sant’Eufemia; ITCS019A, Crotono-Rossano, northern segment

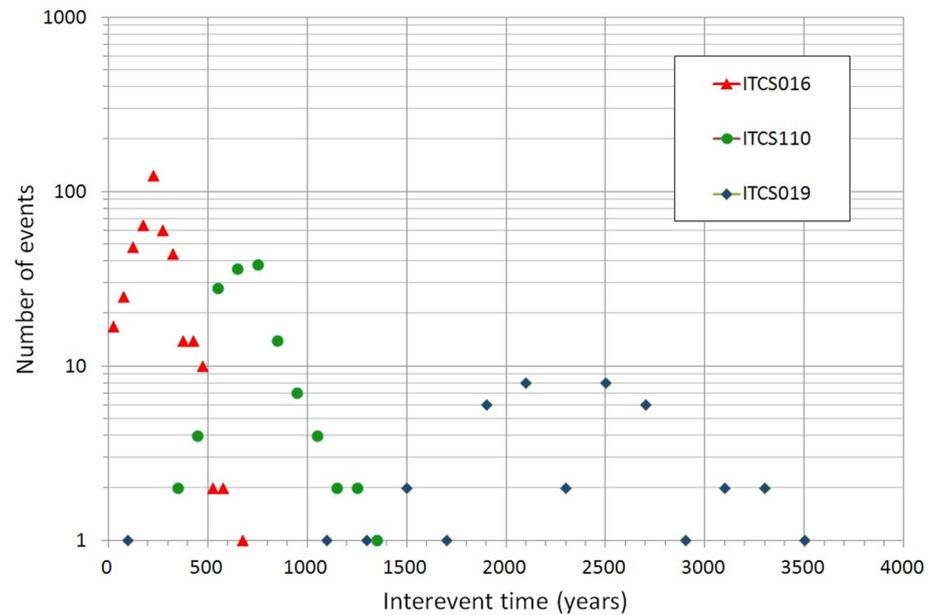


Table 4 Statistical parameters (the mean inter-event time T_r , the Poisson probability, P_{100} in %, the standard deviation σ and the coefficient of variation C_v) of the synthetic catalog for the CRFS

| DISS code | Source name | T_r ($M \geq 6.0$) (years) | P_{100} (%) | σ (years) | C_v |
|-----------|-------------------------|--------------------------------|---------------|------------------|-------|
| ITCS015 | Crati Valley | 292.9 | 28.9 | 61.5 | 0.21 |
| ITCS016 | Aspromonte-Peloritani | 230.4 | 35.2 | 102.6 | 0.45 |
| ITCS019A | Crotono-Rossano | 2312.6 | 4.2 | 616.6 | 0.27 |
| ITCS019B | Crotono-Rossano | 1951.2 | 5.0 | 610.9 | 0.31 |
| ITCS019C | Crotono-Rossano | 2117.2 | 4.6 | 568.8 | 0.27 |
| ITDS030 | Catena Costiera | 295.0 | 28.8 | 69.6 | 0.24 |
| ITCS033 | Pollino | 692.4 | 13.4 | 213.6 | 0.31 |
| ITCS042A | Patti-Eolie | 501.1 | 20.0 | 251.4 | 0.50 |
| ITCS042B | Patti-Eolie | 660.9 | 18.1 | 188.0 | 0.28 |
| ITCS053 | Serre | 720.6 | 13.0 | 165.0 | 0.23 |
| ITDS053A | Cecita Lakes | 3325.0 | 3.0 | 832.7 | 0.25 |
| ITDS053B | Cecita Lakes | 3093.9 | 3.2 | 965.2 | 0.31 |
| ITDS053C | Cecita Lakes | 2475.1 | 4.0 | 939.6 | 0.38 |
| ITCS055 | Bagnara-Bovalino | 934.5 | 10.1 | 383.2 | 0.41 |
| ITCS068 | Caraffa-Squillace | 1362.2 | 7.1 | 349.4 | 0.26 |
| ITCS080 | Nicotera-Rocella Ionica | 1071.0 | 8.9 | 371.8 | 0.35 |
| ITCS082 | Gioia Tauro Plain | 716.6 | 13.0 | 220.7 | 0.31 |
| ITCS104A | Amendolara | 1034.9 | 9.2 | 540.5 | 0.52 |
| ITCS104B | Amendolara | 946.2 | 9.0 | 433.2 | 0.46 |
| ITCS104A | Amendolara | 791.3 | 11.9 | 114.4 | 0.14 |
| ITCS110 | Sant’Eufemia | 720.6 | 13.9 | 165.0 | 0.23 |

year, and mean misfit of -0.018 mm/year. The input characteristic distribution naturally puts much more weight into higher magnitude earthquakes, which tends to match the observed slip rates better given the low annual $M \geq 4.5$ earthquake budget. The characteristic input also matches

the observed magnitude–frequency distribution of the Calabria region, which has a very low apparent b -value (Fig. 10). The physics-based simulator achieves a much closer fit to the observed magnitude–frequency (Fig. 6) distribution than the slip-balance approach. We could apply

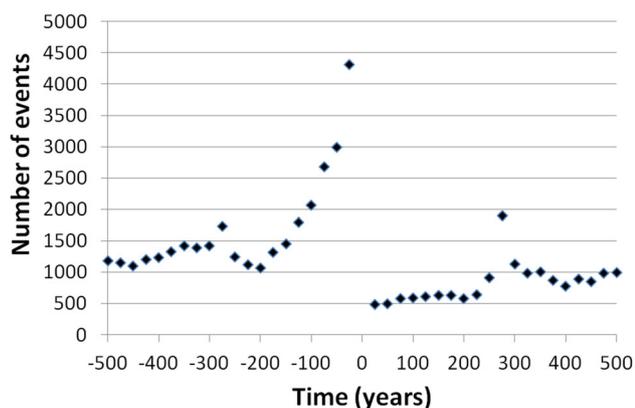
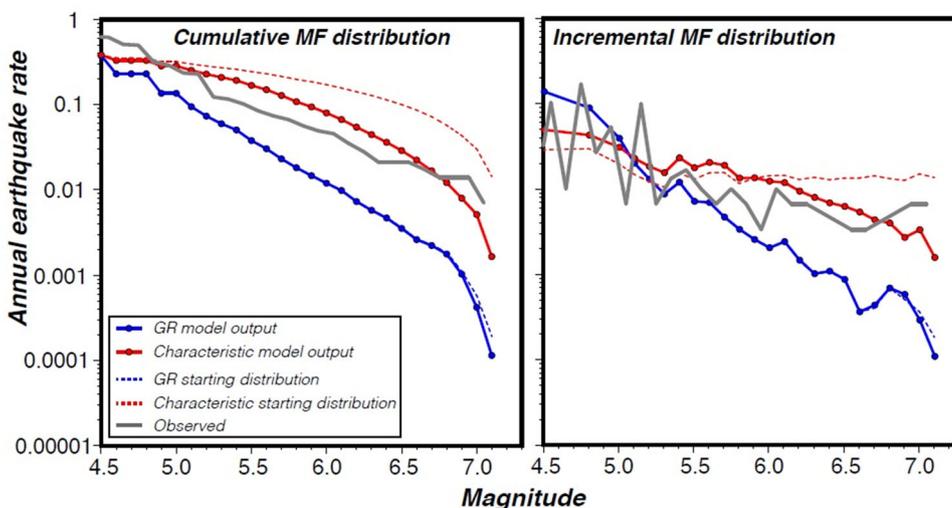


Fig. 9 Stacked number of $M3.0+$ events in the 500 years period before and after an $M6.0+$ mainshock in a 100 kyr simulated catalog of Calabria region

the observed distribution directly for a solution, but it is difficult to know if the shorter term historical and instrumental catalogs are truly representative of the long term (100 kyr) given the mostly slow ~ 1 mm/year slip rates of the Calabria region and implied long inter-event times. If the observed magnitude–frequency distribution is representative of the long-term, then the Calabria region has an unusually stronger tilt toward larger earthquakes than most regions of comparable size, which tend to have b -values closer to 1.0.

There does appear to be method dependence on the calculated rates of $M \geq 6.0$ earthquakes on individual faults (Table 5). The modeled magnitude–frequency distributions have significantly different weighting across the spectrum such that significant differences emerge at a specific threshold like $M \geq 6.0$ (compare Figs. 6 and 10). As might be expected, the slip-rate balance method results are more heavily dependent on observed slip rates than is the stress-based method.

Fig. 10 Comparison of input and model output magnitude–frequency for cumulative (left panel) and incremental (right panel) distributions. Blue curves represent input–output for a Gutenberg–Richter $b = 1$ assumption, whereas red curves represent characteristic distributions with $b = 0$. Gray lines indicate the observed earthquake distribution. The characteristic assumption is a better fit to the observed, though neither model fits as well as the physics-based results shown in Fig. 6



The simulated catalog applied to seismic hazard assessment

To explore the potential use of a synthetic catalog obtained by the simulation algorithm for seismic hazard assessment, we adopted the Cornell (1968) method applied to the $M \geq 4.5$, 100 kyr simulated catalog.

