DE GRUYTER OPEN

Noise Reduction in Radon Monitoring Data Using Kalman Filter and Application of Results in Earthquake Precursory Process Research

Mojtaba NAMVARAN[1] and Ali NEGARESTANI[2]

[1]Kerman Graduate University of Technology, Geophysics Department, Kerman, Iran; e-mail: m.namvaran@kgut.ac.ir (corresponding author)

[2]Kerman Graduate University of Technology, Electrical Engineering Department, Kerman, Iran; e-mail: a.negarestani@kgut.ac.ir

A b s t r a c t

Monitoring the concentration of radon gas is an established method for geophysical analyses and research, particularly in earthquake studies. A continuous radon monitoring station was implemented in Jooshan hot-spring, Kerman province, south east Iran. The location was carefully chosen as a widely reported earthquake-prone zone. A common issue during monitoring of radon gas concentration is the possibility of noise disturbance by different environmental and instrumental parameters. A systematic mathematical analysis aiming at reducing such noises from data is reported here; for the first time, the Kalman filter (KF) has been used for radon gas concentration monitoring. The filtering is incorporated based on several seismic parameters of the area under study. A novel anomaly defined as "radon concentration spike crossing" is also introduced and successfully used in the study. Furthermore, for the first time, a mathematical pattern of a relationship between the radius of potential precursory phenomena and the distance between epicenter and the monitoring station is reported and statistically analyzed.

Key words: radon anomaly, Kalman filter (KF), noise reduction, effective precursory (EP) ratio, seismic activity.

1. INTRODUCTION

The devastations (both financial losses and casualties) caused by frequent and high-magnitude earthquakes in Iran necessitate scientific studies for understanding the geophysical activities in this country. The Kerman province located in south east Iran is one of the most seismically active regions in Iran and the Middle East. This is caused by an influence of stress field that results from convergence and collision between the Arabia and Eurasia plates (Berberian *et al*. 2001). Concentration of the active faults with the north-north western (N-NW) to south-south eastern (S-SE) trend indicates stress-reservoir in this region (Berberian *et al*. 1984).

The Golbaf and Shahdad (located in Kerman province) right lateral strike-slip faults with the N-S trend have raised expectation about seismic reservoir-triggering in this area (Berberian *et al*. 1984). Over the past few decades, some parts of these faults have developed their seismic activities. According to the catalogue by Iranian Seismological Center (IRSC) and International Institute of Earthquake Engineering and Seismology (IIEES), this earthquake-prone region has experienced more than 15 seismic events with M6.0 and higher in the last seven decades (IRSC & IIEES earthquake database).

The radon monitoring in soil and groundwater is a useful tool to study the possible link between the variations in radon concentration and deformation phenomena due to tectonic activities (Ramola 2010, Seyis *et al*. 2010, Hashemi *et al*. 2013). An increasing or decreasing radon concentration before earthquake is apparently caused by movements deep within the earth, following stress, loading and rupture of rocks (Hauksson 1981). So, crustal radon-flux measurement along active faults can provide useful information about movements in micro-fractures of the crust. Decreases in radon concentration were registered, for instance, before the following earthquakes: Japan 1984-1988 (Wakita *et al*. 1991), Taiwan 2008 (Kuo *et al*. 2006), Turkey 2009-2010 (Ali Yalım *et al*. 2012).

In order to distinguish between the variations of radon concentration caused by earthquakes from those caused by other sources, different studies have been reported (Negarestani *et al*. 2002, Richon *et al*. 2004, Ali Yalım *et al*. 2012). Prior to performing any investigation in the field, it is necessary to perform a pre-processing analysis on data to reduce noise (Finkelstein *et al*. 1998). This enhances the reliability of data for pattern recognition between data parameters. A detailed survey was carried out to decide on a reliable noise reduction mechanism that would reflect the complexities of radon gas behavior (Finkelstein *et al*. 1998). Kalman filter (KF) was found to be a very reliable and yet novel approach for radon gas monitoring studies (Simon 2001). So, it has been successfully employed and the results are reported

here. The Kalman filter is an analytical model frequently used for state estimation in state space models (SSM) when the standard Gaussian noise assumption does not apply (Simon and Chia 2002). The linear state space model postulates that an observed time series is a linear function of a state vector and the law of motion for the state vector is first-order vector auto-regression (AR) (Simon and Chia 2002). In recent years, KF has become a very powerful, intelligent and computational tool widely used in signal processing applications such as noise reduction (Brailean *et al.* 1995, Fujimoto and Ariki 2000, Evensen 2003, Brabec and Jílek 2007). The main advantage of this approach is designing an efficient filter to control noisy systems. In other words, the basic idea of a KF is producing less noisy data from a full noisy data. According to KF algorithm, it removes noises by assuming a pre-defined model of a system (Kleinbauer 2004). This means that KF is employed to remove the noises which are mixed with signals during the measurement.

A continuous radon monitoring station for groundwater gas monitoring was set up in Jooshan hot-spring. After few months of monitoring, some anomalies in low concentration of radon could be observed. Hence, alongside the investigation of radon gas concentration, the objectives of this article were set to be:

❏ to reduce commonly observed noises on radon time-series data, using KF,
❏ to define the anomaly shape and discuss a relation between filtered data and seismic parameters, and
❏ to find a correlation between radon decreasing rate and seismic parameters.

2.   EXPERIMENTAL PROCEDURE

Initially, a continuous monitoring of radon gas concentration was performed in Jooshan hot-spring complex situated near Jooshan village in Kerman province. The site location was carefully chosen: it is located between the two active faults Golbaf and Shahdad. The monitoring period was done in a seismically active season of the region from December 2011 to April 2012.

The main device used in this study is a RAD7 detector coupled with a measuring toolbox. The measured air is sucked from the water container using an air pump into the trap, then into the detector. The RAD7 radon monitor (DURRIDGE Company Inc., USA) is a commercial model, widely used in many applications involving continuous radon activities measurement. The detector counts the number of α-disintegrations into its chamber during a specified time (in this study it was 10 minutes).

The Kalman filter was then developed for noise reduction in the process. As this is an analytical model, like in any modeling study some assumption need to be made. Very careful considerations were made to make the re-

quired assumptions so that they would reflect the complexities of radon gas monitoring data in the closest and most realistic way. The KF model was then incorporated on the observational results using a commercial MATLAB® package, to analyze the noise reduction effect.

3. THEORETICAL HYPOTHESIS

3.1 Dobrovolsky equation

Dobrovolsky *et al*. (1979) suggested a theoretical-empirical relationship between size of the effective precursor manifestation zone and the main earthquake magnitude as:

$$D = 10^{0.43M} \; , \tag{1}$$

where M is the magnitude of the earthquake, and D is the effective radius of earthquake [km] called "strain radius". This equation was developed for estimating the deformation and tilts in surface of the earth as a magnitude function of the coming earthquake and distance from the epicenter. In other words, geochemical signals can only be caused by earthquakes with epicentral distances less than or equal to this empirical "magnitude-epicentral distance" relation (Dobrovolsky *et al*. 1979).

3.2 State space model

3.2.1 Non-linear SSM

State space model (SSM) refers to a class of probabilistic graphical models that describe the probabilistic dependence between the latent state variable (x) and the observed measurement (z) (Kleinbauer 2004). In other words, SSM is a mathematical simulation of a process, where the state of process is represented by a numerical vector. The state or the measurement can be either continuous or discrete. SSM is subdivided into two separate models (Simon 2001):

❑ the process model, which describes how the state propagates in time in response to the external influences (for example input and noise), and

❑ the measurement model, which describes how measurements (z) are taken from the process, typically simulating noisy and/or inaccurate measurement which could occur due to different reasons.

The most general from of an SSM is the non-linear state. The two main function of non-linear SSM are:

$$x_{t+1} = f\left(x_t, u_t, w_t\right) \; , \tag{2}$$

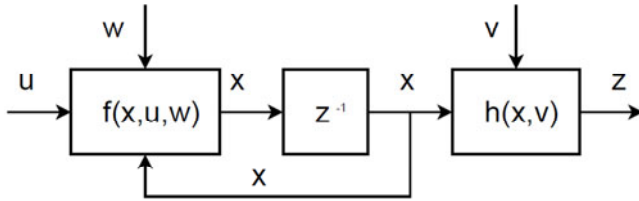$$z_t = h\left(x_t, v_t\right) \; , \tag{3}$$

Fig. 1. Fundamentals of the state space model (SSM). Notice: z^{-1} is z-transform in digital signal processing which governs unit delay function.

where f and h functions govern state propagation (x) and measurements (z), respectively, u is the process input, w and v are state and measurement noise vectors, respectively, and t is the discrete time. Functions f and h are usually based upon a set of discretized differential equations, governing the dynamics and observations during the process. Fundamentals of SSM are illustrated in Fig. 1.

3.2.2 Linear state space model (LSSM)

Suppose that a multiple parallel time series are observable in a process. Usually, the source series of interest are not directly measureable, but hidden in them. In addition, the mixing system generating the observable series from the source is unknown. For simplicity, the mixing system is assumed linear. The aim is to recover these sources, as well as to model their dynamics (Roesser 1975).

This is referred to as blind source separation (BSS) (Cichocki and Amari 2002). A matrix as A is assumed, which is the transfer function between sources and signal receivers. Most array processing techniques rely on the modeling of A: each column of A is assumed to depend on a small number of parameters. This information may be provided either by physical modeling or by direct array calibration. BSS consists in identifying A and/or retrieving the source signals without resorting to any *a priori* information about mixing matrix A; it exploits only the information carried by the received signals themselves, hence, the term "blind" (Cichocki and Amari 2002*)*. According to literature, statistical independence has played a great role in BSS; in most BSS algorithms the sources are assumed to be statistically independent (Belouchrani *et al*. 1997, Lee *et al*. 1999, Zibulevsky and Pearlmutter 2001). In the noiseless case, certain techniques have been proposed to solve this problem efficiently (Zhang and Hyvärinen 2011). For example, if the sources are non-Gaussian or at most one of them is Gaussian, BSS can be solved (Hyvärinen and Oja 2000). As it is explained in the following sections, the recorded data in this study showed the Gaussian distribution. Therefore, the above description could be applied to them (Fig. 2).
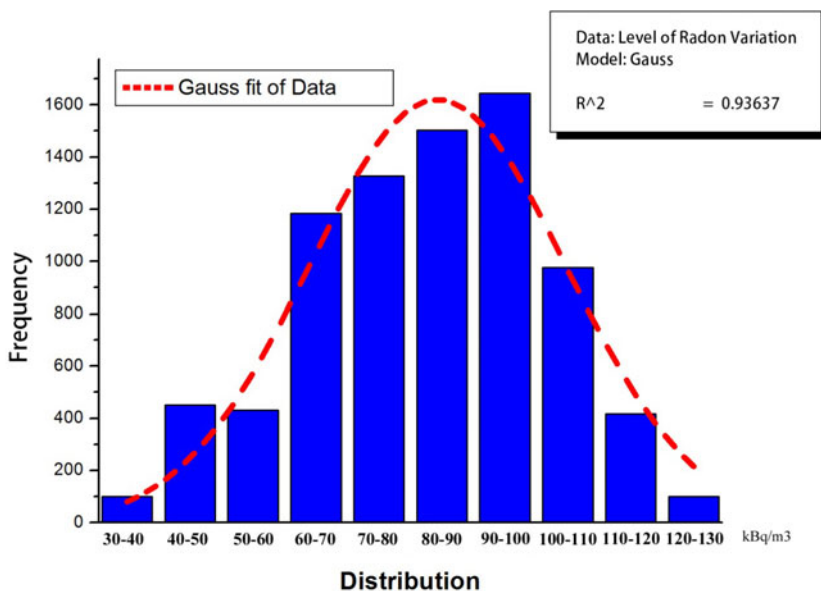
Fig. 2. Gaussian statistical distribution of radon gas concentration and its Gaussian fit. According to the distribution of this kind, BSS can be solved by the independent component analysis technique for these data.
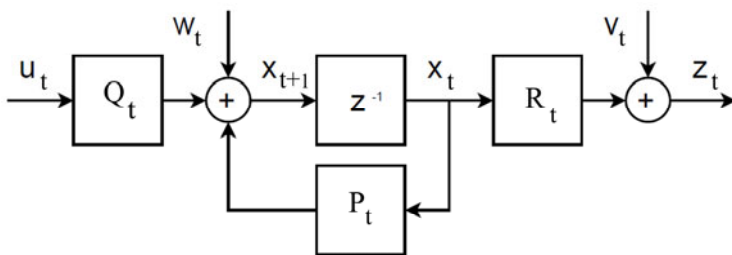

Fig. 3. Fundamentals of linear state space model (LSSM). This linear model is easier both to calculate and analyze.

Considering Fig. 3, an LSSM is a model where functions f and h are linear in both state and input. The functions can then be described by using the matrices P, Q, and R, reducing the state propagation calculations to linear algebra. Hence, the SSM would follow (Kleinbauer 2004):

$$x_{t+1} = P_t x_t + Q_t u_t + w_t \; , \tag{4}$$

$$z_t = R_t x_t + v_t \; . \tag{5}$$

This linear model is easier both to calculate and analyze and also enables investigating properties such as control-ability, observability, and frequency response.

3.3 A random walk process

When faced with a time series that shows irregular growth, the best strategy may not be to try to directly predict the level of the series at each period (*i.e.*, the quantity y_t). Instead, it may be better to try to predict the change that occurs from one period to the next (*i.e.*, the quantity $y_t - y_{t-1}$). In other words, it may be helpful to look at the first difference of the series, to see if a predictable pattern can be discerned there (Simon 2001, Spitzer 2001). For practical purposes, it is just as good to predict the next change as to predict the next level of the series, since the predicted change can always be added to the current level to yield a predicted level (Kleinbauer 2004, Spitzer 2001). Hence, the forecasting model suggested in this description is

$$y_t - y_{t-1} = \alpha \ , \tag{6}$$

where α is the mean of the first difference, *i.e.*, the average change from one period to the next. If we rearrange this equation to put y_t by itself on the left, we get

$$y_t = y_{t-1} + \alpha \ . \tag{7}$$

In other words, we predict that this period's value will equal the last period's value plus a constant representing the average change between periods (Spitzer 2001). This is the so-called "random walk" model. It assumes that, from one period to the next, the original time series merely takes a random step away from its last recorded position (Spitzer 2001).

According to the above description, a random walk is defined as a process where the current value of a variable is composed of the past value plus an error term known as "white noise" (ε). White noise is a random signal with a flat power spectral density and is called white because it affects all the frequency components of a signal equally (Spitzer 2001). Random walk is thereby analytically represented as:

$$y_t = y_{t-1} + \varepsilon_t \ . \tag{8}$$

The implication of a process of this type is that the best prediction of y for next period is the current value, or in other words the process does not allow predicting the change ($y_t - y_{t-1}$). That is, the change of y is absolutely random. It can be shown that the mean of a random walk process is constant but its variance is not. Therefore, a random walk process is non-stationary, and its variance increases with discrete time (t).

Considering the radon migration to follow random walk, this mechanism is used to model radon diffusion and migration through the earth crust. The soil – as a radon propagation medium prior to monitoring – is represented by a system of randomly oriented baffles. The mean distance d over which the atom travels between two collisions takes on the role of a mean free path. The effective mean time between two collisions, or in other words the migration time/orientation/velocity, is strongly random and depends on medium sectional properties, such as temperature, humidity, porosity, permeability of soil, *etc*.

3.4 Kalman filter

The KF is a recursive predictive filter which is based on the use of state-space technique and recursive algorithms. This filter estimates the state of a "dynamic system". This dynamic system can be disturbed by noise of different kind, mostly assumed as white noise (white noise is a random signal with a flat power spectral density) (Kleinbauer 2004). To upgrade the estimated state, the KF uses measurements that are related to the state but disturbed as well (Simon and Chia 2002).

Overall, the KF consists of two stages, prediction and correction. In the first stage, the state of the system is predicted with the dynamic model. In the second stage, the state of the system is corrected with the observation model. Thereby, it minimizes the error covariance of the estimator. The KF is called as a recursive filter; because the procedure of the system is repeated for each time step with the state of the previous time step as initial value (Fig. 4).
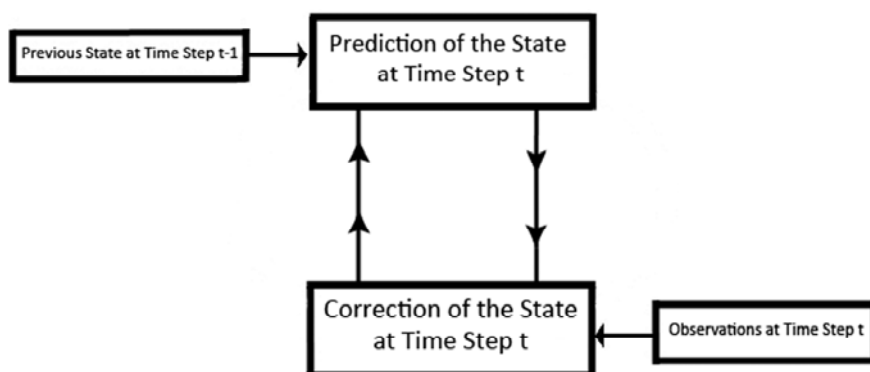

Fig. 4. Circuit of the Kalman filter. This procedure is repeated for each time step with the state of the previous time step as initial value. Therefore, the KF is called a recursive filter (Kleinbauer 2004).