The peak ground acceleration (PGA) at a dense grid of points covering the territory of the Calabria region was estimated for each earthquake of the catalog through a typical attenuation law (Sabetta and Pugliese 1987):

$$\log_{10}(\text{PGA}) = -1.562 + 0.306M - \log_{10}(\sqrt{d^2 + 5.8^2}) + 0.169S_1 + 0.169S_2 \pm 0.173 \tag{4}$$

where M is the earthquake magnitude, d is the epicentral distance, and S_1 and S_2 are parameters taking into account the soil dynamic features at the site. A sample of the estimated PGA in the first 100 years of simulation for the site of Messina is given in Table 6.

At each node of the grid we obtained the distribution of the number of times that a given PGA was exceeded in 100 kyr, and from that the probability of exceedence of the given PGA in 50 years. The results for $\text{PGA} = 0.2$ g are shown in Fig. 11, as an example of how the method can work.

Conclusions

The application of our physics-based simulation algorithm to the Calabria fault system has allowed the compilation of synthetic seismic catalogs lasting 100 kyr for $M \geq 4.5$ and 10 kyr for $M \geq 3.0$. These catalogs contain hundreds of thousands events whose magnitude distribution and time–

Table 5 Comparison of the number of $M \geq 6$ earthquake nucleating on DISS 3.2.0 source faults over 100 kyr simulations

| DISS code | Physics-based | Slip-rate char | Slip-rate GR | Obs. slip rate (mm/year) |
|-----------|---------------|----------------|--------------|--------------------------|
| ITCS015 | 301 | 77 | 22 | 0.55 |
| ITCS016 | 330 | 585 | 53 | 1.45 |
| ITCS019 | 62 | 157 | 38 | 0.30 |
| ITCS033 | 301 | 56 | 10 | 0.40 |
| ITCS042 | 207 | 171 | 13 | 0.30 |
| ITCS053 | 45 | 126 | 25 | 0.60 |
| ITCS055 | 30 | 128 | 26 | 0.30 |
| ITCS068 | 38 | 134 | 31 | 0.30 |
| ITCS080 | 32 | 93 | 15 | 0.30 |
| ITCS082 | 95 | 14 | 4 | 0.15 |
| ITCS104 | 151 | 230 | 34 | 0.70 |
| ITCS110 | 105 | 22 | 10 | 0.20 |
| ITCS111 | 245 | 76 | 11 | 0.55 |
| ITDS053 | 45 | 199 | 15 | 0.60 |

The physics-based simulation rates are shown along with two versions of the slip-rate balance method with Gutenberg–Richter (GR) and characteristic starting magnitude–frequency distributions. The observed slip rates on each fault are also given

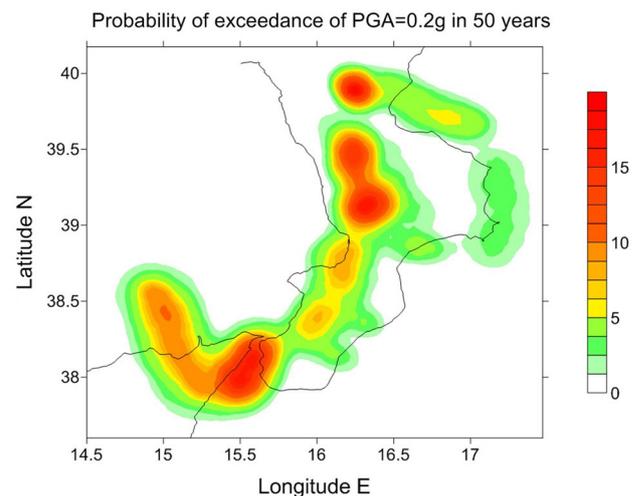
Table 6 Estimated effects on the city of Messina for the 9 earthquakes of the first 100 years of the 100 kyr synthetic catalog

| Time (years) | Magnitude | Ep. distance (km) | PGA (g) |
|--------------|-----------|-------------------|---------|
| 1.22 | 4.63 | 119.32 | 0.009 |
| 12.68 | 6.34 | 157.00 | 0.022 |
| 20.27 | 6.48 | 129.13 | 0.030 |
| 47.58 | 4.46 | 196.75 | 0.005 |
| 48.96 | 4.80 | 30.65 | 0.038 |
| 53.12 | 5.57 | 21.78 | 0.091 |
| 71.28 | 4.46 | 37.87 | 0.024 |
| 88.92 | 4.55 | 189.24 | 0.005 |
| 97.36 | 4.46 | 192.07 | 0.005 |

space features resemble those of the observed seismicity, but without the limitations that real catalogs suffer in terms of completeness and time duration.

The main conclusions that can be drawn on the results of this study can be resumed in the following.

1. The spatial distribution of the simulated seismicity is constrained by the geometrical parameters of the source model (DISS 3.2.0) adopted in the study.
2. The total seismic moment released by the earthquakes of the synthetic catalogs is constrained by the value of slip rate assumed for each fault segment, a critical aspect of the simulation process.
3. The frequency–magnitude distribution of the simulated seismicity is rather consistent with observations in the intermediate magnitudes range ($4.0 \leq M \leq 6.0$), and

**Fig. 11** Probability of exceedance of $\text{PGA} = 0.2 \text{ g}$ in 50 years in Calabria region, obtained from the method described in the text

supports a sort of characteristic earthquake hypothesis in the high magnitude range ($M > 6.0$).

4. The long period of simulations allowed us to obtain the statistical distribution of repeat time—an indispensable component for any time-dependent seismic hazard assessment which cannot be obtained from real observations, due to their short duration. The statistical distribution of earthquakes with $M \geq 6.0$ on single segments exhibits a fairly clear pseudo-periodic behavior, with a coefficient of variation C_v on the order of 0.2–0.5.
5. We have found in our synthetic catalogs a clear trend of acceleration of seismic activity in the 150 years

preceding $M \geq 6.0$ earthquakes and quiescence in a period of at least 250 years following those earthquakes.

6. The results of the physics-based simulator have been compared with those of a simpler slip-rate balanced technique, obtaining a substantial, even if not general, agreement.
7. By the application of a simple attenuation law to all the events contained in the synthetic catalog lasting 100 kyr, it has been possible to obtain a statistical hazard assessment for the Calabria region, in terms of a map reporting the probability of exceedence of a peak ground acceleration of 0.2 g in 50 years.
8. Even if this study was only aimed to explore methodological aspects and potential capabilities of new generation simulator algorithms, the initial results shown in this paper encourage the application of simulators in support to other methodologies of time-independent and time-dependent seismic hazard assessment. In every way, we would like to stress that such hazard assessments rely on physical assumptions (as the value of the slip rate on the sources and the attenuation law adopted in the PGA estimate) in the simulator to be realistic and complete: a condition that can probably never be fully met.

Acknowledgements We are very grateful to two anonymous reviewers for their comments that help to clarify some aspects of the work.

References

- Aloisi M, Bruno V, Cannavò F, Ferranti L, Mattia M, Monaco C, Palano M (2013) Are the source models of the M 7.1 1908 Messina Straits earthquake reliable? Insights from a novel inversion and a sensitivity analysis of leveling data. *Geophys J Int* 192:1025–1041. doi:10.1093/gji/ggs062
- Aloisi M, Bruno V, Cannavò F, Ferranti L, Mattia M, Monaco C (2014) Reply to ‘Comments on the paper “Are the source models of the M 7.1 1908 Messina Straits earthquake reliable? Insights from a novel inversion and sensitivity analysis of levelling data” by Aloisi et al. (2012)’. *Geophys J Int* 197:1403–1409. doi:10.1093/gji/ggu116
- Amoruso A, Crescentini L, Scarpa R (2002) Source parameters of the 1908 Messina Straits, Italy, earthquake from geodetic and seismic data. *J Geophys Res* 107(B4):2080. doi:10.1029/2001JB000043
- Basili R, Valensise G, Vannoli P, Burrato P, Fracassi U, Mariano S, Tiberti MM, Boschi E (2008) The Database of Individual Seismogenic Sources (DISS), version 3: summarizing 20 years of research on Italy’s earthquake geology. *Tectonophysics*. doi:10.1016/j.tecto.2007.04.014
- Boccaletti M, Nicolich R, Tortorici L (1984) The Calabrian Arc and the Ionian Sea in the dynamic evolution of the central Mediterranean. *Mar Geol* 55:219–245
- Boschi E, Pantosti D, Valensise G (1989) Modello di sorgente per il terremoto di Messina del 1908 ed evoluzione recente dell’area dello Stretto, in *Atti VIII Convegno G.N.G.T.S., Roma*, pp 245–258
- Bottari A (2008) Osservazioni macrosismiche e studi, in *Riassunti estesi del Convegno 1908–2008 Scienza e Società a cento anni dal Grande Terremoto, Reggio Calabria 10–12 dicembre 2008*, Miscellanea I.N.G.V., 3:19–20
- Bottari A, Carapezza E, Carapezza M, Carveni P, Cefali F, Lo Giudice E, Pandolfo C (1986) The 1908 Messina Strait earthquake in the regional geostructural frame work. *J Geodyn* 5:275–302
- Capuano P, De Natale G, Gasparini P, Pingue F, Scarpa R (1988) A model for the 1908 Messina Straits (Italy) earthquake by inversion of leveling data. *Bull Seism Soc Am* 78:1930–1947
- Console R, Carluccio R, Papadimitriou E, Karakostas V (2015) Synthetic earthquake catalogs simulating seismic activity in the Corinth Gulf, Greece, fault system. *J Geophys Res* 120(1):326–343. doi:10.1002/2014JB011765
- Cornell CA (1968) Engineering seismic risk analysis. *Bull Seism Soc Am* 58:1583–1606
- D’Agostino N, D’Anastasio E, Gervasi A, Guerra I, Nedimović Mladem R, Seeber L, Steckler M (2011) Forearc extension and slow rollback of the Calabrian arc from GPS measurements. *Geophys Res Lett* 38:L17304. doi:10.1029/2011GL048270
- De Natale G, Pingue F (1991) A variable slip fault model for the 1908 Messina Straits (Italy) earthquake, by inversion of leveling data. *Geophys J Int* 104:73–84
- De Natale G, Pino NA (2014) Comment on ‘Are the source models of the M 7.1 1908 Messina Straits earthquake reliable? Insights from a novel inversion and sensitivity analysis of levelling data’ by M. Aloisi, V. Bruno, F. Cannavò, L. Ferranti, M. Mattia, C. Monaco and M. Palano. *Geophys J Int* 197:1399–1402. doi:10.1093/gji/ggu063
- DISS Working Group (2015) Database of Individual Seismogenic Sources (DISS), Version 3.2.0: A compilation of potential sources for earthquakes larger than M 5.5 in Italy and surrounding areas. Istituto Nazionale di Geofisica e Vulcanologia. <http://diss.rm.ingv.it/diss/>. doi:10.6092/INGV.IT-DISS3.2.0
- Field EH (2015) Computing elastic-rebound-motivated earthquake probabilities in unsegmented fault models: a new methodology supported by physics-based simulators. *Bull Seism Soc Am* 105(2A):544–559. doi:10.1785/0120140094
- Galli P, Bosi V (2004) Catastrophic 1638 earthquakes in Calabria (southern Italy): new insights from paleoseismological investigation. *J Geophys Res* 108(B1):2003. doi:10.1029/2001JB001713
- Ghisetti F (1984) Recent deformations and the seismogenic source in the Messina Straits (southern Italy). *Tectonophysics* 109:191–208
- Ghisetti F (1992) Fault parameters in the Messina Strait (southern Italy) and relations with the seismogenic source. *Tectonophysics* 210:117–133
- Hanks TC, Kanamori H (1979) A moment magnitude scale. *J Geophys Res* 84(B5):2348–2350
- Mattei M, Sagnotti L, Faccenna C, Funicello R (1997) Magnetic fabric of weakly deformed clay-rich sediments in the Italian peninsula: relationships with compressional and extensional tectonics. *Tectonophysics* 271:107–122
- Mele FM, Marocci C, Bono A, Marchetti A (2010) ISIDe, Italian seismological instrumental and parametric data-base. INGV, CNT. <http://iside.rm.ingv.it/iside/standard/index.jsp>
- Minelli L, Vecchio A, Speranza F, Nicolosi I, D’Ajello Caracciolo F, Chiappini S, Carluccio R, Chiappini M (2015) Aeromagnetic investigation of Southern Calabria and the Messina Straits (Italy): tracking seismogenic sources of 1783 and 1908 earthquakes. *J Seism Res Solid Earth*. doi:10.1002/2015JB012305

- Mulargia F, Boschi E (1983) The 1908 Messina earthquake and related seismicity. In: Kanamori H, Boschi E (eds) Earthquakes: observation, theory and interpretation, proceedings of the international school of physics Enrico Fermi, course, vol 85. Elsevier, Amsterdam, pp 493–518
- Murru M, Akinci A, Falcone G, Pucci S, Console R, Parsons T (2016) $M \geq 7$ earthquake rupture forecast and time-dependent probability for the sea of Marmara Region, Turkey. *J Geophys Res*. doi:[10.1002/2015JB012595](https://doi.org/10.1002/2015JB012595)
- Parsons T, Geist EL (2009) Is there basis for preferring characteristic earthquakes over Gutenberg–Richter distributions on individual faults in probabilistic earthquake forecasting? *Bull Seism Soc Am* 99:2012–2019. doi:[10.1785/0120080069](https://doi.org/10.1785/0120080069)
- Parsons T, Console R, Falcone G, Murru M, Yamashina K (2012) Comparison of characteristic and Gutenberg–Richter model for time-dependent $M \geq 7.9$ earthquake probability in the Nankai-Tokai subduction zone, Japan. *Geophys J Int*. doi:[10.1111/j.1365-246X.2012.05595.x](https://doi.org/10.1111/j.1365-246X.2012.05595.x)
- Presti D, Billi A, Orecchio B, Totaro C, Faccenna C, Neri G (2013) Earthquake focal mechanisms, seismogenic stress and seismotectonics of the Calabrian Arc, Italy. *Tectonophysics* 602:153–175
- Rovida A, Camassi R, Gasperini P, Stucchi M (eds) (2011) CPTI11, the 2011 version of the Parametric Catalogue of Italian Earthquakes. INGV, Milano, Bologna. doi:<http://doi.org/10.6092/INGV.IT-CPTI11>, <http://emidius.mi.ingv.it/CPTI11>
- Sabetta F, Pugliese A (1987) Attenuation of peak horizontal acceleration and velocity from Italian strong motion records. *Bull Seism Soc Am* 77:1491–1513
- Schick R (1977) Eine seismotektonische Bearbeitung des Erdbebens von Messina im Jahre 1908. *Geologisches Jahrbuch, Reihe E* 11, 74 pp
- Serpelloni E, Bürgmann R, Anzidei M, Baldi P, Mastrolembo Ventura B, Boschi E (2010) Strain accumulation across the Messina Straits and kinematics of Sicily and Calabria from GPS data and dislocation modeling. *Earth Plan. Sci Lett* 298:347–360. doi:[10.1016/j.epsl.2010.08.005](https://doi.org/10.1016/j.epsl.2010.08.005)
- Tortorici L, Monaco C, Tansi C, Cocina O (1995) Recent and active tectonics in the Calabrian arc (Southern Italy). *Tectonophysics* 243:37–55
- Tullis TE (2012) Preface to the focused issue on earthquake simulators. *Seism Res Lett* 83(6):957–958
- Valensise G, Pantosti D (1992) A 125-Kyr-long geological record of seismic source repeatability: the Messina Straits (southern Italy) and the 1908 earthquake (MS 7 1/2). *Terra Nova* 4:472–483
- Valensise G, Basili R, Burrato P (2008), La sorgente del terremoto del 1908 nel quadro sismotettonico dello Stretto di Messina, in *Terremoto di Messina del 1908. Sorgente sismogenetica*, edited by G. Bertolaso et al., pp 161–182, Dipartimento Protezione Civile, Istituto Nazionale di Geofisica e Vulcanologia, Roma
- Wells DL, Coppersmith KJ (1994) Empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. *Bull Seism Soc Am* 84:974–1002
- Westaway R (1992) Seismic moment summation for historical earthquakes in Italy: tectonic implications. *J Geophys Res* 97(B11):15437

High-resolution computed microtomography for the characterization of a diffusion tensor imaging phantom

Łukasz Kaczmarek^{1,2}  · Tomasz Wejrzanowski² · Jakub Skibiński² · Michał Maksimczuk² · Artur Krzyżak³

Received: 26 December 2016 / Accepted: 17 February 2017 / Published online: 6 March 2017
© The Author(s) 2017. This article is published with open access at Springerlink.com

Abstract This paper addresses the issue of the quantitative characterization of the structure of the calibration model (phantom) for b-matrix spatial distribution diffusion tensor imaging (BSD-DTI) scanners. The aim of this study was to verify manufacturing assumptions of the structure of materials, since phantoms are used for BSD-DTI calibration directly after manufacturing. Visualization of the phantoms' structure was achieved through optical microscopy and high-resolution computed microtomography (μ CT). Using μ CT images, a numerical model of the materials structure was developed for further quantitative analysis. 3D image characterization was performed to determine crucial structural parameters of the phantom: porosity, uniformity and distribution of equivalent diameter of capillary bundles. Additionally calculations of hypothetical flow streamlines were also performed based on the numerical model that was developed. The results obtained in this study can be used in the calibration of DTI-BST measurements. However, it was found that the structure of the phantom exhibits flaws and discrepancies from the assumed geometry which might affect BSD-DTI calibration.