The state vector includes the variables of interest and describes the state of the dynamic system and also suggests its degree of freedom (Brabec and Jílek 2007). The variables in the state vectors can be inferred from values that are measurable (Kleinbauer 2004). In other words, they cannot be measured directly. State vectors contain two different values at the same time. These are the predicted value before the update (*a priori* value) and the corrected value after the update (*a posteriori* value) (Kleinbauer 2004).

The time update equations are responsible for projecting forward (in time) using the current state and error covariance estimates to obtain the *a priori* estimates for the next time step (Kleinbauer 2004).

An important point regarding the state vector is its transformation over time, which is known as a dynamic model. Another term that should be described here is "observation model". This model illustrates the significant relation between the state and the measurement. In scientific terms, the KF estimates the state of a linear system. In order to design and implement a KF, an estimate of the process variables is required. Furthermore, in order to design a KF to eliminate noise from a signal, the process of measuring must be describable as a linear phenomenon.

3.5 KF model for tracking applications

According to above descriptions, monitoring of radon during diffusion/migration is an exact example of random walk. Each monitored sample consists of two parts; a real value that indicates the released radon from the earth plus a white noise which is produced due to the effect of environmental parameter on radon concentration and/or measurement process.

According to several publications, the original predict-update KF equations for tracking applications of a linear system, as assumed for our case, are as below (Cichocki and Amari 2002, Kleinbauer 2004, Leśniak *et al.* 2009):

Predict

$$\hat{x}_{t|t-1} = F_t \hat{x}_{t-1|t-1} + B_t u_t \ , \tag{9}$$

$$P_{t|t-1} = F_t P_{t-1|t-1} F_t^T + Q_t \ . \tag{10}$$

Update

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t \left(y_t - H_t \hat{x}_{t|t-1} \right) , \tag{11}$$

$$K_t = P_{t|t-1} H_t^T \left(H_t P_{t|t-1} H_t^T + R_t \right)^{-1} , \tag{12}$$

$$P_{t|t} = \left(I - K_t H_t \right) P_{t|t-1} , \tag{13}$$

where $\hat{x}$ is the estimated state, F the state transition matrix, u the control variables, B the control matrix, P the state variance matrix (*i.e.*, error of estimation), Q the process variance matrix (*i.e.*, error due to process), y the measurement variables, H the measurement matrix, K the Kalman gain, R the measurement variance matrix (*i.e.*, error from measurement), and the subscripts are as follows: ($t|t$) the current time period, ($t–1|t–1$) the previous time period, and ($t|t–1$) the intermediate steps.

All variants are designated according to their appearance: Normal (a) denotes scalars, and bold-italic (***a***) denotes vectors. According to the above explanations, the KF eliminates noise by assuming a pre-defined model for a system. Obviously, such a model must be reliable. In order to satisfy the reliability criteria for the model, the following stages are considered during development of any such model (so considered in this study):

- ❑ model the state process,
- ❑ model the measurement process,
- ❑ model the noise (this needs to be done for both the state and measurement process),
- ❑ test the filter,
- ❑ refine the filter.

4. PROPOSED KF MODEL FOR RADON CONCENTRATION MONITORING

In this Section, we are trying to estimate the level of radon released from the earth during the experiment. The measurements obtained are from the RAD7-device outputs.

As explained, monitoring of radon during diffusion/migration is an exact example of random walk. Each monitored sample consists of two parts: a real value that indicates the released radon from the earth, plus a white noise which is produced due to effect of environmental parameters on radon concentration and/or measurement process. On the other hand, as explained in the following section, the data recorded in this study showed Gaussian distribution, facilitating the BSS analysis.

Considering the above, a KF model is developed here to estimate the level of radon gas released during the experiments. Generally, the concentration of radon in the earth could be:

- ❑ increasing, decreasing or static (*i.e.*, the level of radon concentration could be varied prior to an earthquake);

- ❑ sloshing or stagnant (*i.e.*, the relative level of radon to the average level is changing over time, or is static).

The simplified scheme of the modeled system is given in Fig. 5.
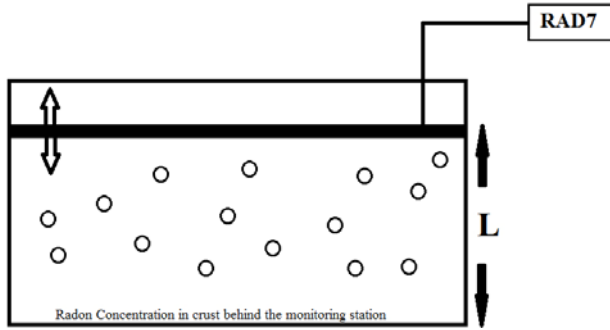
Fig. 5. A schematic plan of radon level behind the monitoring station. The level of radon can be varied (increased, decreased or static) by different phenomena, especially before earthquakes.

As per any modeling or simulation, some assumptions are required to be made. Such assumptions should not broadly affect the reliability or the real-life situation of the process. On the other hand, no investigations on using the KF model for radon gas monitoring have been reported in the literature. So, a simple yet reliable benchmark KF model for radon gas monitoring is developed. The following assumptions and formulizations are considered:

❑ the level of released radon is considered to be constant (*i.e.*, $L = C$);
❑ the predict-update equations would then be reduced to scalar (*i.e.*, $\hat{x} = x$ where x is the estimate of L);
❑ also, for the constant model, $x_{t+1} = x_t$, so $F_t = 1$ for any $t \geq 0$;
❑ also in this case, control variables, B and u, are not used;
❑ also, the level of radon concentration is represented by $y = $ y.
❑ the scale of measurement (z) and state estimate (x) are the same; therefore $H = 1$;
❑ noise is assumed to be from the measurement, so $R = r$;
❑ the process is scalar; therefore $P = p$.
❑ the noise is also considered as $Q = q$.

Considering the above, the predict-update equations can be rewritten as:
Predict:

$$x_{(t|t-1)} = x_{(t-1|t-1)} \, , \tag{14}$$

$$P_{(t|t-1)} = P_{(t-1|t-1)} + q_{(t)} \, . \tag{15}$$

Update:

$$K_t = \frac{P_{(t|t-1)}}{P_{(t|t-1)} + r_{(t)}} \, , \tag{16}$$

$$x_{(t|t)} = x_{(t|t-1)} + K_{(t)} * \left(y_{(t)} - x_{(t|t-1)} \right) , \tag{17}$$

$$P_{(t|t)} = \left(1 - K_{(t)} \right) * P_{(t|t-1)} . \tag{18}$$

The KF filter is now completely modeled. Initially, the state progress is considered to be an arbitrary number, with an extremely high variance as it is completely unknown: $x_0 = 0$ and $P_0 = 1000$. It is worth noting that the
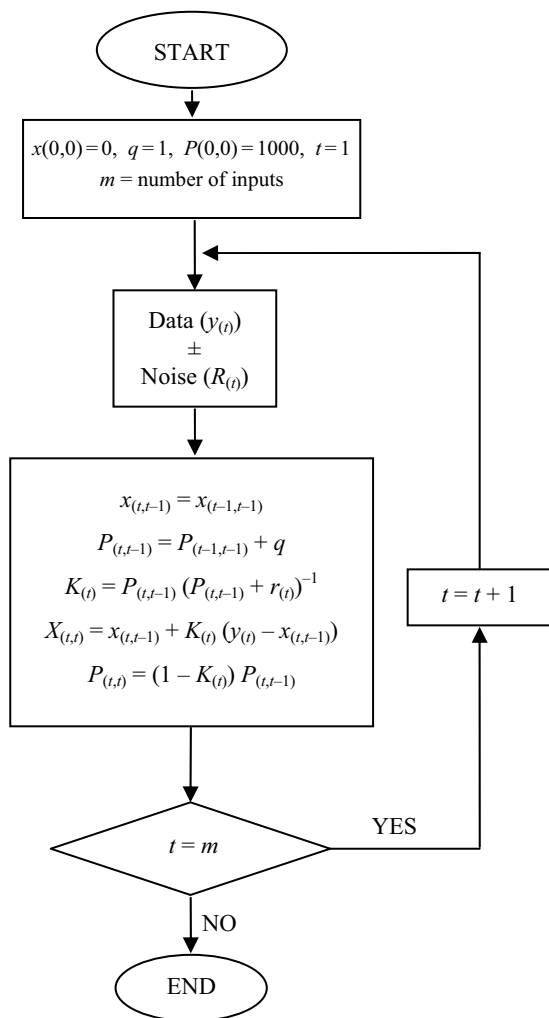

Fig. 6. The algorithm of proposed Kalman filter. Based on the state equation, the data has been read and filtered in each step and the next value has been read as initial value.

more realistic the variable is the faster the convergence of the model would be. The system noise is assumed to be $q = 1$, because the systematic error of the measurement device (RAD7) is exactly 1 Bq/m^3 (according to RAD7 instruction catalogue). The proposed filter could then be illustrated schematically as Fig. 6.

5. RESULTS AND DISCUSSION

During the monitoring of radon concentration in Jooshan hot-spring (December 2011 to April 2012) more than 35 seismic events with different
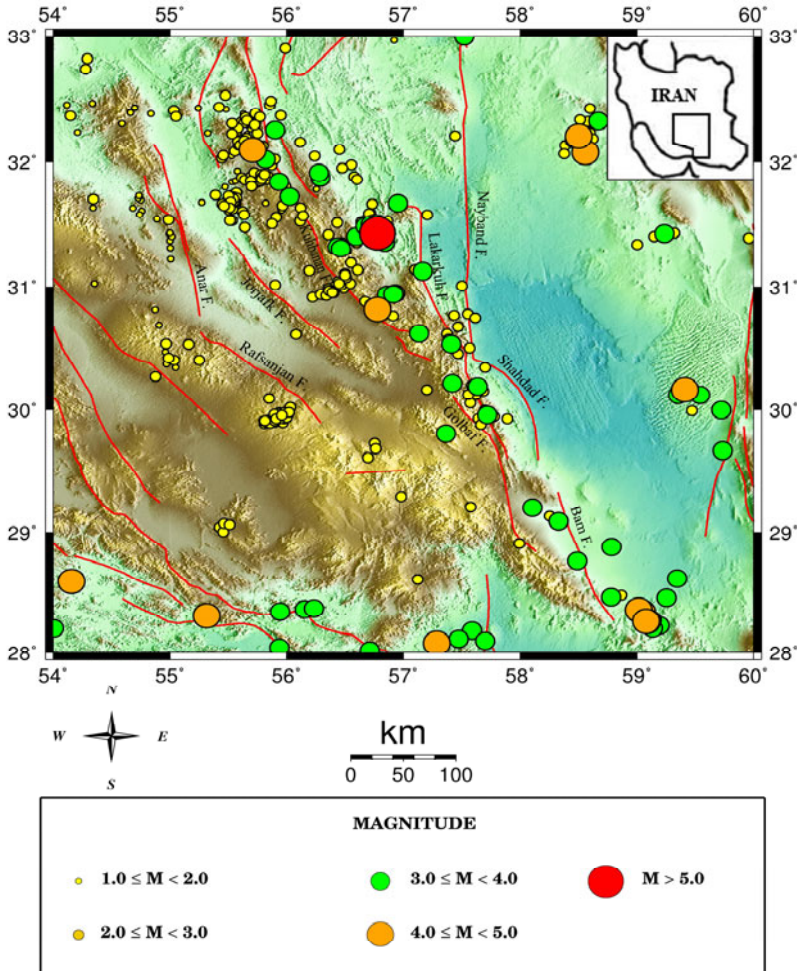

Fig. 7. Seismic activity during the study period (from December 2011 until April 2012) in study area. During this period, more than 35 events with different magnitudes occurred in the 100 km radius around the station.

magnitudes were recorded in a radius span of 100 km around the station. Figure 7 indicates the seismicity of regions around Golbaf and Shahdad faults during the study period.

In this study, the effective precursory (EP) ratio is introduced and applied as:

$$EP = D / R \; , \tag{19}$$

where D [km] is calculated from Eq. 1 and R [km] is the distance from the epicenters of each earthquake to the monitoring station. The calculated value of EP indicates the ability of each seismic event to be used as a precursor. The higher EP means the larger magnitude earthquake and closer epicenters to the monitoring station. Summary of all seismic events with $M \geq 2.5$ during the study and their effective precursory ratio is given in Table 1. All the seismic data and their relevant parameters (*e.g.*, epicenter, magnitude, and depth) are taken from IRSC and IIEES catalogues. In order to narrow down the data to a handful and reliable number of inputs, the events with $EP \geq 0.4$ were used for the KF model analysis.

Table 1

The parameters of reported earthquakes
in vicinity of the monitoring station during the study period*

Date	Local time	Lat.	Long.	Depth [km]	Magni-tude	D	R [km]	EP
29 Dec 2011	20:40	29.9	57.8	10.1	2.5	11.8	35.2	0.33
05 Jan 2012	23:00	30.5	57.5	5.0	2.9	17.6	39.4	0.44
06 Jan 2012	03:41	29.9	57.7	6.1	3.0	19.4	23.1	0.83
09 Jan 2012	08:59	30.1	57.6	7.9	3.3	26.2	3.9	6.71
09 Jan 2012	12:56	29.9	57.6	10.2	2.7	14.4	18.2	0.79
09 Jan 2012	14:08	29.9	57.7	10.1	2.6	13.1	26.8	0.48
13 Jan 2012	05:06	29.2	58.1	13.5	3.7	38.9	115.1	0.33
17 Jan 2012	02:06	30.1	57.2	12.2	2.8	15.9	41.8	0.38
28 Jan 2012	13:32	30.2	57.4	6.1	3.2	23.7	21.9	1.08
05 Feb 2012	11:55	30.6	57.1	10.0	3.1	21.5	70.7	0.30
10 Feb 2012	14:11	30.1	57.6	6.2	2.7	14.4	4.1	3.51
15 Feb 2012	18:57	32.0	58.5	5.4	4.3	70.6	230.9	0.30
17 Feb 2012	04:26	29.8	57.6	15.5	2.5	11.8	30.7	0.38
21 Feb 2012	08:58	29.8	57.3	6.1	3.2	23.7	46.7	0.50
23 Feb 2012	07:59	30.4	57.4	5.1	2.9	17.6	37.5	0.46
24 Feb 2012	04:51	30.0	57.5	10.2	2.7	14.4	12.6	1.14
27 Feb 2012	18:48	31.4	56.7	10.4	5.4	209.8	163.7	1.28

*)The source of seismic data: IRSC (with permission).

5.1 Noise reduction

Statistical distribution of the monitored level of radon gas concentration and its Gaussian fit are given in Fig. 2. As reported in statistical papers for such distributions, BSS can be solved by the independent component analysis technique (Evensen 2003, Kownacki 2011). In different studies, it has been reported that the deviations exceeding ±1, ±1.5, and $\pm2\,\sigma$ from the average concentration level are considered as anomalies, and thereby potentially linked to the geodynamics of the area. Zmazek *et al*. (2005) proposed the $\pm\,\sigma$ threshold approach for anomaly descriptions. Summary of the thresholds and the associated type of anomalies of this approach are given in Table 2. According to this approach, in $\pm1\,\sigma$ group, 57% of anomalies is correct and related to seismic events (CA), 38% of them appeared without seismic events (FA), and finally in 5% cases no anomaly is observed for an earthquake (NA) (Zmazek *et al*. 2005). On the other hand, in the $\pm2\,\sigma$ group, 50, 27, and 22% of anomalies are CA, FA, and NA, respectively (Zmazek *et al*. 2005). Thereby, for this study, the $\pm1\,\sigma$ threshold is selected to investigate the relation of the filtered data with geodynamics events.

Table 2

Different thresholds and the associated anomaly
in $\pm\sigma$ approach (Zmazek *et al*. 2005)

Anomaly	$\pm1.0\,\sigma$	$\pm1.5\,\sigma$	$\pm2.0\,\sigma$
CA	12	10	9
FA	8	8	5
NA	1	3	4

Note: CA refers to radon concentration anomalies correct and related to seismic activity, NA refers to those anomalies caused by sources other than seismic activity, and FA refers to non-active period.

A commercial version of MATLAB® is used to evaluate and analyze the monitored results alongside the noise reduction data (incorporating the developed KF filter). A summary of the data analysis is given in Fig. 8. A comparison of monitored data before and after filtering is given in Fig. 8A. The differences between these two curves (before and after noise reduction) are proportional to the measurement system (RAD7) noise (q). It can be seen that the filtered signal is much smoother than the monitored data. This can be explained, as in the spectral analysis it is reported that the KF model decreases the high frequency components of the recorded signal. It also shows the reliability of the filtered data for further investigations. The magnitudes of earthquakes and the EP ratio of the seismic activities are pro-
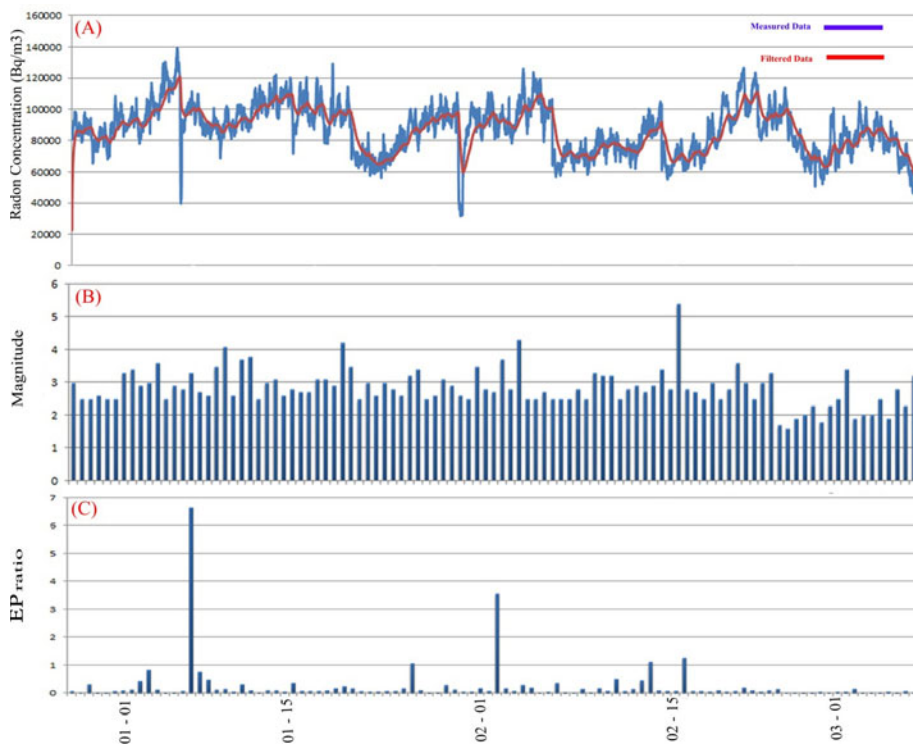
Fig. 8: (A) Measured radon signal and filtered radon signal, (B) magnitude of earth-quakes that occurred during the measurement period, and (C) effective precursory (EP) ratio for each earthquake.

vided in Fig. 8B and C, respectively; all are in the same time span (together with Fig. 8A, covering the whole period of study. From Fig. 8, it is clear that there is good correlation between sharp anomalies and earthquake occur-rence. As can be observed, in a few hours to a few days prior to an earth-quake, a variation in radon level is visible. Figure 9 illustrates the variation of radon gas concentration during the monitoring period, after applying the KF filter. As can be seen, the overall number of sudden changes (A-V) was 22. These are investigated in further details to realize their potential relation-ship to the reported earthquakes. In this time series signal, the average and standard deviation of data are 82214.18 and 16776.92, respectively. So, the values of $\bar{x}+\sigma$, $\bar{x}-\sigma$, $\bar{x}+2\sigma$, and $\bar{x}-2\sigma$ are inferable, which are 98991.10, 65437.26, 115768.02, and 48660.34, respectively. Point A is not a real peak in radon concentration, but it is due to the operation of KF. The KF initialized the process of filtering with an arbitrary number with an ex-
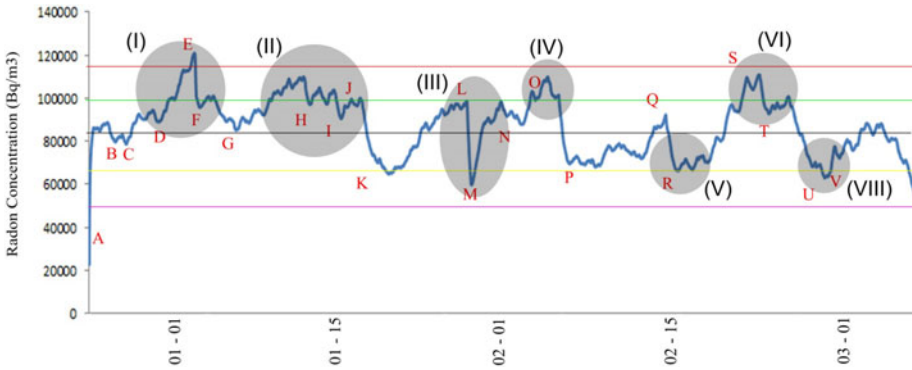
Fig. 9. Filtered signal whose all peaks have been labeled. The values of average (black), $\bar{x} + \sigma$ (green), $\bar{x} - \sigma$ (yellow), $\bar{x} + 2\sigma$ (red), and $\bar{x} - 2\sigma$ (violet) have been drawn. The sharp increase in the initial part of signal (point A) is due to the operation of KF.

tremely high variance, as it is completely unknown. The value of variance controls the velocity of convergence. It means that initializing with a more meaningful variable results in faster convergence. Among the A-V labeled peaks, only a few provide good correlation with the reported seismic events. These are categorized in 7 different ranges and labeled as (I-VIII) in Fig. 9. For instance, 53 hours after the recorded peak F (categorized in the range I), an earthquake with M3.3 and an EP value of 6.71 occurred. By comparison to the unfiltered data (Fig. 8A) this sudden variation in radon concentration is also obvious. This event was about 4 kilometers away from the monitoring station, with radon activity crossing the σ level above the average value. Point M (range III) in Fig. 9 shows the bottom of a sharp decrease. This happened between sample numbers 3174 and 3477 and prolonged for 50 hours. During this period, the level of radon concentration decreased from 96 kBq/m^3 down to 60 kBq/m^3 (minimum point) and then started to increase. About 92 hours after this peak, an earthquake M3.2 with EP = 1.08 occurred. The location of the earthquake was about 22 kilometers away from the Jooshan hot-spring. Peak P (Fig. 9 – just after range IV), occurred after a sharp decrease in level of radon concentration monitored in sample 4048. This decrease started from 3929 sample until 4048 samples and then started increasing toward the normal value. An earthquake M2.7 with EP = 3.51 occurred about 4 kilometers away from the station, 105 hours after this peak. During this fluctuation, the level of radon declined from 101 to 71 kBq/m^3 at the minimum state and then started to increase. According to the above description of data fluctuations, an anomaly is introduced as "radon concentration spike crossing":

$$RS = \bar{x} \pm \sigma \ , \tag{20}$$

where $\bar{x}$ is the average value of measured values, and σ is the standard deviation. As described, all earthquakes happened after the $\bar{x} \pm \sigma$ threshold line in filtered data and these points illustrate anomalous behavior in radon signal. Among the ranges in Fig. 9, there are few ranges in which the radon concentration behaved anomaly but no earthquake occurred after these behaviors. This suggests the uncertainties involved with the Dobrovolsky equation and its need to be modified.

5.2 Pattern recognition

Another important result of this study is recognizing the pattern of relation between filtered radon concentration signal and the EP ratio. As explained in Fig. 8A, the filtered data is smoother than the measured data. Thereby, the rate of variation in filtered radon concentration is introduced to be changed by the angle of the decreasing line, which we denote as θ. As illustrated in Fig. 10, between points E to F, $\theta = 90°$. After this stage, an event M3.3 with EP = 6.71 occurred. Also, between points J to K, L to M, O to P, and Q to R, the decreasing angles are 79°, 88°, 84°, and 81°, respectively. After these anomalies, seismic events occurred with M3.2, M2.7, M2.7, and M5.4, respectively. Therefore, θ could be described as the rate of variation in radon level. In other words, the higher slope of radon signal after noise cancelation indicates that the fluctuation is sharper, and *vice versa*. Different studies, *e.g.*, Wakita *et al.* (1991), Kuo *et al.* (2006), and Tsunomori and Kuo (2010) suggest that the level of radon concentrated in soil and groundwater decreases from background levels to a lower level prior an earthquake and then
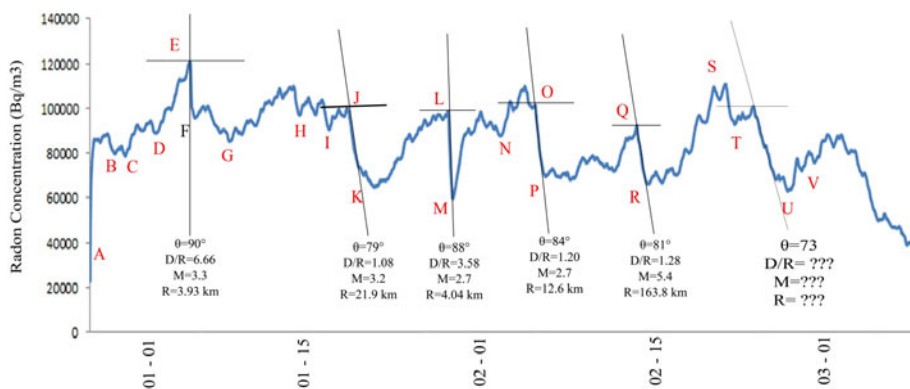

Fig. 10. Decreasing angles of radon signal (θ). Different values of θ could describe the rate of variation in radon level. In this study, θ has been in good correlated with EP ratio.
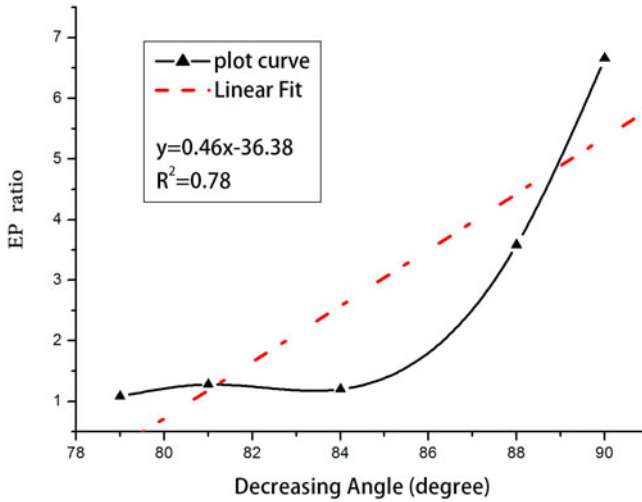
Fig. 11. The decline angles before earthquake (θ) plotted as a function of *EP* ratio.

starts to increase before reaching its previous normal value. Based on this, in order to describe the correlation between θ and EP, values of decreasing angles before earthquakes were plotted as a function of EP ratio (Fig. 11). Therefore, the equation of the best straight fit-line of the figure would be

$$EP = \alpha + \beta\theta \; , \tag{21}$$

where EP is the effective precursory ratio (Eq. 19), θ is the decreasing angle (degree) on the filtered radon concentration data, β is the slope of straight fit-line, and α is determined by points values. Also the correlation coefficient r^2 was determined to be 0.78, indicating that θ and EP ratio were correlated to each other. Based on previous descriptions and the proposed pattern, it is concluded here that the probability of an earthquake occurrence increases proportional to the increased value of θ.

6.   CONCLUSION

A continuous radon concentration monitoring analysis was performed at the widely reported earthquake-prone area of Jooshan hot-spring (located near Golbaf and Shahdad faults), Kerman province, south east Iran. Analytical analysis using Kalman filter modeling on noise reduction from the radon concentration signal was reported for the first time. The main scientific conclusions of the study are as summarized here:

❏ During the measurement, parameters of different kind result in noises on signals. So, before interpreting the data, the effects of these noises must be reduced as much as possible.

□ KF is a powerful tool which is successfully reported to decrease the high-frequency components of the recorded signal of the monitored radon gas concentration.

□ A novel threshold on the filtered radon concentration data is defined as "radon concentration spike crossing" beyond which a seismic event seems inevitable.

□ Occasionally, anomalies were observed on the radon concentration figure but no seismic events were reported. This suggests uncertainties involved with the Dobrovolsky equation.

□ A new ratio between the effective radius (as proposed by Dobrovolsky) and the distance of the earthquake from the monitoring station was defined as "effective precursory" ratio. This proved a much better understanding of the mechanism of the radon concentration signal.

□ The use of KF filter also facilitated much detailed analytical analysis of data. From the filtered data, a decreasing angle was reported to have proportional effect on the possibility of forthcoming earthquake.

□ Effective precursory (EP) ratio and the decreasing angles were found to be remarkably correlated.

□ Generally, the level of radon concentration was found to be a useful measure as a precursor, few hours to few days prior to an earthquake.

□ The delay time between the anomalies and seismic event depends on different parameters such as geological features of case study, environmental parameters, *etc*. Therefore, in order to increase the reliability of the findings on precursory role of radon concentration level, it can be useful to increase the number of stations and duration of measurement in future analyzes.

R e f e r e n c e s

Ali Yalım, H., A. Sandıkcıoğlu, O. Ertuğrul, and A. Yıldız (2012), Determination of the relationship between radon anomalies and earthquakes in well waters on the Akşehir-Simav Fault System in Afyonkarahisar province, Turkey, *J. Environ. Radioactiv.* **110**, 7-12, DOI: 10.1016/j.jenvrad.2012.01.015.

Belouchrani, A., K. Abed-Meraim, J.-F. Cardoso, and E. Moulines (1997), A blind source separation technique using second-order statistics, *IEEE Trans. Signal Process.* **45**, 2, 434-444, DOI: 10.1109/78.554307.

Berberian, M., J.A. Jackson, M. Ghorashi, and M.H. Kadjar (1984), Field and teleseismic observations of the 1981 Golbaf–Sirch earthquakes in SE Iran, *Geophys. J. Roy. Astr. Soc.* **77**, 3, 809-838, DOI: 10.1111/j.1365-246X.1984.tb02223.x.

Berberian, M., J.A. Jackson, E. Fielding, B.E. Parsons, K. Priestley, M. Qorashi, M. Talebian, R. Walker, T.J. Wright, and C. Baker (2001), The 1998 March 14 Fandoqa earthquake (M_w 6.6) in Kerman province, southeast Iran: Re-rupture of the 1981 Sirch earthquake fault, triggering of slip on adjacent thrusts and the active tectonics of the Gowk fault zone, *Geophys. J. Int.* **146**, 2, 371-398, DOI: 10.1046/j.1365-246x.2001.01459.x.

Brabec, M., and K. Jílek (2007), State-space dynamic model for estimation of radon entry rate, based on Kalman filtering, *J. Environ. Radioactiv.* **98**, 3, 285-297, DOI: 10.1016/j.jenvrad.2007.05.006.

Brailean, J.C., R.P. Kleihorst, S. Efstratiadis, A.K. Katsaggelos, and R.L. Lagendijk (1995), Noise reduction filters for dynamic image sequences: a review, *Proc. IEEE* **83**, 9, 1272-1292, DOI: 10.1109/5.406412.

Cichocki, A., and S.-I. Amari (2002), *Adaptive Blind Signal and Image Processing. Learning Algorithms and Applications*, John Wiley & Sons, Chichester.

Dobrovolsky, I.P., S.I. Zubkov, and V.I. Miachkin (1979), Estimation of the size of earthquake preparation zones, *Pure Appl. Geophys.* **117**, 5, 1025-1044, DOI: 10.1007/BF00876083.

Evensen, G. (2003), The Ensemble Kalman Filter: theoretical formulation and practical implementation, *Ocean Dynam.* **53**, 4, 343-367, DOI: 10.1007/s10236-003-0036-9.

Finkelstein, M., S. Brenner, L. Eppelbaum, and E. Ne'Eman (1998), Identification of anomalous radon concentrations due to geodynamic processes by elimination of Rn variations caused by other factors, *Geophys. J. Int.* **133**, 2, 407-412, DOI: 10.1046/j.1365-246X.1998.00502.x.

Fujimoto, M., and Y.Ariki (2000), Noisy speech recognition using noise reduction method based on Kalman filter. **In:** *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, ICASSP'2000, 5-9 June 2000, Istanbul, Turkey*, Vol. 3, 1727-1730, DOI: 10.1109/icassp. 2000.862085.

Hashemi, S.M., A. Negarestani, M. Namvaran, and S.M. Musavi Nasab (2013), An analytical algorithm for designing radon monitoring network to predict the

location and magnitude of earthquakes, *J. Radioanal. Nucl. Chem*. **295**, 3, 2249-2262, DOI: 10.1007/s10967-012-2310-0.

Hauksson, E. (1981), Radon content of groundwater as an earthquake precursor: Evaluation of worldwide data and physical basis, *J. Geophys. Res.* **86**, B10, 9397-9410, DOI: 10.1029/JB086iB10p09397.

Hyvärinen, A., and E. Oja (2000), Independent component analysis: algorithms and applications, *Neural Networks* **13**, 4-5, 411-430, DOI: 10.1016/S0893-6080 (00)00026-5.