Keywords Phantom · Internal structure · Non-destructive analysis · Numerical models

Introduction

BSD-DTI is one of the magnetic resonance imaging (MRI) approaches to diffusion tensor imaging, taking into consideration the actual spatial distribution of the b-matrix, which corresponds to the distribution of the magnetic field gradients (Krzyżak and Olejniczak 2014). MRI techniques are used for porous material saturation analysis and diffusion study. This technique is applied in various issues from medical investigations (e.g., Tomanek et al. 1996; Das 2004 or Krzyżak et al. 2005, 2008) to petroleum geology reservoir rock characterization (e.g., Wei et al. 2015; Xiao et al. 2013). The principles and basis of classic MRI technology were described in detail by Callaghan (1994), Straley et al. (1995) and Coates et al. (1999). Results obtained through MRI depend, to a large extent, on scanner properties, sequence parameters and characteristics of samples. To obtain accurate MRI results, the scanner needs to be properly calibrated, involving the determination of the relation of T1, T2 relaxation times with sample porosity. The calibration process requires the application of phantoms: samples with a precisely defined structure (Komlosh et al. 2011). Furthermore, in BSD-DTI, well-defined anisotropic phantoms are central to improving accuracy and determining diffusion gradient directions and magnitudes (Komlosh et al. 2011). This is due to impact of the inhomogeneity of the analyzed material on the inhomogeneity of the magnetic field, which influences the relaxation processes and timing of each sequence. In general, the registered MRI signal errors of the calibration phantom are composed of two main components: (1)

✉ Łukasz Kaczmarek
lukasz.kaczmarek@uw.edu.pl;
Lukasz.Kaczmarek@wimpw.edu.pl

¹ Faculty of Geology, University of Warsaw, Żwirki i Wigury 93, 02-089 Warsaw, Poland

² Faculty of Materials Science and Engineering, Warsaw University of Technology, Wołoska 141, 02-507 Warsaw, Poland

³ Faculty of Geology, Geophysics and Environmental Protection, University of Science and Technology in Kraków, Mickiewicza 30, 30-059 Kraków, Poland

systematic errors induced by the nonlinearity of imaging gradients (e.g., diffusion gradients), and (2) imperfections of the phantom. Knowing the exact structure of the phantom, we can accurately and spatially determine its effect on the signal, and thus identify the actual distribution of imaging gradients. This can significantly reduce the influence of the heterogeneity of the magnetic field gradients for MRI experiments, particularly in DTI. In principle, BSD-DTI assumes carrying out standard DTI imaging with at least six different directions of diffusion gradients (plus one without) for at least six positions of anisotropic phantom—three orthogonal, aligned with the laboratory frame axes and three oblique. The measurement results are sufficient to determine the diffusion tensor, assuming its symmetry. Due to the structural complexity of the phantoms and possible defects which may occur during the manufacturing process, the phantom obtained can differ from the designed geometry. Therefore, verification and quantitative description of phantom geometry is an essential step. Microtomography is a non-destructive tool that provides information about the internal structure of phantoms, enabling detailed analysis of the homogeneity and beam of capillaries geometry, and lending valuable insight for the development of numerical models. The theoretical basis of μ CT was very well described by Ketcham and Carlson (2001), Baker et al. (2012) and by Gnudde and Boone (2013) in the review articles. Examples of μ CT analysis of various types of porous media were presented in articles by Appoloni et al. (2007), Bielecki et al. (2013), Kaczmarek et al. (2015). μ CT has also been used for analysis of flow through rock specimens by: Petchsingto and Karpyn (2009), Dvorkin et al. (2009). Literature data and our own investigations indicate that μ CT analysis has a wide range of applications in geology.

This study presents μ CT analysis of phantoms' internal structure and its parameterization after the manufacturing process. The main goal of this study was to verify the structure of the material and identify potential discrepancies that could influence the BST-DTI calibration process. Subsequently, numerical simulations of fluid flow through the phantom sample were performed to study the linearity of the capillaries and influence of flaws on the streamlines of fluid inside the phantom. The geometry and size of the internal elements of phantoms depend on the specific MRI approach. Therefore, the presented methodology is a universal solution and can be adapted for use with a variety of phantom types.

Materials and methods

The internal structure of phantoms relates to the particular research technique employed. Accordingly, the presented study procedure can be easily adopted for use with a variety

of MRI scanners. For initial recognition, we used a light microscope followed by a non-destructive technique—computed microtomography (μ CT). Results obtained from μ CT were visualized with commercial software (SkyScan[®] 1.13.11.0 and Avizo[®] 8.0). Figure 1 shows a schematic chart of the multistage analysis that was conducted.

Sample

The cylindrical phantom dedicated for the calibration of BST-DTI scanners was constructed by the Military University of Technology in Warsaw. More detailed information about manufacturing can be found in Krzyzak et al. (2015). The linear structure is determined by capillary bundles. This structure enables lateral free diffusion along capillaries and restricts diffusion crosswise without any distortion, a crucial condition for the improvement of accuracy and better determination of the diffusion gradient directions and, therefore, the usefulness of phantoms. Figure 2 shows the investigated cylindrical glass phantom, which was filled with distilled water. The sample was a 3.6 cm long cylinder with a diameter of 3.4 cm, densely packed with hexagonal high precision acrylic glass bundles of capillaries (each with a diameter of approx. 920 μ m; total number equal 1502). Each bundle of capillaries consisted of dozens of capillaries (each with a diameter of approx. 36 μ m). The bundles of capillaries are enclosed in a glass sphere to prevent water loss.

Additionally, it should be noted that it is possible to define and manufacture various types of phantom structures and use them as models of the real structures of rocks. It is the characteristic features of the sedimentary rock structure like capillaries and lamination, which are important for flow media, as shown by Sun et al. (2015), where the relationship between the permeability, microfractures and the stresses in the rock is studied. These structural features can be reproduced in the phantoms by bundles of capillaries.

Imaging

Light microscopy

A Nikon EPIPHOT 200 was used for initial observations, providing images of the top or bottom surface of the

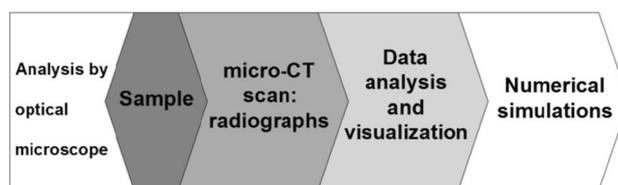


Fig. 1 Workflow of phantom analysis

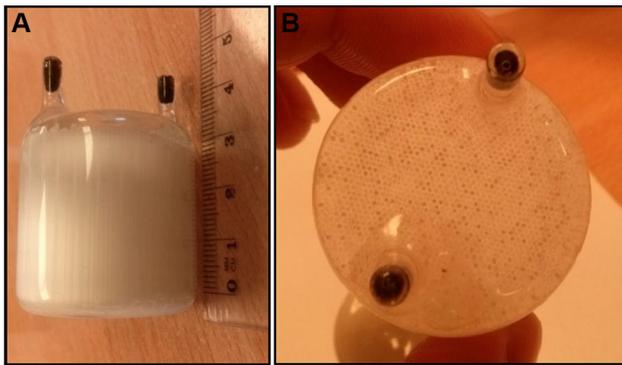


Fig. 2 **a** Sidewall of cylindrical phantom with visible longitudinal moving bundles of capillaries, **b** Phantom's upper surface with visible cellular cross-sections of bundles of capillaries

phantom. The advantage of this method was the ability to obtain images of the individual capillaries, through $75\times$ and $8\times$ visual magnification of sample. The drawback of this method was distorted and locally blurred images. Furthermore, light causes an increase in temperature of the sample during observation and longer observation can lead to the rupture of phantom glass.

X-ray microtomography

The main advantage of μ CT is the non-destructive character of internal structure analysis. High-resolution computed microtomography is based on imaging the variable coefficient of linear absorption of the material depending on density. We can also simulate fluid flow inside the sample on the basis of a numerical model obtained from high-resolution images. The numerical model obtained from tomography images precisely represents the geometry of the studied material. Xradia Micro XCT-400 was used for the analysis. The acquisitions were performed using a Hamamatsu L8121-03, which generates X-rays in the range of 40–150 kV. The scanned image is converted to digital data using a CCD video system, which has a resolution of 1024×1024 pixels, with 16-bit image depth of the detector. Selected samples were scanned using different acquisition parameters to select the appropriate settings for the technical equipment, producing images. A large field of view (LFOV) lens with 0.5 zoom was used during tests to obtain the broadest area of recognition. The manufacturer's LE3 custom glass filters were used for beam filtration. Scanning parameters are shown in Table 1.