Kleinbauer, R. (2004), Kalman filtering implementation with Matlab, Study Report, Universität Stuttgart, Institute of Geodesy.

Kownacki, C. (2011), Optimization approach to adapt Kalman filters for the real-time application of accelerometer and gyroscope signals' filtering, *Digit. Signal Process.* **21**, 1, 131-140, DOI: 10.1016/j.dsp.2010.09.001.

Kuo, T., K. Fan, H. Kuochen, Y. Han, H. Chu, and Y. Lee (2006), Anomalous decrease in groundwater radon before the Taiwan M6.8 Chengkung earthquake, *J. Environ. Radioactiv.* **88**, 1, 101-106, DOI: 10.1016/ j.jenvrad.2006.01.005.

Lee, T.-W., M.S. Lewicki, M. Girolami, and T.J. Sejnowski (1999), Blind source separation of more sources than mixtures using overcomplete representations, *IEEE Signal Process. Lett.* **6**, 4, 87-90, DOI: 10.1109/ 97.752062.

Leśniak, A., T. Danek, and M. Wojdyła (2009), Application of Kalman filter to noise reduction in multichannel data, *Schedae Informaticae* **17**, 18, 63-73, DOI: 10.2478/v10149-010-0004-3.

Negarestani, A., S. Setayeshi, M. Ghannadi-Maragheh, and B. Akashe (2002), Layered neural networks based analysis of radon concentration and environmental parameters in earthquake prediction, *J. Environ. Radioactiv*. **62**, 3, 225-233, DOI: 10.1016/S0265-931X(01)00165-5.

Ramola, R.C. (2010), Relation between spring water radon anomalies and seismic activity in Garhwal Himalaya, *Acta Geophys*. **58**, 5, 814-827, DOI: 10.2478/s11600-009-0047-0.

Richon, P., F. Perrier, J.-C. Sabroux, M. Trique, C. Ferry, V. Voisin, and E. Pili (2004), Spatial and time variations of radon-222 concentration in the atmosphere of a dead-end horizontal tunnel, *J. Environ. Radioactiv.* **78**, 2, 179-198, DOI: 10.1016/j.jenvrad.2004.05.001.

Roesser, R.P. (1975), A discrete state-space model for linear image processing, *IEEE Trans. Automat. Contr.* **20**, 1, 1-10, DOI: 10.1109/tac.1975.1100844.

Seyis, C., S. İnan, and T. Streil (2010), Ground and indoor radon measurements in a geothermal area, *Acta Geophys.* **58**, 5, 939-946, DOI: 10.2478/s11600-010-0012-y.

Simon, D. (2001), Kalman filtering, *Embedded Sys. Program*. **14**, 6, 72-79.

Simon, D., and T.L. Chia (2002), Kalman filtering with state equality constraints, *IEEE Trans. Aero. Elec. Sys*. **38**, 1, 128-136, DOI: 10.1109/7.993234.

Spitzer, F. (2001), *Principles of Random Walk*, 2nd ed., Graduate Texts in Mathematics, Vol. 34, Springer, New York.

Tsunomori, F., and T. Kuo (2010), A mechanism for radon decline prior to the 1978 Izu-Oshima-Kinkai earthquake in Japan, *Radiat. Meas.* **45**, 1, 139-142, DOI: 10.1016/j.radmeas.2009.08.003.

Wakita, H., G. Igarashi, and K. Notsu (1991), An anomalous radon decrease in groundwater prior to an M6.0 earthquake: A possible precursor?, *Geophys. Res. Lett.* **18**, 4, 629-632, DOI: 10.1029/91GL00824.

Zhang, K., and A. Hyvärinen (2011), A general linear non-Gaussian state-space model: Identifiability, identification, and applications. **In:** C.-N. Hsu and W.S. Lee (eds.), *JMLR Workshop and Conference Proc., Asian Conf. on Machine Learning 2011, Tokyo, Japan*, 113-128.

Zibulevsky, M., and B.A. Pearlmutter (2001), Blind source separation by sparse decomposition in a signal dictionary, *Neural Comput.* **13**, 4, 863-882, DOI: 10.1162/089976601300014385.

Zmazek, B., M. Živčić, L. Todorovski, S. Džeroski, J. Vaupotič, and I. Kobal (2005), Radon in soil gas: How to identify anomalies caused by earthquakes, *Appl. Geochem.* **20** 6, 1106-1119, DOI: 10.1016/j.apgeochem.2005.01.014.

Passive Seismic Experiment "13 BB Star"
in the Margin of the East European Craton,
Northern Poland

Marek GRAD[1], Marcin POLKOWSKI[1], Monika WILDE-PIÓRKO[1],
Jerzy SUCHCICKI[2], and Tadeusz ARANT[2]

[1]Institute of Geophysics, Faculty of Physics, University of Warsaw, Warsaw, Poland
e-mail: mgrad@mimuw.edu.pl (corresponding author)

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

A b s t r a c t

The lithosphere-asthenosphere boundary (LAB) is investigated re-
cently very effectively, mostly using seismic methods because of their
deep penetration and relatively good resolution. The nature of LAB is
still debated, particularly under "cold" Precambrian shields and plat-
forms. Passive experiment "13 BB star" is dedicated to study deep struc-
ture of the Earth's interior in the marginal zone of the East European
craton in northern Poland. The seismic network consists of 13 broadband
stations on the area of *ca.* 120 km in diameter. The network is located in
the area of well-known sedimentary cover and crustal structure. Good re-
cords obtained till now, and expected during next 1-year long recording
campaign, should yield images of detailed structure of the LAB, „410",
"?520", and „660" km discontinuities, as well as mantle-core boundary
and inner core.

Key words: East European craton, lithosphere, asthenosphere, seismic
passive experiment, broadband seismology.

---

1. INTRODUCTION

The nature of a transition from a rigid plate − the lithosphere, to a weaker layer below, the asthenosphere, is still debated. Till now our understanding of the lithosphere-asthenosphere boundary (LAB) is limited in comparison to other interfaces in the Earth, *e.g.*, Moho, "410 km", "670 km", and mantle-core boundary. The thickest lithosphere, in excess of 200 km, is observed under "cold" Precambrian shields and platforms, while the thinnest lithosphere, of 50-100 km, was found under "hot" oceans and oceanic and continental rifts. The depth of oceanic LAB is widely accepted, why continental LAB is very often difficult to detect. For example, lithospheric thickness of the Baltic Shield has estimates varying between 160 and 350 km. Another controversy relates to nature of LAB. Recent estimates and modellings suggest LAB to be a first-order structural discontinuity that accommodates differential motion between tectonic plates and the underlying mantle (*e.g.*, Eaton *et al*. 2009). According to other opinions, LAB is not a sharp discontinuity, but rather gradual and wide transition zone (see, *e.g.*, Meissner 1986, Bartzsch *et al.* 2011).

The lithosphere-asthenosphere boundary (LAB) is investigated recently very effectively, mostly using seismic methods because of their deep penetration and relatively good resolution. Comparison of *S*-wave velocity changes with depth in the upper mantle obtained from global tomographic models shows that velocities under cratons are greater than under oceans (Fig. 1a). This figure shows also that LAB is rather gradual and wide transition zone, and velocity-depth curve illustrates the difficulty in identifying a signature of the lithosphere-asthenosphere boundary (Romanowicz 2009). In general, velocities under cratons are greater than under oceans down to 200 km depth, and velocity decreases with depth, with a minimum centred at 100 to 150 km under the oceans and 200 to 250 km in the cratons. Regional 1D *S*-wave velocity models obtained for the central part of Baltic Shield from SVEKALAPKO array shows that velocities beneath the craton are significantly higher (about 4% down to a depth of 250 km) compared to standard model *iasp91* (Bruneton *et al.* 2004). On the other hand, lowering of the velocity with depth is not visible (Fig. 1b; Sandoval *et al.* 2004, Kozlovskaya *et al.* 2008). This fact could be interpreted as a lack of asthenosphere, or that LAB in this area occurs deeper. Shear wave velocity in southern Tibet shows clearly lowering velocity in the depth range of 120-300 km, which could be interpreted as the asthenosphere (Fig. 1c; Jiang *et al.* 2011).

Passive seismic experiment "13 BB star" is dedicated to study LAB and deep structure of the Earth's interior in the marginal zone of the East European craton in northern Poland. The Precambrian East European craton (EEC) is the coherent part of Europe that occupies the north-eastern half of
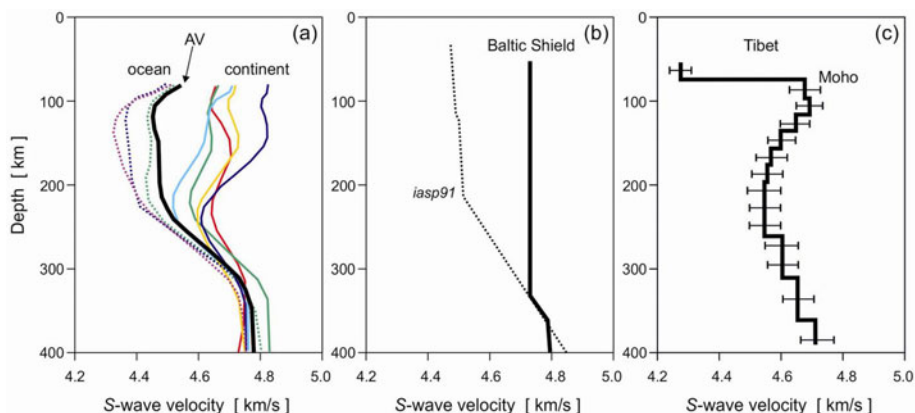
Fig. 1. Comparison of shear-wave slowdown for different regions of the Earth: (a) Shear-wave velocity changes with depth in the upper mantle from global tomographic model. Dashed lines are for oceanic mantle (pink for Pacific, blue for mid-Atlantic ridge, and green for Indian Ocean, south of Madagascar) and solid lines are for continental mantle under cratons (light blue for northern area of South America, green for NW part of Africa, red for Western Siberia, orange for Central Australia, and deep blue for NE part of North America). Thick black line (AV) is global average. Velocities under cratons are greater than under oceans down to 200 km depth, and velocity decreases with depth, with a minimum centred at 100 to 150 km under the oceans and 200 to 250 km in the cratons; (b) 1D *S*-wave velocity model obtained for the central part of Baltic Shield at the station FJ01 from SVEKALAPKO array (solid line) compared to standard model *iasp91* (thin dashed line; Kennett and Engdahl 1991); (c) 1D average shear wave velocity in southern Tibet (solid line with error bars). Compilation from Romanowicz (2009), Jiang *et al.* (2011), and Kozlovskaya *et al.* (2008).

the continent. In the south and west of EEC, the Trans-European suture zone (TESZ) separates the Precambrian terranes of the craton from the younger Phanerozoic terranes of middle to late Palaeozoic accretion and deformation (Fig. 2; Znosko 1975, 1979, Pożaryski *et al.* 1982, Ziegler 1990, Pharaoh *et al.* 1997, Berthelsen 1998, Grad *et al*. 2002, Bogdanova *et al.* 2006). The TESZ correlates with a gravity minimum (*e.g.*, Królikowski and Petecki 1997, Krysiński *et al.* 2000, Grabowska *et al.* 2011), high heat flow values (*e.g.*, Majorowicz *et al.* 2003), and electromagnetic properties of the lithosphere (*e.g.*, Jóźwiak 2013).

2. "13 BB STAR" EXPERIMENT

To study the deep structure of the Earth's interior, location of "13 BB star" network was chosen in northern Poland, in the marginal zone of EEC (Figs. 2 and 3). This area is well recognised in terms of both the velocities of
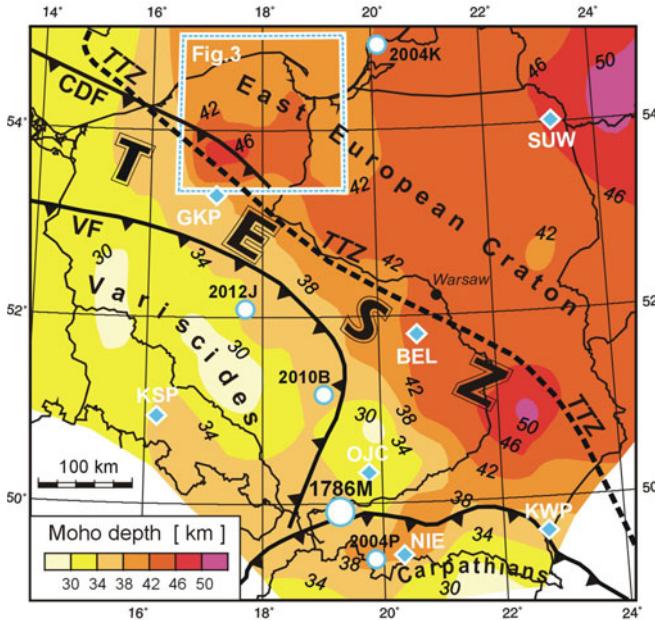
Fig. 2. Location of seismic stations in Poland on the background of the Moho depth map (Guterch and Grad 2006) with main tectonic elements of the area: CDF – Caledonian Deformation Front, TESZ – Trans-European Suture Zone, TTZ – Teisseyre-Tornquist Zone, VF – Variscan Front. Seven permanent broadband seismic stations are marked by blue diamonds (SUW, GKP, BEL, KSP, OJC, NIE, KWP). White dots show location of five strongest seismic events which occurred during the past decade in the study area: 2004K (two events near Kaliningrad, 21 September 2004, M_W = 5.1 and 5.2; Domański 2007), 2012J (Jarocin, 6 January 2012, M_L = 3.8; Lizurek *et al.* 2013); 2010B (Bełchatów, 22 January 2010, Mw = 4.3; Wiejacz and Rudziński 2010), 2004P (Podhale, 30 November 2004, M_b = 4.7; Wiejacz and Dębski 2009). Big dot shows location of largest earthquake known to have ever taken place in Poland 1786M (Myślenice, 3 December 1786, M = 5.6; Pagaczewski 1972). White-blue rectangular shows location of "13 BB star" passive experiment in northern Poland, in the edge of the East European craton (for details see Fig. 3).

sedimentary layer and the crustal structure (Skorupa 1974, Grad *et al.* 2003, Grad and Polkowski 2012). Precise knowledge of velocities and thickness of sediments is based on borehole data (VSP – vertical seismic profiling). The mean velocity of the sedimentary cover increases with thickness and varies from 2.5 km/s where these strata are 1 km thick to about 4.3 km/s where they are 8 km thick (Grad 1986). In northern Poland, the depth of the crystalline basement is 1-2 km and increases toward the south-west to 7-8 km in the margin of the EEC, which is known both from borehole data and seismic re-
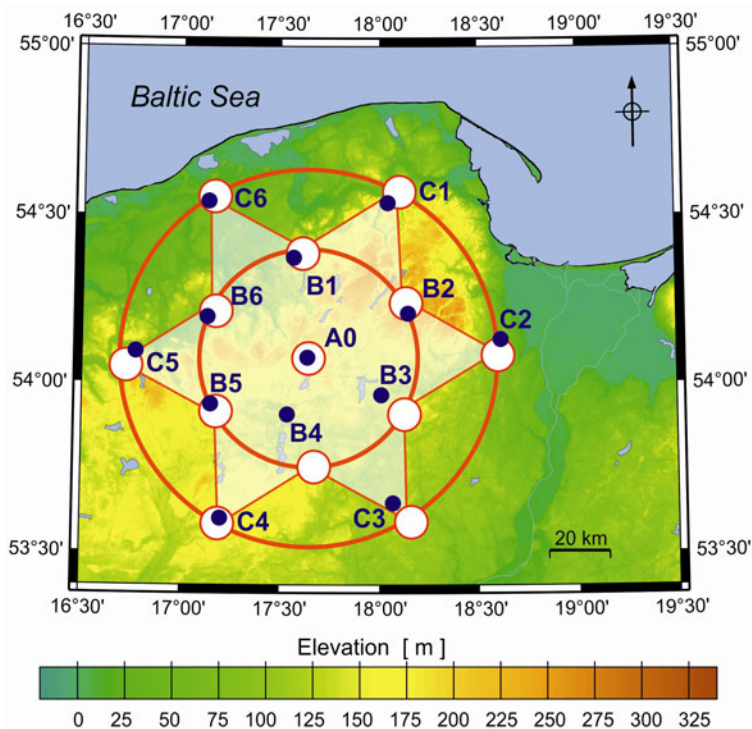
Fig. 3. Location map of "13 BB star" network in northern Poland on the background of the topography map. Red dots and circles show the planed regular geometry of the network where broadband seismometers are placed in basic equilateral triangles with side lengths of 35 km. The navy-blue dots are final locations of the stations (A0, B1-B6, C1-C6).

fraction (Skorupa 1974, Młynarski 1984). In the area of "13 BB star" network the Proterozoic basement has an age of 1.85-1.80 Ga (Bogdanova *et al.* 2006). The crystalline crust has a three-layer structure that is typical for EEC, with *P*-wave velocities of 6.1-6.4, 6.5-6.8, and 6.9-7.2 km/s in the upper, middle, and lower parts, respectively. Surrounding Moho depth is also well known there (see Fig. 2; Guterch and Grad 2006, Grad *et al.* 2009, Majdański 2012, Malinowski 2013). The Moho topography is relatively flat (about 40-45 km beneath stations), which also is favourable from the point of view of deep structure studies. The assumed shape of network is shown in Fig. 3. Red circles and red/white dots show the assumed regular geometry of the network where seismometers are placed in basic equilateral triangles with side lengths of 35 km. All places for stations were found in forests, on the glades or fenced nurseries, far away from local roads, villages, traffic, and other man-made disturbances. To be independent of electric power sup-

ply we use solar panels with batteries, which permit long term field operation. "Good quality" recording places should guarantee high signal-to-noise ratio. The navy-blue dots in Fig. 3 show final locations of the stations (A0, B1-B6, C1-C6). Such a system of 13 stations apart of basic small triangles with side lengths of 30-40 km creates larger triangles with side lengths of 60-80 and 90-110 km.

Installation of a temporary broadband seismic station in northern Poland is shown in Fig. 4, and a scheme with elements of each station in the network is shown in Fig. 5.

The "13 BB star" network consists 13 stations, equipped with broadband seismometers: model Reftek 151-120 – "Observer". The "Observer" is a low noise seismometer, which contains three independent sensors (one vertical and two horizontal) with built-in electronic feedback circuit, control, and power conversion circuits, and large dynamic range. It is a force-balance feedback sensor with frequency bandwidth of 0.0083-50 Hz (which corresponds to periods 120-0.02 s), flat to velocity. It has built-in levelling and

Fig. 4. Installation of the temporary broadband station in northern Poland: (a) Levelling broadband seismometer Reftek 151-120 on the granitic base in the *ca.* 1 m deep case; (b) Batteries, converter and recorder on the shelf, over the seismometer; (c) Solar panel, GPS and antenna for cell phone connection; (d) Closing the "garden" of seismic station in the forest area (photos taken by M. Grad, June 2013).
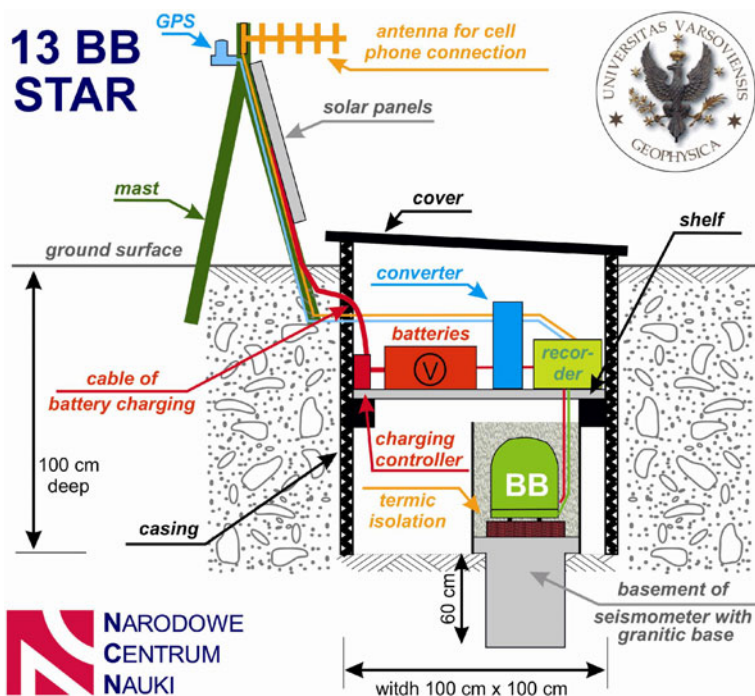
Fig. 5. Scheme of the station elements in temporary "13 BB star" network.

automatic mass zero-position adjusting mechanism. Monitoring and mass adjustment is performed via high resolution recorder Reftek 130B. All seismometers were placed on the 60 cm high reinforced concrete basement with granitic base in the *ca.* 1 m deep case with open bottom (Fig. 5). Above the seismometer, on the shelf, the recorder, own-design ARM Linux based data transmission set equipped with 3G/EDGE wireless modem, and batteries were placed. The case was covered and masked with natural materials like sand, soil, and grass making the case invisible. Over the ground, solar panels, GPS and antenna for cell phone connection were installed. Finally, the "garden" of seismic station (size 4 m × 4 m) was fenced and closed.

In the "13 BB star" project a special attention will be paid to high-quality seismic network observations and modern interpretation methods. Solving of this problem has fundamental contribution for seismology, tectonics, and geodynamics, not only in the regional scale but also worldwide. The results of "13 BB star" project shall contribute to better basic knowledge on the structure of the Earth's interior, its physical properties, mantle dynamics, tectonic evolution, and present-day processes. The data acquired during the 3-years long recording campaign will be analysed using different modern seismic

methods and integrated interpretations. They will yield images of detailed lithosphere-asthenosphere system hitherto unknown for the marginal zone of EEC in northern Poland. By applying surface waves, as well as the receiver function techniques, we expect to reveal variations of seismic velocities within the upper mantle. Regularity of network, its symmetry and size were chosen to enable gathering propagation properties of short, intermediate, and long period surface waves for all azimuths (Cotte *et al.* 2002, Bruneton *et al.* 2004). As shown, *e.g.*, by Jiang *et al.* (2011) Rayleigh wave phase-velocity sensitivity depends on the period of wave. Short wave (25 s period) has a peak sensitivity at about 25 km depth, and the longest wave (143 s period) has a peak sensitivity at about 200 km depth (see also Pasyanos 2010, Bartzsch *et al.* 2011). So, this is a depth we expect to reach using surface waves. Using receiver function technique "410", "?520", and "660" km discontinuities beneath the marginal zone of EEC will be investigated. As part of the study in the marginal part of EEC, beneath our network, the anisotropy will be delineated. *S*-wave splitting, observations of *SKS*, and *SKKS* phases will be collected from the network, and the polarization of the fast-waves will be determined. For strong events, in distance range of 30–40°, multiples of *ScS* phase could be observed (mantle reverberations), and high frequency *PKP*, *SKS*, and other core phases (Kennett and van der Hilst 1996, Shearer *et al.* 2011).
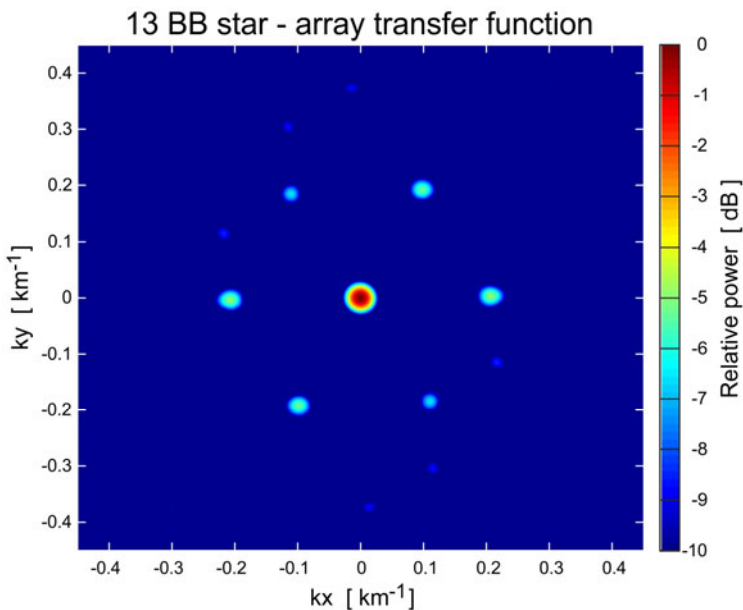

Fig. 6. Array transfer function of "13 BB star" seismic network. Colours are related to power of the array response normalized with its maximum.

The "13 BB star" network shape and station distribution may allow using some array analysis techniques. The evaluated array transfer function for "13 BB star" network is shown in Fig. 6. The wavenumber range of seismic signal, which can be investigated by "13 BB star" network using the array techniques, is *ca.* 0.05-0.2 1/km (Schweitzer *et al.* 2012). So, the high resolution in measuring apparent velocities and backazimuth can be obtained only for teleseismic waves.

The "13 BB star" network could be also used to help in the detection of small natural seismic events (see, *e.g.*, Wiszniowski *et al.* 2014). Poland is a region of very low natural seismicity. The largest earthquake with magnitude estimated at 5.6 occured on 3 December 1786, near Myślenice in southern Poland (see Fig. 2; Pagaczewski 1972). Over the last decade a few relatively strong earthquakes in Poland and surroundings have been widely felt in the area of Poland. The location of five events is shown in Fig. 2: two events close to Kaliningrad, near Jarocin, Bełchatów, and at Podhale. For more information see Domański (2007), Lizurek *et al.* (2013), Wiejacz and Rudziński (2010), and Wiejacz and Dębski (2009).

3. FIRST RESULTS

The "13 BB star" network started its operation on 19 July 2013. The recorded data are transmitted to central server dedicated for the project and available near on-line. All seismic data are sampled with 0.01 s (corresponding to 100 Hz). Beside the seismic data, additional state of health information is constantly transmitted and includes temperature, voltage, and power measurements. Server is equipped with proprietary web interface for monitoring network status and allows preliminary seismic data analysis.

Station noise analysis is presented in Fig. 7. Probabilistic Power Spectral Densities (PPSD) allow comparison of noise level between stations and to permanent stations. For example, station C6 is closest to sea shore (13.75 km in straight line), A0 is located close (*ca.* 100 m) to local paved road with traffic of about 20-30 cars per day. Permanent station GKP is closest to "13 BB star" network, while KSP is located in southern Poland and has lowest noise along Polish Seismic Network. For all stations, PPSDs are calculated for the same period of 11 months (July 2013 – June 2014) for one-hour segments (PSDs) with 50% overlap (Beyreuther *et al.* 2010, Megies *et al.* 2011). Only full one-hour segments are considered. All recorded data is considered for each station including seismic events (McNamara and Buland 2004). New high and low noise models (NHNM, NLNM) are shown for reference (Peterson 1993).

According to USGS/NEIC PDE Catalogue, in the period from 19 July 2013 to 24 June 2014, a total of 1543 earthquakes of magnitude ≥ 5.0 and 440 earthquakes of magnitude ≥ 5.5 occurred all over the word (Fig. 8).
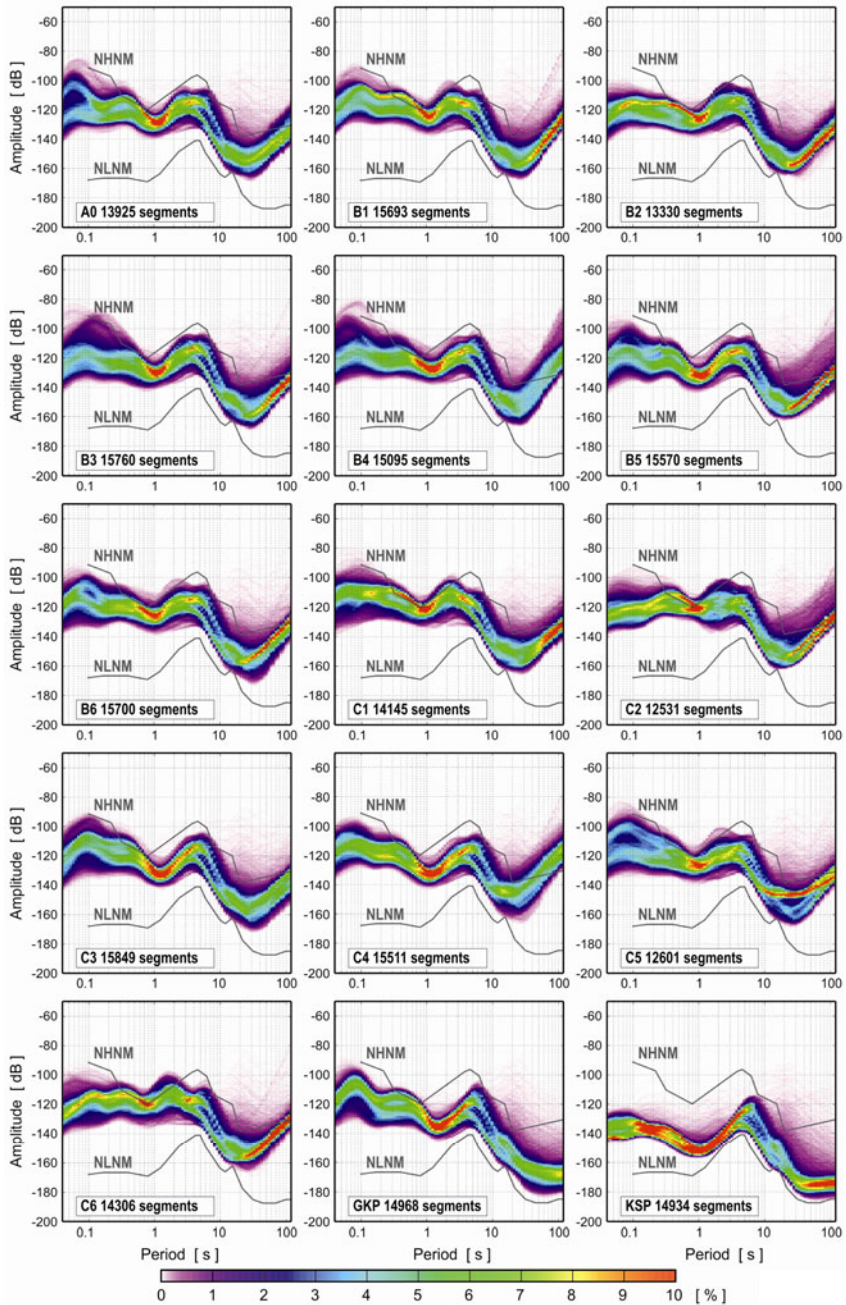
Fig. 7. Probabilistic Power Spectral Densities (PPSD) for all 13 stations and 2 permanent stations of Polish Seismic Network. New high and low noise models (NHNM, NLNM) are shown for reference (Peterson 1993).
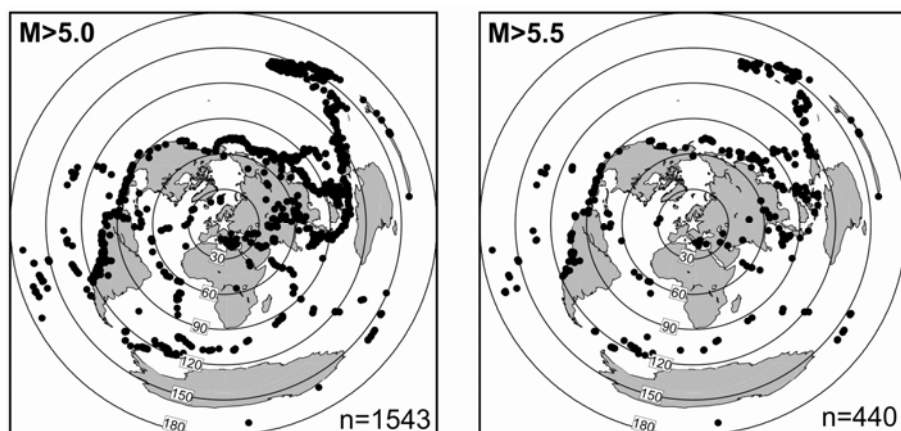
Fig. 8. Distribution of the earthquakes epicentres of magnitude ≥ 5.0 (left plot, 1543 events) and of magnitude ≥ 5.5 (right plot, 440 events) from 19 July 2013 to 24 June 2014 (according to USGS/NEIC PDE Catalogue). The azimuthal-equidistant projection shows the true distances and backazimuths of the epicentres with respect to the centre of "13 BB star" network (marked by star). Circles show epicentral distances of 30°, 60°, 90°, 120°, 150°, and 180°.

Examples of records of "13 BB star" network for five events are shown in Figs. 9-13 and their focal parameters are collected in Table 1. Figure 9 shows broadband seismograms of a teleseismic earthquake ($\Delta \approx 66°$) from the area of Okhotsk Sea. Figure 9a shows one hour records of Z, N, and E components at station C6. The event has strong P arrivals at Z component, and strong S arrivals at N and E components. The big depth of the earthquake ($h = 570$ km) could explain a total lack of surface waves. Figure 9b shows a zoom of iP and iS arrivals for all stations which confirms the uniform motion (indicated by arrows): "down" for P waves at Z component,

Table 1

Focal parameters of events recorded by "13 BB star" network

Event no.	Date	Time UT hh:mm:ss	Lat. ϕ [°N]	Long. λ [°E]	Region	Depth [km]	M	Distance from A0 [deg]
1	1 Oct 2013	03:38:21	53.17	152.88	Okhotsk Sea	570	6.7	66.52
2	6 Oct 2013	01:37:21	35.28	26.69	Romania	134	5.4	10.24
3	24 Sep 2013	11:29:50	27.06	65.56	Pakistan	30	7.4	44.05
4	13 Aug 2013	15:43:14	5.78	−78.23	South of Panama	2	6.6	88.75
5	15 Oct 2013	00:12:31	9.93	124.16	Philippines	7	7.1	91.42

Fig. 9. Seismograms of a teleseismic earthquake from Okhotsk Sea region recorded by broadband network in northern Poland during passive experiment "13 BB star". Panel (a) shows one hour records of *Z*, *N*, and *E* components at station C6. Note strong *P* arrivals at *Z* component, strong *S* arrivals at *N* and *E* components, and no surface waves. Panel (b) with zoom of *iP* and *iS* arrivals shows for all stations uniform motion (indicated by arrows): "down" for *P* waves at *Z* component, "up" for *S* waves at *N* component, and "down" for *S* waves at *E* component. All traces are lowpass filtered (< 1 Hz).