Table 1 Parameters of X-ray microtomography

| Scan number | Voltage (kV) | Intensity (μ A) | Exposure time (s) | Type of objective (–) | Voxel size (μ m) |
|-------------|--------------|----------------------|-------------------|-----------------------|-----------------------|
| 1 | 110 | 90 | 2 | LFOV | 34 |
| 2 | 140 | 90 | 9 | LFOV | 11 |

Image processing

During radiograph processing, noise removal processes were performed, i.e., beam hardening (0.8) and center shift (2.0). The reconstruction parameters were binning 1 with bright spots correction and smoothing (Kernel size = 0.5). Then radiographs were reconstructed to obtain a set of 1024 high-resolution bitmap images. Then, the images were processed from 16 bit to 8 bit to reduce the size of images for faster processing. Subsequently, the set of images was used to develop a numerical model enabling visualization of the samples. After reconstruction, the thresholding method of 8-bit (256 gray scale) images was performed to distinguish phantom bundles of capillaries. Owing to μ CT processing software limitations, the analysis area had to be reduced to the working area (region of interest—ROI) in the cuboid form. Additionally, it should be noted that the size of voxel is the result of the size of ROI. The sample was earmarked for further DTI studies, and so singular bundles of capillaries could not be cut off for additional μ CT analysis with smaller ROI and smaller voxel size. In the study, the voxel size of the analyzed images was $11 \mu\text{m}$. ROI was selected in the center of the Phantom ($13 \times 16 \times 10 \text{ mm}$) and the prepared images were smoothed using a 3D median filter. A filter was used at a very low level (2 pixels), with the aim being to remove single pixels. The pixels present in the image were the result of the binarization. The low level of median filter, on the other hand, meant that the sharp edges of the phantom structure could be retained. Based on a binary set of image slices, a three-dimensional numerical model could be generated. A schematic chart of image preparation for numerical analysis is shown in Fig. 3.

μ CT analysis revealed information regarding the incomplete adhesion of capillary bundles, which was a discrepancy from the design assumptions. Due to these defects, empty channels without capillary bundles were created. Image analysis was performed that pertains to this issue, as shown in Fig. 3.

Phantom parameters

One of the most important parameters describing the storage capacity of rock is porosity. This paper concerns the porosity parameter considered as a fraction of voxels assigned to the pore regions with regard to the total number of voxels: $V_v = V_p/V$ where: V_v is the porosity, V_p is the pore volume and V is the total volume of the sample.

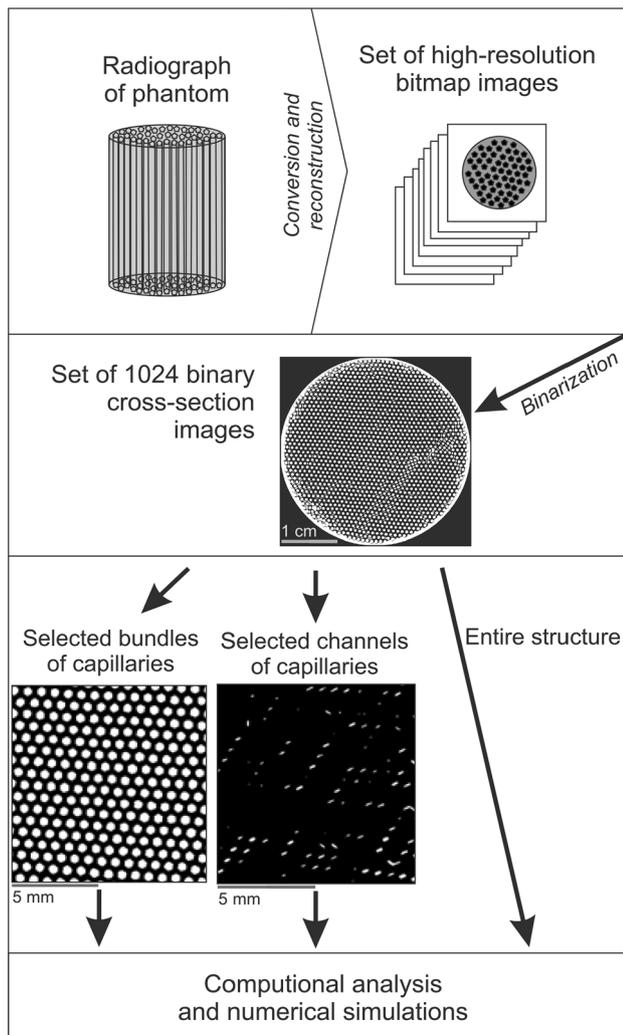


Fig. 3 Schematic chart of quantitative characterization and visualization of the phantom structure

The structure of the phantom with directional anisotropy could be described by the same parameters, based on the stereology of materials (Wejrzanowski et al. 2008). The mean values of calculated parameters used to describe the 3D internal structure of the phantom are shown in Table 2.

The homogeneity of the phantom was analyzed using Voronoi tessellation (Aurenhammer 1991). Cross-sections capillary bundles exhibit a cellular structure and, thus, the distribution of each “cell” size is expressed by the statistical parameter – coefficient of variation $CV(f)$, which described the homogeneity of the studied sample.

Fluid flow simulation

Numerical simulations of fluid flow and the subsequent visualization of streamlines (flow paths) were performed

using the finite volume method (FVM) employing the developed model. Based on μ CT results, we created a model of the sample for application in numerical simulations of fluid flow through a phantom structure.

The continuity and Navier–Stokes equations for the flow of incompressible Newtonian fluid at steady-state can be expressed as:

$$\rho u \cdot \nabla u = -\nabla p + \mu \nabla^2 u, \quad (1)$$

$$\nabla \cdot u = 0, \quad (2)$$

where μ and ρ are the viscosity and density of the fluid, while u and p are local velocity and pressure, respectively. To solve the above equations and calculate fluid flow in the studied sample, Avizo software was employed using FVM. This method transforms partial differential equations into algebraic ones. Values are calculated at discrete places on meshed geometry.

The integral form of the conservation law can be rewritten using the Gauss divergence theorem:

$$\partial/\partial t \int_{\Omega} U d\Omega + \int_{\Omega} \vec{\nabla} \cdot \vec{F} d\Omega = 0, \quad (3)$$

where U represents a vector of states and F represents the corresponding flux tensor. The volume integrals of divergence terms in a partial differential equation are converted to surface integrals of fluxes all around the control volume (4).

$$\partial/\partial t \int_{\Omega} U d\Omega + \oint_{\partial\Omega} \vec{F} \cdot \vec{n} d\Gamma = 0, \quad (4)$$

where Γ represents the total surface area of the cell and n is a unit vector normal to the surface and pointing outward. The resulting equations can be solved numerically via discretization into a set of discrete volumes which can be well defined for the porous structures in question.

In the case of further detailed analysis, it is possible to perform a digital experiment of permeability. The test is based on fluid flow simulation through a numerical model of the rock sample. Due to the previous μ CT scanning of the sample, it is possible to reconstruct the exact shapes of the flow paths in the model. A similar study of permeability determination was presented by Dvorkin et al. (2009).

In this study, the flow across both designed and real structures was simulated using Avizo software with the computational domain of length $13 \times 16 \times 10$ mm (Fig. 4). The inlet velocities for FVM simulations correspond to the Reynolds number from 1 to 10. The Newtonian fluid flowing through the channel was oil, in accordance with the archive viscosity data of oil to mirror real flow conditions from the Baltic Basin shale formations in Poland (Twarduś and Nowicka 2014).