Fig. 10. Seismograms of a regional earthquake in Romania recorded by broadband network in northern Poland during passive experiment "13 BB star". Seismograms are sorted with distance from event. Note good signal-to-noise ratio and different shape of first *P*-wave arrivals (shown by arrows). All traces are highpass filtered (> 1 Hz).

"up" for *S* waves at *N* component, and "down" for *S* waves at *E* component. Differentiation of arrival times is easily seen for this event: the arrival time difference across the network is ca. 6.3 s for *P* waves and ca. 12.7 s for *S* waves.

Seismograms of a regional earthquake from Romania ($\Delta \approx 10°$) recorded by broadband "13 BB star" network are shown in Fig. 10 with first *P*-wave arrivals marked by arrows. Note good signal-to-noise ratio and different shape of arrivals. Seismograms are sorted with distance from the event epicenter. For this event the time difference of *P*-wave arrivals across the network is *ca.* 12 s. Seismograms of a teleseismic earthquake from Pakistan ($\Delta \approx 44°$) are shown in Fig. 11. Note good signal-to-noise ratio, clear *P* and *S* arrivals, and dominant surface waves. All traces are lowpass filtered (< 1 Hz). Figure 12 shows seismograms of a teleseismic earthquake from Panama ($\Delta \approx 88°$) recorded by "13 BB star" network. Note strong *P*, *PP*, and surface waves with distinct dispersion. All traces are lowpass filtered (< 1 Hz). Seismograms of the Philippines earthquake ($\Delta \approx 92°$) recorded by all stations are shown in Fig. 13. Note strong *P*, *PP*, *S*, and *SS* arrivals, as
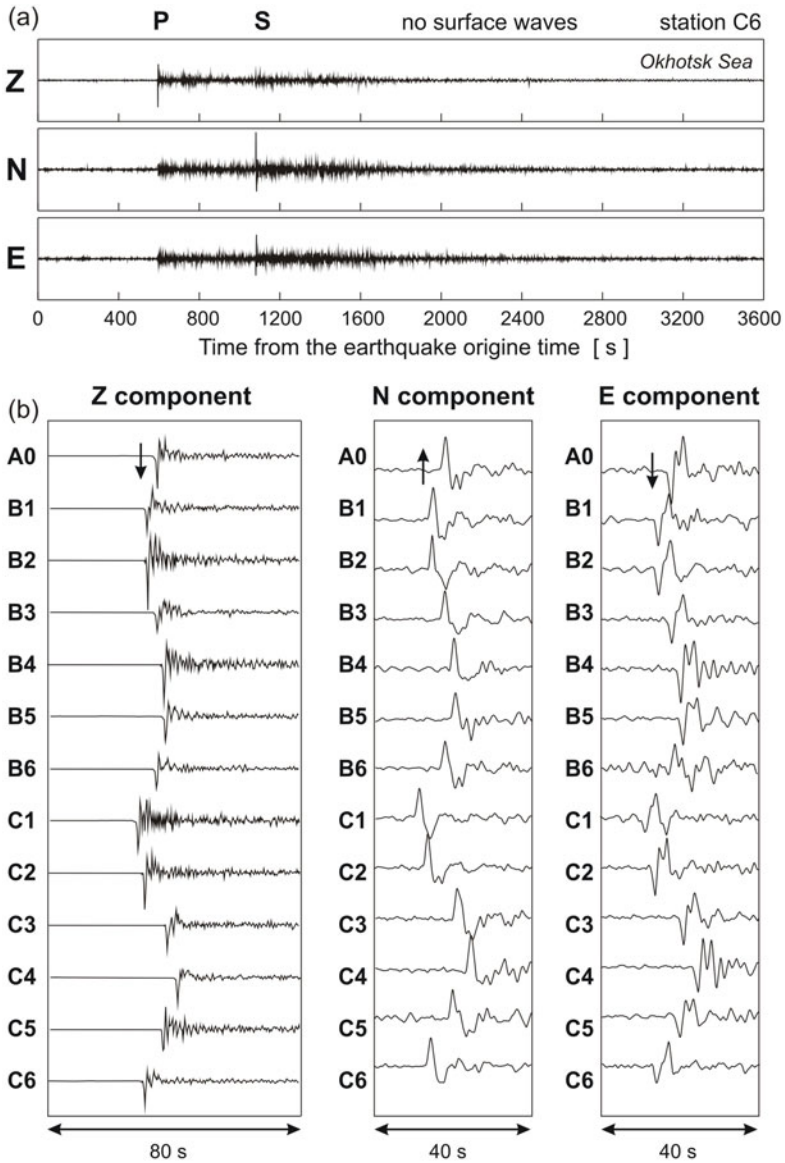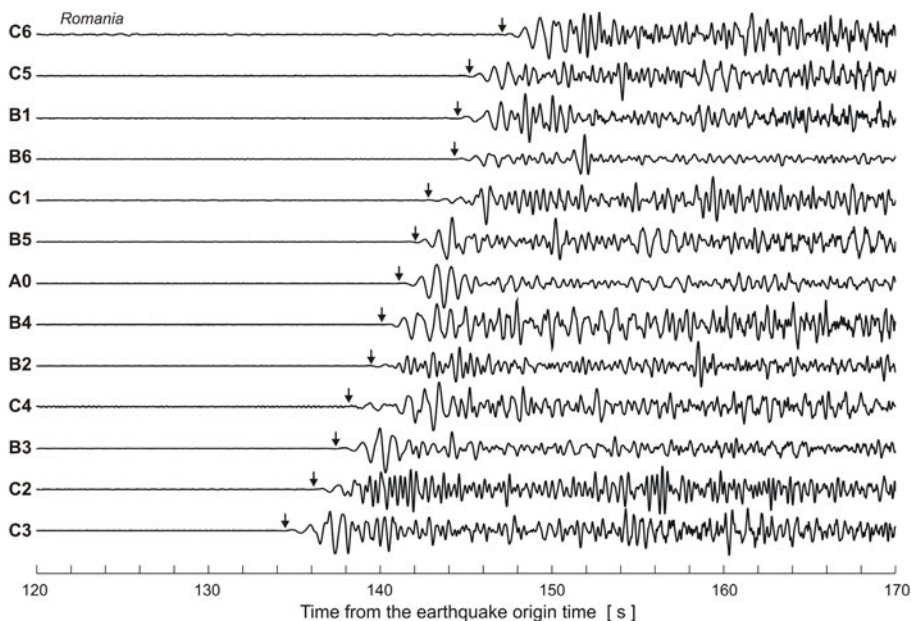
Fig. 11. Seismograms of a teleseismic earthquake in Pakistan recorded by broadband network in northern Poland during passive experiment "13 BB star". Note good signal-to-noise ratio and dominant surface waves. All traces are lowpass filtered (< 1 Hz).

well as full wave field of surface waves with distinct dispersion and separation of *LQ* and *LR* waves. This earthquake was strong ($M = 7.1$) and relatively shallow ($h = 7$ km); traces are not filtered.

The examples of seismograms presented in Figs. 9-13 show frequency differentiation for different seismic phases. More sophisticated representations of the records are displayed in Fig. 14 – spectral seismograms, which allow quantitative analyses of the frequency content of the recorded seismic phases. This technique was also employed for the determination of magnitude spectra, the source parameters, and discrimination between earthquakes and explosions (*e.g.*, Duda *et al.* 1996, Domański 2007, Wilde-Piórko *et al.* 2011). Spectral seismograms obtained from the broadband seismograms of
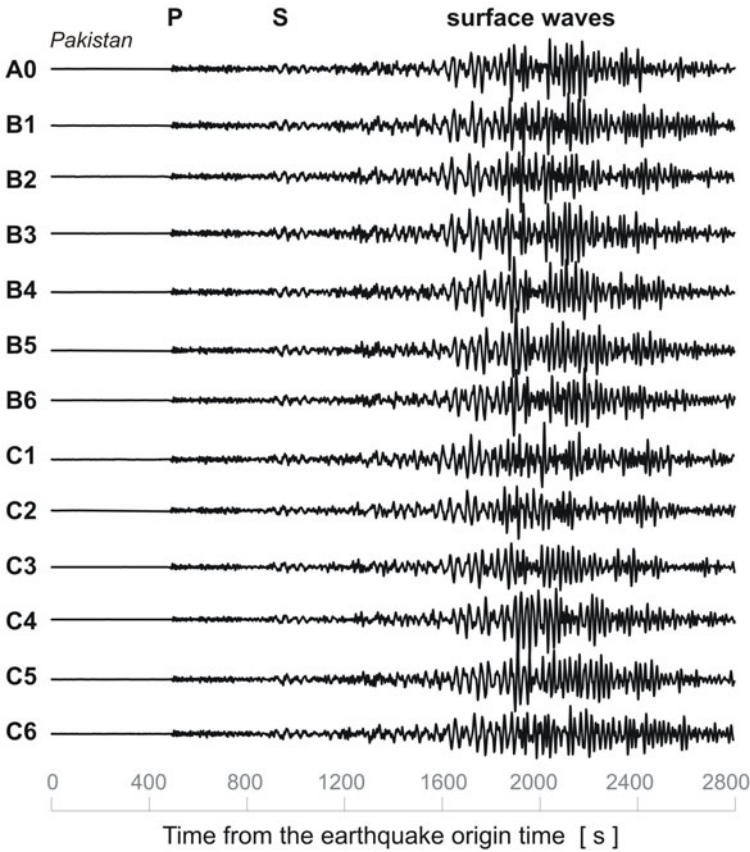
Fig. 12. Seismograms of a teleseismic earthquake in Panama recorded by broadband network in northern Poland during passive experiment "13 BB star". Note strong *P*, *PP*, and surface waves. All traces are lowpass filtered (< 1 Hz).

four teleseismic events are shown in Fig. 14. They are obtained by applying Butterworth filters of order 2, with the widths of 0.5 octave (Fasthoff and Guo 2001).

An interesting case is the earthquake from Okhotsk Sea (Fig. 14a-c). The pulse of body waves is relatively short and strong. The maximum amplitude of *P*-wave radiation of this event is in the period range of 1-5 s (Fig. 14a; *Z* component), while for stronger *S*-waves this period is 3-10 s (Fig. 14b and c; *N* and *E* components, respectively). In the range of longer periods, a total lack of energy is observed, related with very deep source, which do not produce surface waves. In the spectral seismogram of the Pakistan earthquake (Fig. 14d; *Z* component) dominant amplitudes are related to surface waves. Dispersion is easily seen: at time of 1400 s from the beginning of earthquake the maximum amplitude is observed for period *ca.* 90 s, while for the time of 2500 s it is 20 s only. For the Panama earthquake (Fig. 14e; *Z* component) a separation of surface waves is clear: the maximum amplitude of *LQ*-waves is observed for period *ca.* 90 s at time of 2600 s, and for *LR*-waves this period is 20-10 s at time 2800-3400 s. For the Philippines earthquake (Fig. 14f;
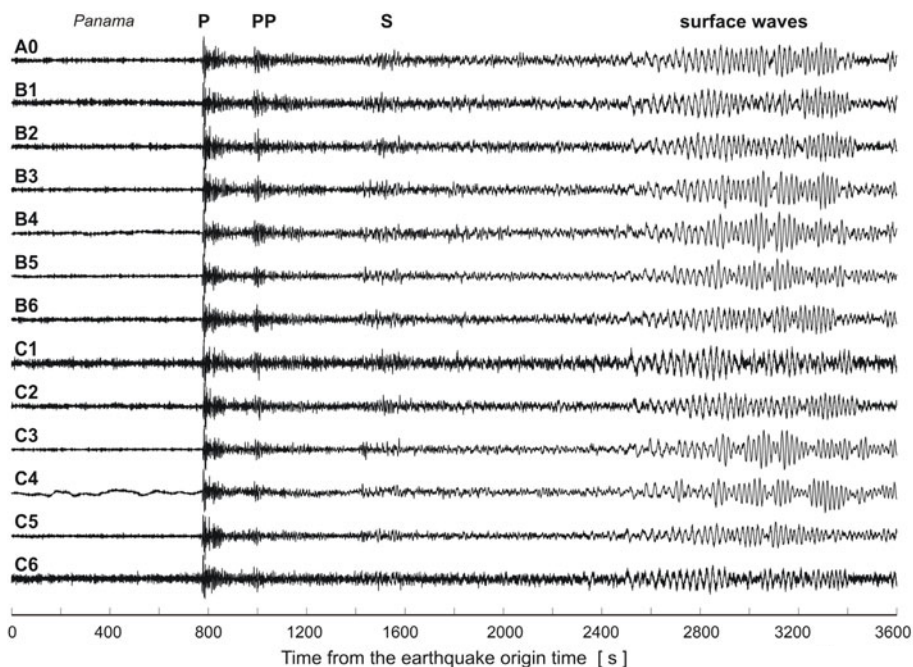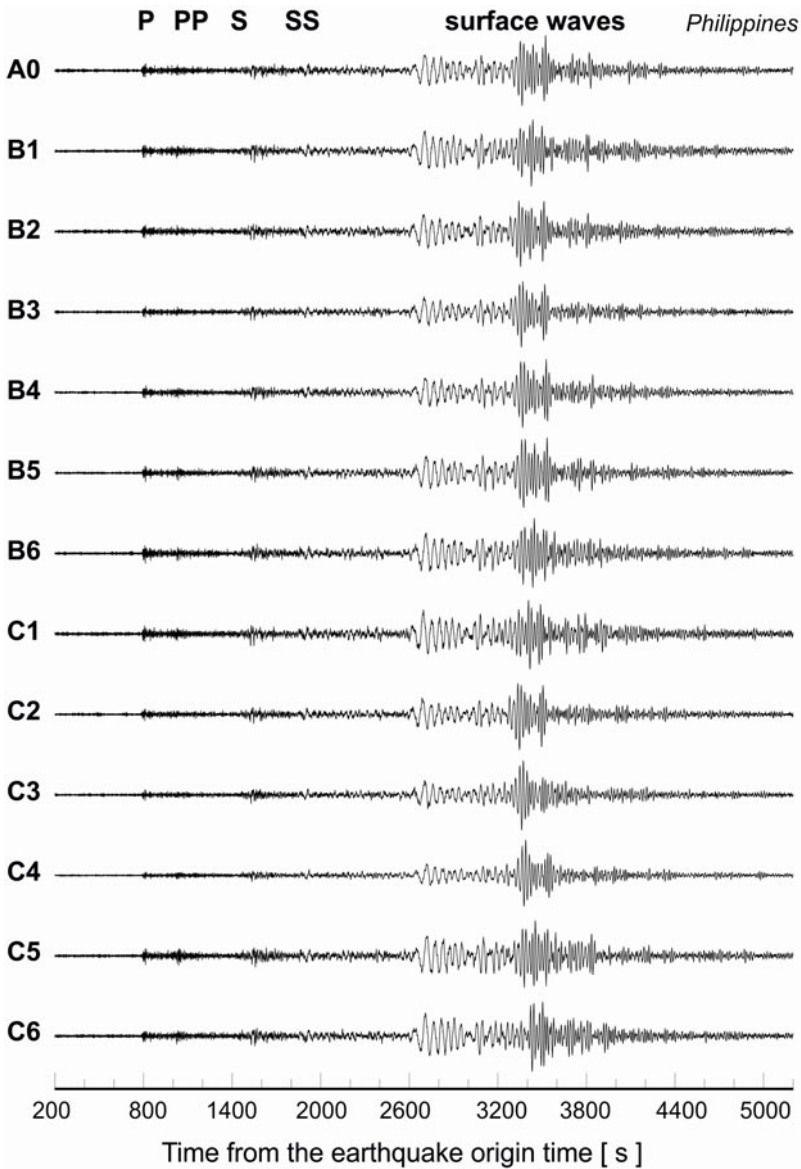
Fig. 13. Seismograms of a teleseismic earthquake from Philiphines recorded by broadband network in northern Poland during passive experiment "13 BB star". Note strong *P*, *PP*, *S*, and *SS* arrivals, as well as full wave field of surface waves with distinct dispersion. Traces are not filtered.

Z component) body waves are relatively weak, and dominant are *LQ*-waves with period *ca.* 70 s at time of 2800 s.
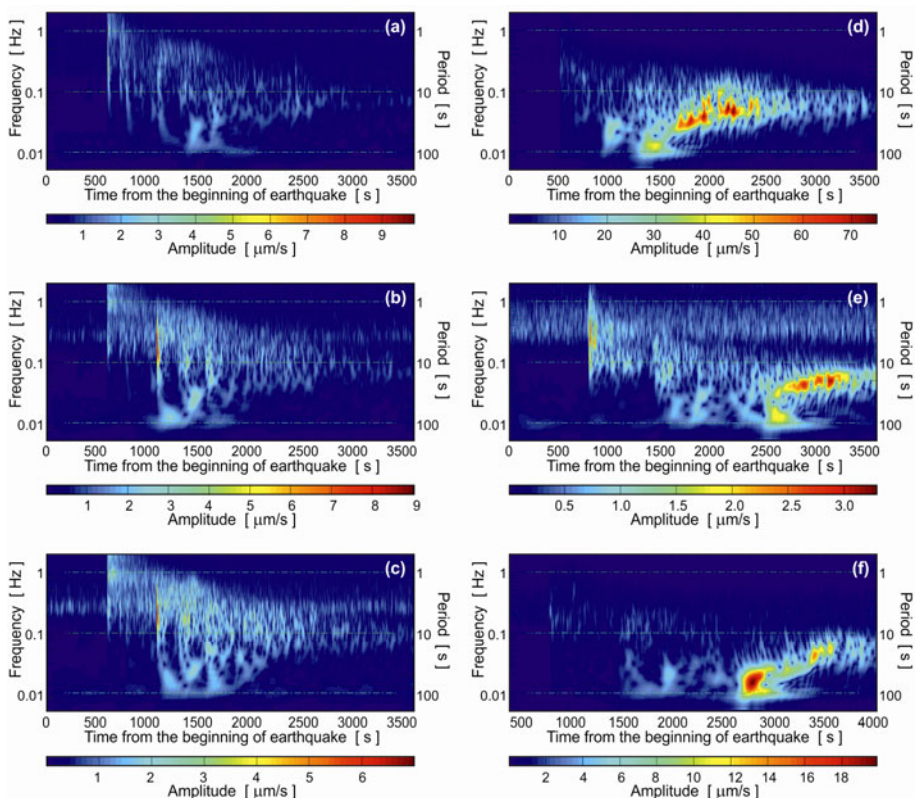
Fig. 14. Spectral seismograms obtained from records of the broadband network in northern Poland during passive experiment "13 BB star": (a), (b), and (c) spectral seismograms of the Okhotsk Sea earthquake recorded at station C6, for *Z*, *N*, and *E* components, respectively; (d) spectral seismograms of the Pakistan earthquake recorded at station A0, *Z* component; (e) spectral seismograms of the Panama earthquake recorded at station C5, *Z* component; (f) spectral seismograms of the Philippines earthquake recorded at station A0, *Z* component. All spectral seismograms were obtained by applying Butterworth filters of order 2, with the widths of 0.5 octaves. Note different amplitude scales.

4. SUMMARY

The knowledge of the Earth's lithosphere-asthenosphere system is far from complete. This is the main task of "13 BB star" experiment, and to solve this problem surface wave techniques and receiver function will be in use (see, *e.g.*, Cotte *et al.* 2002, Bruneton *et al.* 2004, Wilde-Piórko *et al.* 2010, Trojanowski and Wilde-Piórko 2012). Using array techniques of broadband seismology, also deeper Earth's structure will be investigated: "410", "?520", and "670" km boundaries (Gossler and Kind 1996), and, if possible,

also deep mantle and mantle-core boundary with *S*-wave splitting, *ScS*, *SKS*, and *SKKS* phases (Kennett and van der Hilst 1996). The good knowledge of the crustal structure beneath the "13 BB star" network is an advantage of its location. The knowledge of seismic wave velocity distribution in sedimentary cover is of great importance: for receiver function technique because of big sediments-basement velocity contrast (*e.g.*, 2.5 to 6 km/s), and for seismic tomography, particularly in the cases of very deep basins, because of strongly differentiated velocities in thick strata. In the area of "13 BB star" network, velocities in sediments are well known from vertical seismic profiling in deep boreholes and shallow refraction.

Finally, good records obtained till now, and expected during next 1-year long recording campaign, should yield images of detailed structure in the marginal zone of the EEC in northern Poland.

R e f e r e n c e s

Bartzsch, S., S. Lebedev, and T. Meier (2011), Resolving the lithosphere-asthenosphere boundary with seismic Rayleigh waves, *Geophys. J. Int.* **186**, 3, 1152-1164*,* DOI: 10.1111/j.1365-246X.2011.05096.x.

Berthelsen, A. (1998), The Tornquist Zone northwest of the Carpathians: An intraplate pseudosuture, *GFF* **120**, 2, 223-230, DOI: 10.1080/11035899801202223.

Beyreuther, M., R. Barsch, L. Krischer, T. Megies, Y. Behr, and J. Wassermann (2010), ObsPy: A Python toolbox for seismology, *Seismol. Res. Lett.* **81**, 3, 530-533, DOI: 10.1785/gssrl.81.3.530.

Bogdanova, S., R. Gorbatschev, M. Grad, T. Janik, A. Guterch, E. Kozlovskaya, G. Motuza, G. Skridlaite, V. Starostenko, L. Taran, and EUROBRIDGE and POLONAISE Working Groups (2006), EUROBRIDGE: new insight into the geodynamic evolution of the East European Craton, *Geol. Soc. London Mem.* **32**, 599-625, DOI: 10.1144/GSL.MEM.2006.032.01.36.

Bruneton, M., H.A. Pedersen, R. Farra, N.T. Arndt, P. Vacher, U. Achauer, A. Alinaghi, J. Ansorge, G. Bock, W. Friedrich, M. Grad, A. Guterch, P. Heikkinen, S.E. Hjelt, T.L. Hyvonen, J.P. Ikonen, E. Kissling, K. Komminaho, A. Korja, E. Kozlovskaya, M.V. Nevsky, H. Paulssen, N.I.

Pavlenkova, J. Plomerova, T. Raita, O.Y. Riznichenko, R.G. Roberts, S. Sandoval, I.A. Sanina, N.V. Sharov, Z.H. Shomali, J. Tiikkainen, E. Wielandt, K. Wylegalla, J. Yliniemi, and Y.G. Yurov (2004), Complex lithospheric structure under the central Baltic Shield from surface wave tomography, *J. Geophys. Res.* **109**, B10, B10303, DOI: 10.1029/2003JB 002947.

Cotte, N., H.A. Pedersen, and TOR Working Group (2002), Sharp contrast in lithospheric structure across the Sorgenfrei–Tornquist Zone as inferred by Rayleigh wave analysis of TOR1 project data, *Tectonophysics* **360**, 1-4, 75-88, DOI: 10.1016/S0040-1951(02)00348-7.

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

Duda, S.J., J. Saul, and M. Grad (1996), The influence of near-source and near-receiver structure on the spectrum of *P* waves from nuclear explosions, *Acta Geophys. Pol.* **44**, 1, 1-18.

Eaton, D.W., F. Darbyshire, R.L. Evans, H. Grütter, A.G. Jones, and X. Yuan (2009), The elusive lithosphere-asthenosphere boundary (LAB) beneath cratons, *Lithos* **109**, 1-2, 1-22, DOI: 10.1016/j.lithos.2008.05.009.

Fasthoff, S., and L. Guo (2001), On spectral seismograms – 3D display of seismogram, *CT Theor. Appl.* **10**, 3, 47-51.

Gossler, J., and R. Kind (1996), Seismic evidence for very deep roots of continents, *Earth Planet. Sci. Lett.* **138**, 1-4, 1-13, DOI: 10.1016/0012-821X(95)00215-X.

Grabowska, T., G. Bojdys, M. Bielik, and K. Csicsay (2011), Density and magnetic models of the lithosphere along CELEBRATION 2000 profile CEL01, *Acta Geophys.* **59**, 3, 526-560, DOI: 10.2478/s11600-011-0007-3.

Grad, M. (1986), Seismic P-wave velocities of the East European Platform in Poland, *Acta Geophys. Pol.* **34**, 1, 21-29.

Grad, M., and M. Polkowski (2012), Seismic wave velocities in the sedimentary cover of Poland: Borehole data compilation, *Acta Geophys.* **60**, 4, 985-1006, DOI: 10.2478/s11600-012-0022-z.

Grad, M., G.R. Keller, H. Thybo, A. Guterch, and POLONAISE Working Group **(**2002), Lower lithospheric structure beneath the Trans-European Suture Zone from POLONAISE'97 seismic profiles, *Tectonophysics* **360**, 1-4, 153-168, DOI: 10.1016/S0040-1951(02)00350-5.

Grad, M., S.L. Jensen, G.R. Keller, A. Guterch, H. Thybo, T. Janik, T. Tiira, J. Yliniemi, U. Luosto, G. Motuza, V. Nasedkin, W. Czuba, E. Gaczyński, P. Środa, K.C. Miller, M. Wilde-Piórko, K. Komminaho, J. Jacyna, and L. Korabliova (2003), Crustal structure of the Trans-European suture zone region along POLONAISE'97 seismic profile P4, *J. Geophys. Res.* **108**, B11, 2541, DOI: 10.1029/2003JB002426.

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

Guterch, A., and M. Grad (2006), Lithospheric structure of the TESZ in Poland based on modern seismic experiments, *Geol. Quart.* **50**, 1, 23-32.

Jiang, M., S. Zhou, E. Sandvol, X. Chen, X. Liang, Y.J. Chen, and W. Fan (2011), 3-D lithospheric structure beneath southern Tibet from Rayleigh-wave tomography with a 2-D seismic array, *Geophys. J. Int.* **185**, 2, 593-608, DOI: 10.1111/j.1365-246X.2011.04979.x.

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

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

Kennett, B.L.N, and R.D. van der Hilst (1996), Using a synthetic continental array to study the Earth's interior, *J. Phys. Earth* **44**, 6, 669-674, DOI: 10.4294/jpe1952.44.669.

Kozlovskaya, E., G. Kosarev, I. Aleshin, O. Riznichenko, and I. Sanina (2008), Structure and composition of the crust and upper mantle of the Archean–Proterozoic boundary in the Fennoscandian shield obtained by joint inversion of receiver function and surface wave phase velocity of recording of the SVEKALAPKO array, *Geophys. J. Int.* **175**, 1, 135-152, DOI: 10.1111/j.1365-246X.2008.03876.x.

Królikowski, C., and Z. Petecki (1997), Crustal structure at the Trans-European Suture Zone in northwest Poland based on the gravity data, *Geol. Mag.* **134**, 5, 661-667, DOI: 10.1017/S0016756897007395.

Krysiński, L., M. Grad, and POLONAISE Working Group (2000), POLONAISE'97 – Seismic and gravimetric modelling of the crustal structure in the Polish basin, *Phys. Chem. Earth A* **25**, 4, 355-363, DOI: 10.1016/S1464-1895(00)00057-0.

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

Majdański, M. (2012), The structure of the crust in TESZ area by kriging interpolation, *Acta Geophys.* **60**, 1, 59-75, DOI: 10.2478/s11600-011-0058-5.

Majorowicz, J.A., V. Čermak, J. Šafanda, P. Krzywiec, M. Wróblewska, A. Guterch, and M. Grad (2003), Heat flow models across the Trans-European Suture Zone in the area of the POLONAISE'97 seismic experiment, *Phys. Chem. Earth* **28**, 9-11, 375-391, DOI: 10.1016/S1474-7065(03)00059-7.

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

McNamara, D.E., and R.P. Buland (2004), Ambient noise levels in the continental United States, *Bull. Seismol. Soc. Am.* **94**, 4, 1517-1527, DOI: 10.1785/ 012003001.

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

Meissner, R. (1986), *The Continental Crust – A Geophysical Approach,* International-al Geophysics Series, Vol. 34, Academic Press Inc., Orlando, 426 pp.

Młynarski, S. (1984), The structure of deep bedrock in Poland on the basis of refraction results, *Publs. Inst. Geophys. Pol. Acad. Sci.* **A-13**, 160, 87-100.

Pagaczewski, J. (1972), Catalogue of earthquakes in Poland in 1000-1970 years, *Publs. Inst. Geophys. Pol. Acad. Sci.* **51**, 3-36.

Pasyanos, M.E. (2010), Lithospheric thickness modeled from long-period surface wave dispersion, *Tectonophysics* **481**, 1-4, 38-50, DOI: 10.1016/j.tecto. 2009.02.023.

Peterson, J. (1993), Observation and modeling of seismic background noise, U.S. Dept. of Interior, Geological Survey, Open-file Rep. 93-322, 95 pp.

Pharaoh, T.C., R.W. England, J. Verniers, and A. Żelaźniewicz (1997), Introduction: geological and geophysical studies in the Trans-European Suture Zone, *Geol. Mag.* **134**, 5, 585-590, DOI: 10.1017/S0016756897007619.

Pożaryski, W., W. Brochwicz-Lewiński, and H. Tomczyk (1982), Sur la caractére hétérochronique de la Ligne Teisseyre-Tornquist, entre Europe centrale et orientale, *C.R. Acad. Sci. Paris II Tect.* **295**, 691-696 (in French).

Romanowicz, B. (2009), The thickness of tectonic plates, *Science* **324**, 5926, 474-476, DOI: 10.1126/science.1172879.

Sandoval, S., E. Kissling, J. Ansorge, and the SVEKALAPKO Seismic Tomography Working Group (2004), High-resolution body wave tomography beneath the SVEKALAPKO array – II. Anomalous upper mantle structure beneath the central Baltic Shield, *Geophys. J. Int.* **157**, 1, 200-214, DOI: 10.1111/ j.1365-246X.2004.02131.x.

Schweitzer, J., J. Fyen, S. Mykkeltveit, S.J. Gibbons, M. Pirlii, D. Kühn, and T. Kværna (2012), Seismic arrays. **In:** P. Bormann (ed.), *New Manual of Seismological Observatory Practice 2 (NMSOP-2)*, Deutsches GeoForschungsZentrum, Potsdam, 1-80, DOI: 10.2312/GFZ.NMSOP-2_ch9.

Shearer, P.M., C.A. Rychert, and Q. Liu (2011), On the visibility of the inner-core shear wave phase *PKJKP* at long periods, *Geophys. J. Int.* **185**, 3, 1379-1383, DOI: 10.1111/j.1365-246X.2011.05011.x.

Skorupa, J. (1974), Seismic velocity map of Poland 1:500 000, Wyd. Geol., War-szawa.

Trojanowski, J., and M. Wilde-Piórko (2012), S-velocity structure beneath the Bohemian Massif from Monte Carlo inversion of seismic receiver function, *Acta Geophys.* **60**, 1, 76-91, DOI: 10.2478/s11600-011-0047-8.

Wessel, P., and W.H.F. Smith (1995), The Generic Mapping Tools GMT Version 3, Technical reference and cookbook, School of Ocean and Earth Science and Technology, University of Hawaii, Manoa, NOAA.

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

Wiejacz, P., and Ł. Rudziński (2010), Seismic event of January 22, 2010 near Bełchatów, Poland, *Acta Geophys.* **58**, 6, 988-994, DOI: 10.2478/s11600-010-0030-9.

Wilde-Piórko, M., M. Świeczak, M. Grad, and M. Majdański (2010), Integrated seismic model of the crust and upper mantle of the Trans-European Suture zone between the Precambrian craton and Phanerozoic terranes in Central Europe, *Tectonophysics* **481**, 1-4, 108-115, DOI: 10.1016/j.tecto.2009.05.002.

Wilde-Piórko, M., S.J. Duda, and M. Grad (2011), Frequency analysis of the 2004 Sumatra-Andaman earthquake using spectral seismograms, *Acta Geophys.* **59**, 3, 483-501, DOI: 10.2478/s11600-011-0010-8.

Wiszniowski, J., B.M. Plesiewicz, and J. Trojanowski (2014), Application of real time recurrent neural network for detection of small natural earthquakes in Poland, *Acta Geophys.* **62**, 3, 469-485, DOI: 10.2478/s11600-013-0140-2.

Ziegler, P.A. (1990), *Geological Atlas of Western and Central Europe*, 2nd ed., Shell Int. Petrol. Maatschappij, Hague, Geological Society Publ. House, London.

Znosko, J. (1975), Tectonic units of Poland against the background of the tectonics of Europe, *Geol. Inst. Anniversary Bull.* **252**, 61-75.

Znosko, J. (1979), Teisseyre-Tornquist tectonic zone: some interpretative implications of recent geological and geophysical investigations, *Acta Geol. Pol.* **29**, 4, 365-382.

Release-Rebound Processes: Vector Motions

Roman TEISSEYRE

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

Abstract

The tensor relations describing the shear deviatoric strains and rotation strains may be presented as vector relations in a special coordinate system, *e.g.*, in the diagonal or off-diagonal one. However, these fields can be also presented in the 4D invariant forms by means of invariant Dirac tensors. We present 4D relativistic relations for the invariant shear deviatoric strain and rotation strain vectors closely related to a fracture process in solids and to the molecular strains (shear and rotational) in fluids. These shear and rotation strains may interact with the radial derivatives of pressure along the propagation directions.