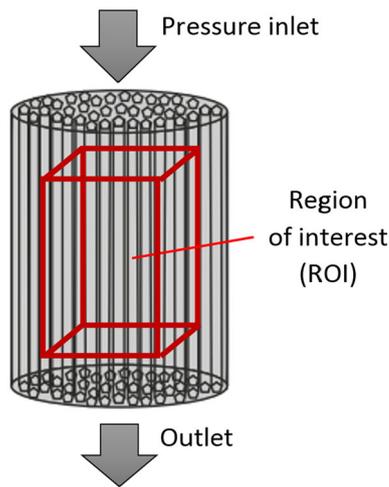


Fig. 4 Scheme of numerical simulation of flow

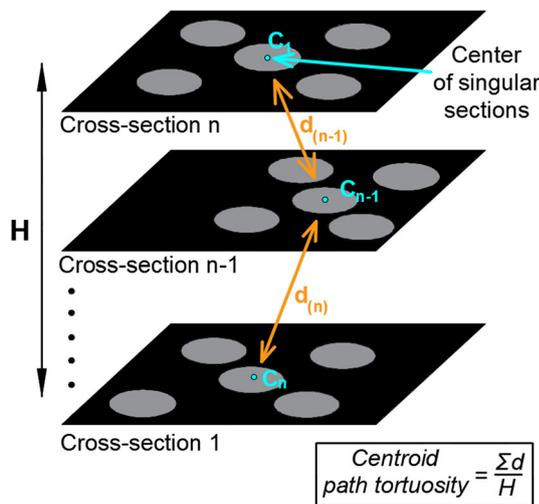


Fig. 5 Scheme of centroid path tortuosity calculation

A mesh consisting of 988,000 elements was generated based on the geometry of the studied sample, which provided an accurate representation of the internal structure without distorting generalizations. Calculations

were performed to visualize streamlines, enabling qualitative verification of the linearity of the structure of capillaries and comparison with rock fissures. Furthermore, the tortuosity can also be calculated, based on the centroid path tortuosity module in AVIZO software. In reference to the software calculation procedure, this general parameter is understood as the ratio of the real length of the flow path to the straight line between the beginning and the end of flow (Fig. 5). Furthermore, Petchsingto and Karpyn (2009) have shown that tortuosity can be interpreted in relation to the pressure gradient profile of flow.

Results and discussion

Light microscopy

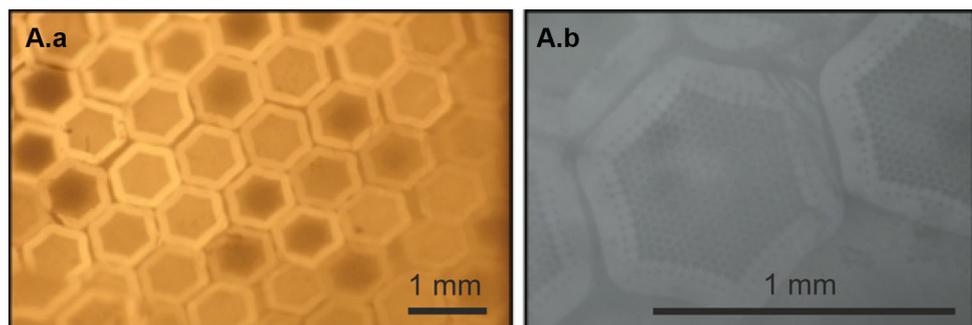
The glass capillary array phantom is shown in Fig. 6 as a photographic enlargement and microscopic image. The cross-section of the phantom reveals its regular internal structure. Owing to the lack of thin wall adhesion, small channels are observed among the capillary bundles. The greatest restriction on optical microscopy image analysis was the thickness of the phantom’s protective casing, which resulted in image blurring.

During the first stage of the research, we could observe slight movement of the capillary bundles in the phantom. The narrow edges of the bundles peeling off from the outer casing caused this damage.

X-ray microtomography and numerical simulations

During analysis of the phantom, we observed disorders of the structure of capillary bundles. Firstly, the shifts between bundles created diversified channels (Fig. 7). Secondly, gaps between bundles were observed. Additionally, size differences were noticed during the image analysis. To accurately describe geometric structure relations, parameters were chosen that described the internal

Fig. 6 Optical microscopy image of sample: A.a. capillary bundles of cylindrical phantom; A.b. singular capillaries of cylindrical phantom



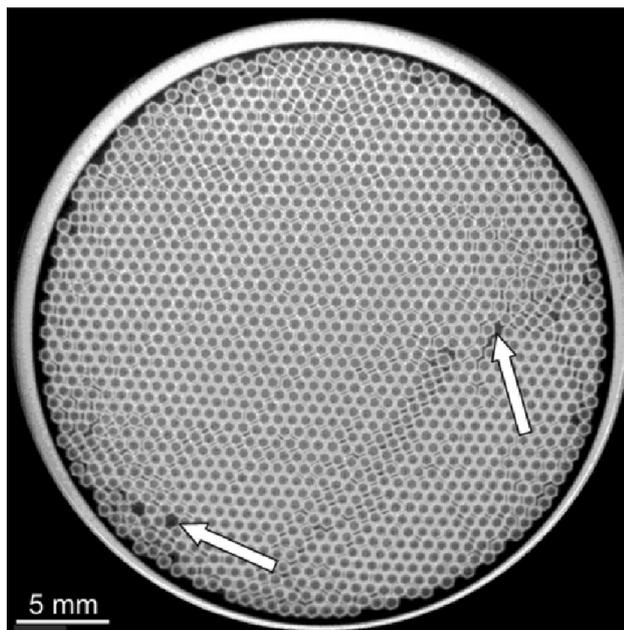


Fig. 7 Cross-section of phantom. Arrows indicate examples of channels and gaps of capillary bundles

structure of phantoms: porosity, surface of channels and capillary bundles, wall thickness of the capillary bundles, equivalent inner diameters (without walls) of channels and capillary bundles, and capillary bundle elongation.

The parameters obtained through the quantification analysis are listed in Table 2. It is noteworthy that the fine channel share is nearly 8% of the entire phantom pore space. The bundles of capillaries are characterized by a relatively high uniformity of equivalent diameter, which is nearly 2.7 times greater than that in the case of fine channels. The large standard deviation of channel equivalent diameters indicates their contingency. The structural orientation of the phantom causes the elongation of defects, which confirms a maximum diameter 2 times greater than the minimum diameter of fine channels. Such disturbances may generate disorder diffusion.

Complementary figures of parameters describing the structure of the phantom, showing various percentage ranges of equivalent inner diameter distribution of capillary bundles and channels, are presented in Figs. 8 and 9.

During the fabrication procedure of phantom, the diameter of bundles was designed to be 920 μm . Microtomography research results indicate, however, that the average value of the inner diameter of bundles is 508 μm and standard deviation of this parameter is 14 μm , whereas the average thickness of walls is 166 μm with an SD equal to 18 μm . In total, the diameter of bundles obtained with the use of microtomography is 840 μm . Taking into account standard deviation values, the diameter of the bundles can reach 872 μm . The difference of 48 μm between the expected and obtained value (taking into

Table 2 Parameterization of phantom microstructure

| Parameter | Unit | Results of the analysis | | |
|---------------------------|-------------------|-------------------------|------------------------|---|
| | | Entire phantom | Bundles of capillaries | Channels between the walls of the capillary bundles |
| Global parameters | | | | |
| V_v | % | 38 | 35 | 3 |
| S_v | 1/mm | 2.29 | 1.96 | 0.32 |
| N_A | 1/mm ² | 2.85 | 1.98 | 0.87 |
| DA | – | 0.89 | 0.89 | 0.90 |
| Local parameters | | | | |
| $E(T)$ | μm | – | 166 | – |
| SD(T) | | | 18 | |
| CV(T) | | | 0.11 | |
| $E(D_C)$ | μms | 437 | 508 | 191 |
| SD(D_C) | | 145 | 14 | 77 |
| CV(D_C) | | 0.33 | 0.06 | 0.40 |
| $E(D_{\max}/D_{\min})$ | – | 1.34 | 1.11 | 2.01 |
| SD(D_{\max}/D_{\min}) | | 0.50 | 0.03 | 0.59 |
| CV(D_{\max}/D_{\min}) | | 0.37 | 0.02 | 0.29 |

Structural parameters: V_v , the ratio of pore volume to total volume (porosity); S_v , the surface of channels and capillary bundles to the total volume of the analyzed area; N_A , both the relative multiplicity of channels and/or capillary bundles per unit area; DA, the degree of anisotropy, where for fully isotropic system DA = 0, and for the fully anisotropic DA = 1 (<http://www.skyscan.be/>); T , wall thickness of the capillary bundles; D_C , the equivalent inner diameters (without walls) of channels and/or capillary bundles; D_{\max}/D_{\min} , the capillary bundle elongation (where D_{\max} is a maximum diameter, and D_{\min} is a minimum diameter). Statistical parameters: $E(x)$, mean value of variable x ; SD(x), standard deviation of variable x ; CV(x) = SD(x)/ $E(x)$ coefficient of variation of variable x

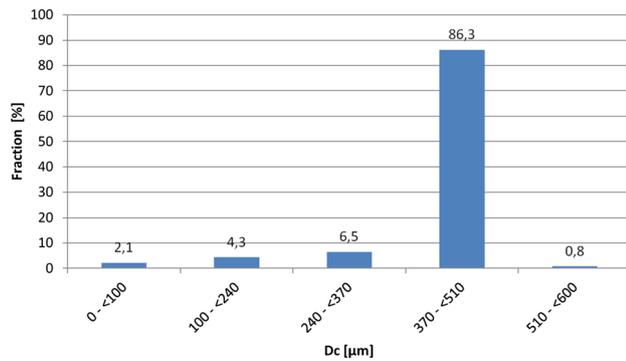


Fig. 8 Total distribution of equivalent inner diameters of capillary bundles and channels for the phantom

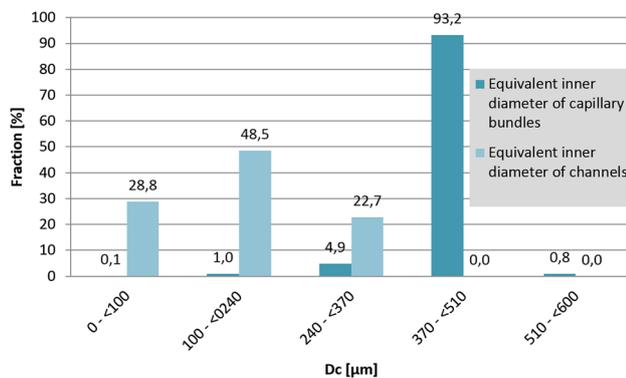


Fig. 9 Phantom: equivalent inner diameter distribution of capillary bundles compared to the equivalent inner diameter of channels

account SD values) might stem from the production process, involving the dense packing of bundles. This fact should be considered in the MRI and DTI studies, due to the relationship between the accuracy of the tests in terms of homogeneity and the linearity of the calibration of the phantom structure. We should also pay attention to the

unimodal distribution of capillary bundle diameters, which indicates a favorable distribution of the dominant value closest to the expected one.

The resolution of μ CT was not sufficiently good to analyze a singular capillary. This was mainly due to the fact that the samples should not have been broken, thereby precluding individual investigation of individual capillaries. Nevertheless, since the size and geometry of capillary bundles are the result of individual capillaries, the internal structure of phantoms can be characterized by the geometrical relations of bundles.

The next stage of the research was to produce an assessment of homogeneity, which was performed by comparing the area (influence zone, A) obtained by binarization and tessellation image processing transformations. Figure 10 shows the homogeneity assessment process. The results are: $CV(A) \approx 0.57$; $E(A) \approx 0.30 \text{ mm}^2$; $SD(A) \approx 0.17 \text{ mm}^2$, where A is the mean size of influence zone. The results of analysis for the bundles of capillaries only are: $CV(A) \approx 0.04$; $E(A) \approx 0.47 \text{ mm}^2$; $SD(A) \approx 0.02 \text{ mm}^2$.

Analysis of the obtained results led to an explicit determination of phantom structure and indicated a homogeneity with slight flaws, which should not influence the lateral direction of the diffusion of particles visualized in BST-DTI. The homogeneity of phantom structure determines its applicability as a reference for the comparison of real non-linear capillaries and flow paths of hydrocarbons in rocks.

Fluid flow simulations

The previous stages, in the case of a cylindrical sample, enabled the visualization of capillary bundles and fine channels. In the next step, we created a three-dimensional numerical model of the cylindrical phantom (Fig. 11) to study fluid flow across the sample. The aim of the

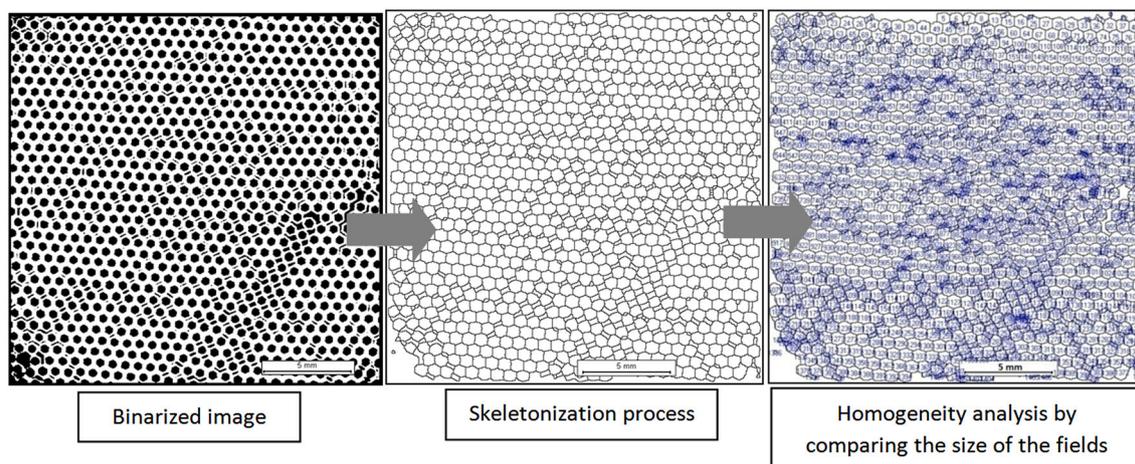


Fig. 10 Transformation process involved in analysis of phantom homogeneity

Fig. 11 Phantom visualization: **a** binarized cross-sections; **b** capillaries without deleted channels

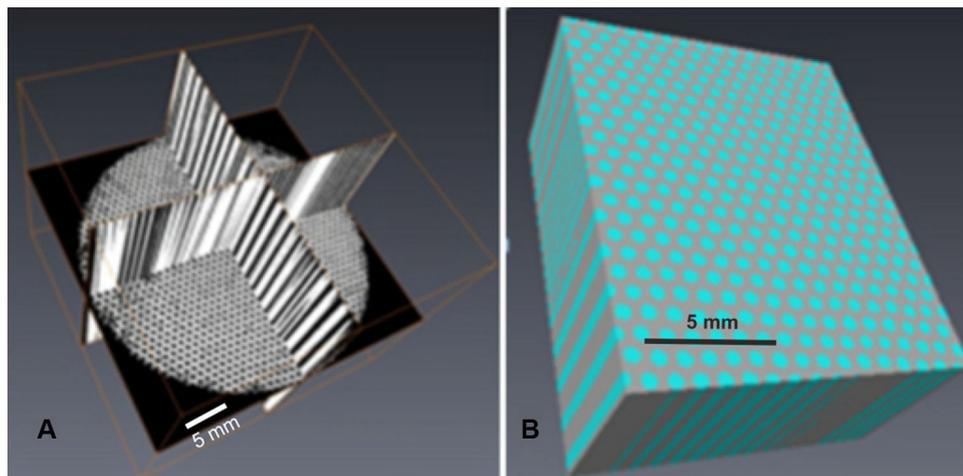
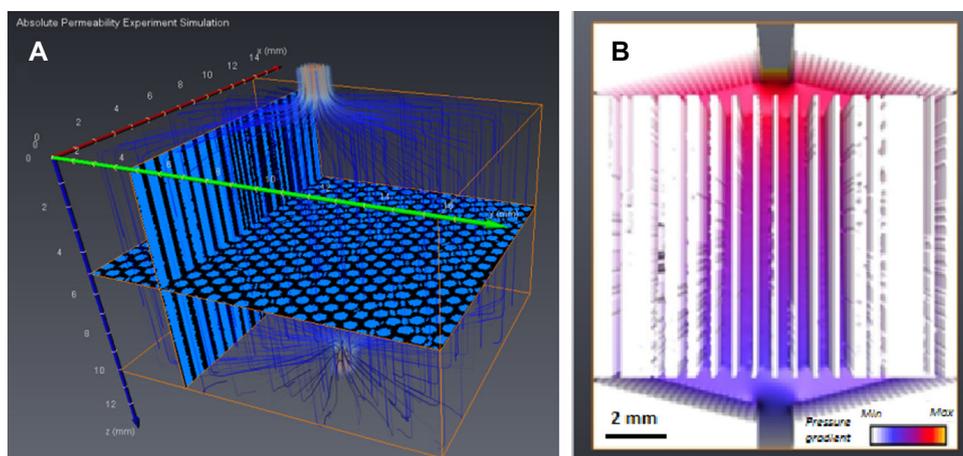


Fig. 12 Visualization of **a** streamlines, **b** pressure gradient in phantom



simulations was to verify the linearity of the capillaries. In the case of straight, undisturbed and undamaged bundles of capillaries, simulation of fluid flow should result in a linear distribution of streamlines and in a uniform pressure drop in the phantom.

It should be noted that diffusion tensor imaging (DTI) visualizes the movement of hydrogen nuclei. In DTI, anisotropic phantoms are crucial for precise determination of the diffusion tensor components. The resulting visualization of the flow streamlines based on μ CT study (Fig. 12) is used to identify potential path movements of nuclei in the DTI method. This approach can be used to verify the impermeability of the materials analyzed, due to the high precision of DTI techniques (even at the nanometric scale). For example, in the case of fractures in the rock samples detected by μ CT, one can analyze the continuity and propagation of potential media; as in the second step, it is possible to verify the results by DTI. In the case of overlapping streams flow obtained by DTI and by simulation based on μ CT, the low permeability of the material can be predicted.

As shown in Fig. 12, the linearity of the flow is maintained. Moreover, the centroid path tortuosity for beams of capillaries and channels is 1.015, which also confirms the linearity of the phantom structure. Likewise, uniform pressure changes provide information that there were no larger defects that could interfere with straightforward flow. In the case of contrary results, further more detailed calculation should be performed. The linearity of the sample is the basic criterion that needs to be fulfilled for further proper DTI calibration; therefore, the results obtained indicate the utility of phantoms for MRI scanner calibration. The inhomogeneity of the materials structure, in the form of fine channels and shifts of capillary bundles, revealed during characterization with the use of μ CT and light microscope suggested that the linearity of the sample could be disturbed. However, the numerical simulations performed indicate that the sample can be successfully applied in further DTI testing.

Furthermore, there is the question of the influence of capillary walls on flow velocity, which should be studied in future considerations. Another question regards possibly

scaling up the study of the properties of the reservoir to the global scale. Here, Dvorkin et al. (2009) suggested a positive correlation analysis of reservoir rocks at different scales.

Conclusions

This study presents universal internal structure parameterization and visualization guideline analysis for BST-DTI calibration phantoms. MRI scanners require a calibration process, which involves determination of the relation between T1, T2 relaxation times distribution, signal magnitudes and sample porosity. Such calibration is possible by the precise definition of DTI phantoms. However, the manufacturing process of the materials can introduce some flaws or discrepancies from the designed geometry, which could seriously skew MRI scanner calibration. The results presented here demonstrate that high-resolution X-ray computed microtomography is an adequate non-destructive technique for the characterization of phantoms designed for rock analysis. The quantitative analysis procedure proposed here serves as a tool for assessing the applicability of phantoms to the application of DTI for the characterization of specific rock structures. According to this procedure, we were able to prove the usefulness of a cylindrical phantom for the calibration of MRI scanners and further in-depth analysis.

The study clearly showed that computed microtomography can be performed without sophisticated treatment of the sample and may provide many parameters that describe the internal structure of phantoms (porosity, inner diameter of capillary bundles). Moreover, the anisotropy of the samples can be described. The must-have factor for this method is good contrast (based on linear coefficient of X-ray absorption) between the liquid and the structure of the phantom. The differences at all stages (acquisition, reconstruction, processing and analysis) show that long acquisition and reconstruction periods are required to obtain higher resolutions. The advantage of μ CT is that it delivers the high precision images of internal microstructure, which are used to create three-dimensional visualizations and simulations. The finite volume method proved to be a useful supplementary tool in the analysis of linearity of the sample, when characterization of the material indicated manufacturing flaws. The high-resolution X-ray computed microtomography analysis and numerical modeling with the use of a finite volume method reveal capabilities for complementary use for DTI research technology.

Acknowledgements This work was supported by the National Center for Research and Development (Contract No. PBS2/A2/16/2013).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Appoloni C, Fernandes C, Rodrigues C (2007) X-ray microtomography study of a sandstone reservoir rock. *Nucl Instrum Methods Phys Res Sect A* 580:629–632. doi:10.1016/j.nima.2007.05.027
- Aurenhammer F (1991) Voronoi diagrams—a survey of a fundamental geometric data structure. *ACM Comput Surv* 23:345–405. doi:10.1145/116873.116880
- Baker D, Mancini L, Polacci M, Higgins M, Gualda G, Hill R, Rivers M (2012) An introduction to the application of X-ray microtomography to the three-dimensional study of igneous rocks. *Lithos* 148:262–276. doi:10.1016/j.lithos.2012.06.008
- Bielecki J, Jarzyna J, Bożek S, Lekki J, Stachura Z, Kwiatek W (2013) Computed microtomography and numerical study of porous rock samples. *Radiat Phys Chem* 93:59–66. doi:10.1016/j.radphyschem.2013.03.050
- Callaghan P (1994) Principles of nuclear magnetic resonance microscopy. Oxford University Press, Oxford
- Coates GR, Xiao L, Prammer MG (1999) NMR logging principles and applications. Halliburton Energy Services Publication, Houston
- Das S (2004) Nuclear magnetic resonance spectroscopy. *Resonance* 9:34–49
- Dvorkin J, Derzhi N, Fang Q, Nur A, Nur B, Grader A, Baldwin C, Tono H, Diaz E (2009) From micro to reservoir scale: permeability from digital experiments. *Lead Edge* 28:1446–1452. doi:10.1190/1.3272699
- Gnudde V, Boone M (2013) High-resolution X-ray computed tomography in geosciences: a review of the current technology and applications. *Earth Sci Rev* 123:1–17. doi:10.1016/j.earscirev.2013.04.003
- Kaczmarek Ł, Maksimczuk M, Wejrzanowski T, Krzyżak A (2015) High-resolution X-ray microtomography and nuclear magnetic resonance study of a carbonate reservoir rock. In: Proc. 15th International Multidisciplinary Scientific GeoConference SGEM 2015, June 18–24, 2015, Albena. doi:10.5593/SGEM2015/B11/S6.099
- Ketcham R, Carlson W (2001) Acquisition, optimization and interpretation of X-ray computed tomographic imagery: applications to the geosciences. *Comput Geosci* 27:381–400. doi:10.1016/S0098-3004(00)00116-3
- Komlosh ME, Özarlan E, Lizak MJ, Horkay F, Schram V, Shemesh N, Cohen Y, Basser PJ (2011) Pore diameter mapping using double pulsed-field gradient MRI and its validation using a novel glass capillary array phantom. *J Magn Reson* 208:128–135. doi:10.1016/j.jmr.2010.10.014
- Krzyżak AT, Olejniczak Z (2014) Improving the accuracy of PGSE DTI experiments using the spatial distribution of b matrix. *Magn Reson Imaging* 33:286–295. doi:10.1016/j.mri.2014.10.007
- Krzyżak AT, Jasiński A, Adamek D (2005) Qualification of the most statistically “sensitive” diffusion tensor imaging parameters for detection of spinal cord injury. *Acta Phys Pol Ser A* 108:207–210
- Krzyżak AT, Jasiński A, Kwieciński S, Kozłowski P, Adamek D (2008) Quantitative assessment of injury in rats spinal cord in vivo using MRI of water diffusion tensor. *Appl Magn Reson* 34:1–24. doi:10.1007/s00723-008-0095-7

- Krzyżak AT, Kłodowski K, Raszewski Z (2015) Anisotropic phantoms in magnetic resonance imaging. In: Proc. 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, August 25–29, 2015, MiCo, Milano
- Petchsingto T, Karpyn T (2009) Deterministic modeling of fluid flow through a CT-scanned fracture using computational fluid dynamic. *Energy Sources Part A Recovery Util Environ Effects* 31:897–905. doi:[10.1080/15567030701752842](https://doi.org/10.1080/15567030701752842)
- Straley C, Rossini D, Vinegar H, Tutunjian P, Morriss C (1995) Core analysis by low field NMR. SCA-9404, pp 43–56
- Sun Q, Xue L, Zhu S (2015) Permeability evolution and rock brittle failure. *Acta Geophys* 63(4):978–999. doi:[10.1515/acgeo-2015-0017](https://doi.org/10.1515/acgeo-2015-0017)
- Tomanek B, Jasiński A, Sulek Z, Muszyńska J, Kulinowski P, Kwieciński S, Krzyżak A, Skórka T, Kibiński J (1996) Magnetic resonance microscopy of internal structure drone and queen honey bees. *J Apic Res* 35:3–9. doi:[10.1080/00218839.1996.11100907](https://doi.org/10.1080/00218839.1996.11100907)
- Twarduś E, Nowicka A (2014) Dokumentacja wynikowa otworu badawczego Opalino. PGNIG, Piła
- Wei D-F, Liu X-P, Hu X-X, Xu R, Zhu L-L (2015) Estimation of permeability from NMR logs based on formation classification method in tight gas sands. *Acta Geophys* 63(5):1316–1338. doi:[10.1515/acgeo-2015-0042](https://doi.org/10.1515/acgeo-2015-0042)
- Wejrzanowski T, Spychalski WL, Roźniatowski K, Kurzydłowski KJ (2008) Image based analysis of complex microstructures of engineering materials. *Int J Appl Math Comput Sci* 18:33–39. doi:[10.2478/v10006-008-0003-1](https://doi.org/10.2478/v10006-008-0003-1)
- Xiao L, Zou C, Mao Z, Shi Y, Liu XP, Jin Y, Guo H, Hu X (2013) Estimation of water saturation from nuclear magnetic resonance (NMR) and conventional logs in low permeability sandstone reservoirs. *J Petrol Sci Eng* 108:40–51. doi:[10.1016/j.petrol.2013.05.009](https://doi.org/10.1016/j.petrol.2013.05.009)