Key words: release-rebound, Dirac tensors, strain vectors, fracture relations.

1. INTRODUCTION

The release-rebound motions originate due to the fracture processes; to present such relations we will follow the 4D Maxwell space-time system. We assume that at a source of elastic waves, the molecular transport motions become transformed to the real fracture motion as a source of the observed radiated fields, that is, the rotation and shear strains.

2. RELEASE-REBOUND PROCESSES; VECTOR MOTIONS

In our considerations on some important relations for vectors and tensors in a special coordinate system, *e.g.*, in the diagonal or off-diagonal one, we can demand that any of such relations may be presented in an invariant 4D form.

Thus, we underline that such important relations, which in the 3D space may be achieved only in a special coordinate system, could become invariant in the 4D system, quite similarly to the electro-magnetic fields; we may call these relations the 4D Maxwell-like invariant relations.

For the total symmetric and antisymmetric strains, the axial, shear, and rotation strains, $\bar{E}, \hat{E}_{(ik)}, \breve{E}_{[ik]}$, can be written as:

$$E_{ik} = E_{(ik)} + E_{[ik]} = \delta_{ik}\bar{E} + \hat{E}_{(ik)} + \breve{E}_{[ik]} \qquad \text{- total strains ,}$$

$$\bar{E}(\bar{U}) = \frac{1}{3}\sum_s E_{ss} = \frac{1}{3}\sum_s \frac{\partial \bar{U}_s}{\partial x_s} \qquad \text{- axial strain ,}$$

$$\hat{E}_{(ik)}(\hat{U}) = E_{(ik)} - \delta_{ik}\bar{E} = \frac{1}{2}\left(\frac{\partial \hat{U}_k}{\partial x_i} + \frac{\partial \hat{U}_i}{\partial x_k}\right) - \delta_{ik}\frac{1}{3}\sum_{s=1}^{3}\frac{\partial \bar{U}_s}{\partial x_s} \quad \text{- shear deviatoric strains ,}$$

$$\breve{E}_{[ik]}(\breve{U}) = \frac{1}{2}\left(\frac{\partial \breve{U}_k}{\partial x_i} - \frac{\partial \breve{U}_i}{\partial x_k}\right) \qquad \text{- rotation strain ,}$$

(1a)

where we have used the potential displacement fields, $\bar{U}_s$, $\hat{U}_k$, $\breve{U}_n$, which could be treated as quite independent potentials.

We write also for pressure related terms:

$$\bar{E} = 1/3\sum E_{ss} , \quad \bar{S} = 1/3\sum S_{ss} , \quad \bar{S} = (2\mu + 3\lambda)\bar{E} , \quad \text{and} \quad p = -\bar{S} , \qquad (1b)$$

where the axial strain relates to pressure.

However, the symmetric, shear strains could be presented in the special off-diagonal system also as a vector field defined in that system; thus, we obtain:

$$E_{(i)} \rightarrow \left\{E_{(23)}, E_{(31)}, E_{(12)}\right\} , \quad E_{[i]} \rightarrow \left\{E_{[23]}, E_{[31]}, E_{[12]}\right\} ,$$

$$E_{[s]} = \frac{1}{2}\varepsilon_{skl}E_{[kl]} , \quad E_{(s)} = \frac{1}{2}\left|\varepsilon_{skl}\right|E_{(kl)}^{D} , \quad E_i = E_{[i]} + iE_{(i)} .$$

(1c)

However, we may present these strains as the complex invariant vectors; to this end, we should define the special off-diagonal vectors (that means, the vectors defined in the off-diagonal system). To achieve this goal we should define the invariant complex rotation and shear strain vectors by using the Dirac tensors in the system $\{\mathbf{x}, ict\}$:

$$
\gamma^1 = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}, \quad
\gamma^2 = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix},
$$

$$
\gamma^3 = i\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}, \quad
\gamma^4 = i\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix},
$$
(2a)

$$
\gamma^1\gamma^3 = \begin{bmatrix} i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & i & 0 \\ 0 & 0 & 0 & -i \end{bmatrix}, \quad
\gamma^2\gamma^3 = \begin{bmatrix} 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & -i & 0 \end{bmatrix}, \quad
\gamma^1\gamma^2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix},
$$

$$
\gamma^1\gamma^4 = \begin{bmatrix} 0 & 0 & 0 & i \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{bmatrix}, \quad
\gamma^2\gamma^4 = \begin{bmatrix} 0 & 0 & i & 0 \\ 0 & 0 & 0 & i \\ -i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{bmatrix}, \quad
\gamma^3\gamma^4 = -\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix},
$$
(2b)

and

$$
\gamma^4\gamma^2\gamma^3 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{bmatrix}.
$$
(2c)

Now, for the vector $E_s = E_{[s]} + iE_{(s)} \to E_\sigma$ we may write

$$
E_{\lambda\kappa} = E_{[\lambda\kappa]} + iE_{(\lambda\kappa)} \to E_{23}\gamma^1 + E_{31}\gamma^2 + E_{12}\gamma^4\gamma^2\gamma^3 = \begin{bmatrix} 0 & -E_3 & -E_2 & -E_1 \\ -E_3 & 0 & E_1 & -E_2 \\ -E_2 & E_1 & 0 & -E_3 \\ -E_1 & -E_2 & -E_3 & 0 \end{bmatrix}
$$
(3a)

or equivalently for $E_{23} \equiv E_1$, $E_{31} \equiv E_2$, $E_{12} \equiv E_3$:

$$
E_\sigma = E_{[\sigma]} + iE_{(\sigma)} = E_1\gamma^1 + E_2\gamma^2 + E_3\gamma^4\gamma^2\gamma^3 = \begin{bmatrix} 0 & -E_3 & -E_2 & -E_1 \\ -E_3 & 0 & E_1 & -E_2 \\ -E_2 & E_1 & 0 & -E_3 \\ -E_1 & -E_2 & -E_3 & 0 \end{bmatrix}.
$$
(3b)

Thus, the wave equations for deviatoric shear and rotation strains could be presented in the following way:

$$\sum_n \frac{\partial^2}{\partial x_n^2} E_{\lambda n} - \frac{\kappa^2 \partial^2}{c^2 \partial t^2} E_{\lambda n} = Y_{\lambda n}, \quad \text{or} \quad \Box E_{\lambda \eta} = Y_{\lambda \eta}, \tag{4}$$

where instead of the related strain wave velocity we introduce the material constant, κ, and the invariant light velocity c, $(V = c/\kappa)$.

We may add the external forces, $Y_{\lambda \kappa}$:

$$Y_{\lambda \kappa} = iY_{23}\gamma^1 + iY_{31}\gamma^2 + Y_{12}\gamma^3 = \begin{bmatrix} 0 & Y_{12} & -Y_{31} & -Y_{23} \\ -Y_{12} & 0 & Y_{23} & -Y_{31} \\ Y_{31} & -Y_{23} & 0 & -Y_{12} \\ Y_{23} & Y_{31} & Y_{12} & 0 \end{bmatrix}, \tag{5a}$$

or

$$Y_{\mu} = iY_1\gamma^1 + iY_2\gamma^2 + Y_3\gamma^3 = \begin{bmatrix} 0 & Y_3 & -Y_2 & -Y_1 \\ -Y_3 & 0 & Y_1 & -Y_2 \\ Y_2 & -Y_1 & 0 & -Y_3 \\ Y_1 & Y_2 & Y_3 & 0 \end{bmatrix}. \tag{5b}$$

In this way we arrive at the equations for the deviatoric shear and rotation strain vectors, $E_s = E_{[s]} + iE_{(s)} \rightarrow E_\sigma$:

$$\sum_n \frac{\partial^2}{\partial x_n^2} E_\sigma - \frac{\kappa^2 \partial^2}{c^2 \partial t^2} E_\sigma = Y_\sigma, \quad \text{or} \quad \Box E_\sigma = Y_\sigma. \tag{6}$$

We assume that these complex relations could be split into the following Maxwell-like equations for the vectors $E_{(s)}$, $E_{[i]}$:

$$\varepsilon_{spq} \frac{\partial}{\partial x_p} E_{[q]} - \frac{\kappa}{c} \frac{\partial}{\partial t} E_{(s)} = \frac{4\pi\kappa}{c} J_s, \quad \varepsilon_{spq} \frac{\partial}{\partial x_p} E_{(q)} + \frac{\kappa}{c} \frac{\partial}{\partial t} E_{[s]} = 0, \tag{7a}$$

where $\partial E_{[s]}/\partial x_s = 0$, $\partial E_{(s)}/\partial x_s = 4\pi\varepsilon$, and we have introduced the defect related current field, J_k, and propagation velocity $V = c/\kappa$, and the big constant, κ, depends on a given material and related conditions.

The Maxwell-like relations can describe both the strain wave propagation and also the fracture events occurring in different conditions and materials.

The related 4D form for these Maxwell-like relations (Eq. 7a) can be presented as follows:

$$\sum_\kappa \frac{\partial}{\partial x^\kappa} E_{\lambda\kappa} = \frac{4\pi\kappa}{c} J_\lambda, \quad x^\lambda = \{x^1, x^2, x^3, x^4\}, \quad x^4 = i\frac{c}{\kappa}t. \tag{7b}$$

At a fracture process, *e.g.*, an earthquake, a wave propagation will be excited (*cf.*, Teisseyre 1985, 2011). The related wave mosaic explains the interrelated propagation of shear and rotation motions.

We may note that a compressional load might induce defects forming the centers with opposite-sense shears and rotations, and, accordingly, the micro-breaks may lead to the fragmentation processes (rebound slip). We should note that the theories of dislocation and dislocation arrays (Eshelby *et al.* 1951, *cf.*, Rybicki 1986) lead to a better understanding of the roles of the defect densities and applied stresses (*cf.*, Kossecka and DeWitt 1977 and Teisseyre and Górski 2012) which brings us closer to our understanding of the fracture processes and dynamic synchronization.

Note that the Gibbs free energy of defect formation has been recently specified by Majewski and Teisseyre (2013) by succesfully extending a thermodynamical model; for a recent review see Varotsos (2007).

We should underline that the propagation of shear and rotation strain waves presents the mutual interaction of shears and rotation strains; such a linear propagation is visualized in Fig. 1.
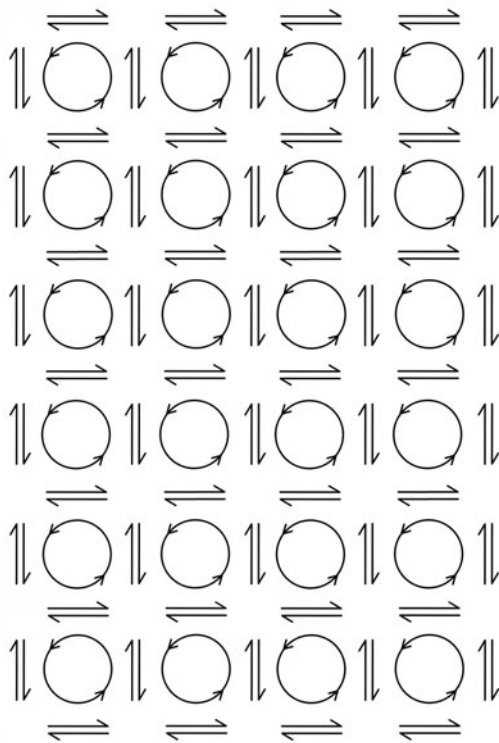
Fig. 1. Propagation of the shear and rotation strains, a linear case.
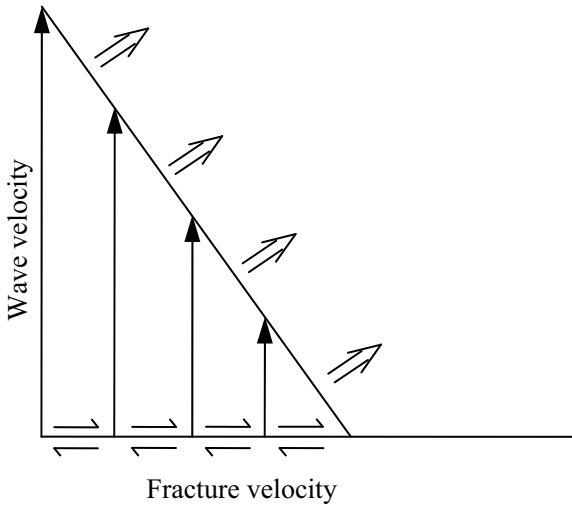
Fig. 2. Sketch showing how the fracture velocity and wave velocity form the propagation front velocity (double arrows) in the near-fracture region.

In a near-fracture region we can assume that a fracture progress at the source induces the waves propagating much faster than the fracture velocity; in this way, a propagation velocity front becomes created (Fig. 2). Thus, a fracture progress and waves propagation lead together to the front propagation; its velocity depends on the velocities of both, fracture and wave. This remark relates to the problems of the fault-plane solutions.

Finally, we should come to the fracture-related equations based on the release-rebound processes applied to the molecular motions: the molecular displacements and simple molecular rotation. The radiated rotation and shear strains can influence, however, any local molecular transport behaviour.

In relation to the real propagation of the shear and rotation strain waves, we may present a behaviour of the molecular transport; its propagation is a consequence of the propagation of shear and rotation strain waves. In this sense, we may define the molecular displacement transport, $w_s = \partial \upsilon_s / \partial x_s$.

In this way we may consider the molecular motions appearing in the plane $z = \text{const}$ and thus we may consider the 2D case with motions on plane (r, φ). The related molecular transport motions, along the propagation plane, relate to the (r, φ) system:

$$w_n \rightarrow \left(w_r, w_\varphi \right) . \tag{8a}$$

In this case we have assumed $w_z = 0$. The rotation transport molecular vector, $\ddot{\omega}_s$, may interact with these fields. Thus, to this local molecular transport components we add the molecular rotation transport with the component, $\ddot{\omega}_z$, oriented along the z-direction:

$$\ddot{\omega}_s \to \ddot{\omega}_z \quad \text{and} \quad \left(\ddot{\omega}_r = 0, \ \ddot{\omega}_\varphi = 0\right) . \tag{8b}$$

This molecular shear transport and molecular rotation transport should be subjected to local release-rebound fracture equations (*cf.*, Eq. 7a):

$$\varepsilon_{spq} \frac{\partial}{\partial x_p} \ddot{\omega}_q - \frac{\tilde{\kappa}\partial}{c\,\partial t} w_s = 0 , \qquad \varepsilon_{spq} \frac{\partial}{\partial x_p} w_q + \frac{\tilde{\kappa}\,\partial}{c\,\partial t} \ddot{\omega}_s = 0 . \tag{9a}$$

To assure the 4D invariance we have put, instead of the molecular velocity, v, the invariant velocity, c, and some material constant, $\tilde{\kappa}$: $v = c / \tilde{\kappa}$.

We include here the molecular rotation transport (Eq. 8b) and in this way the Eqs. 9a become invariant.

For the considered case, related to the molecular motions, we obtain from Eqs. 9a:

$$\frac{\partial \ddot{\omega}_z}{r\partial\varphi} - \frac{\partial w_r}{v\partial t} = 0 , \quad \frac{\partial \ddot{\omega}_z}{\partial r} + \frac{\partial w_\varphi}{v\partial t} = 0 , \quad \frac{\partial w_\varphi}{\partial r} - \frac{\partial \ddot{\omega}_z}{v\partial t} = 0 . \tag{9b}$$

or

$$\frac{\partial}{\partial t}\frac{\partial \ddot{\omega}_z}{r\partial\varphi} - \frac{\partial^2 w_r}{v\partial t^2} = 0 , \quad \frac{\partial}{\partial t}\frac{\partial \ddot{\omega}_z}{\partial r} + \frac{\partial^2 w_\varphi}{v\partial t^2} = 0 , \quad \frac{\partial}{\partial t}\frac{\partial w_\varphi}{\partial r} - \frac{\partial^2 \ddot{\omega}_z}{v\partial t^2} = 0 . \tag{9c}$$

This, quite new system of relations for w_r, w_φ, and $\ddot{\omega}_z$, leads to the equation for the molecular rotation:

$$\frac{\partial^2 \ddot{\omega}_z}{\partial r^2} + \frac{\partial^2 \ddot{\omega}_z}{v^2 \partial t^2} = 0 . \tag{10a}$$

The related source fracture processes may occur with some phase delay, *e.g.*, $v = iv_0$; we obtain:

$$\frac{\partial^2 \ddot{\omega}_z}{\partial r^2} - \frac{\partial^2 \ddot{\omega}_z}{v_0^2 \partial t^2} = 0 . \tag{10b}$$

Similarly we may consider also the radial wave propagation (*cf.*, Fig. 3). In this case, we should consider the local molecular system (r, φ) perpendicular to the R radial direction; for the related local shear and rotation strains we may add the presented transport:
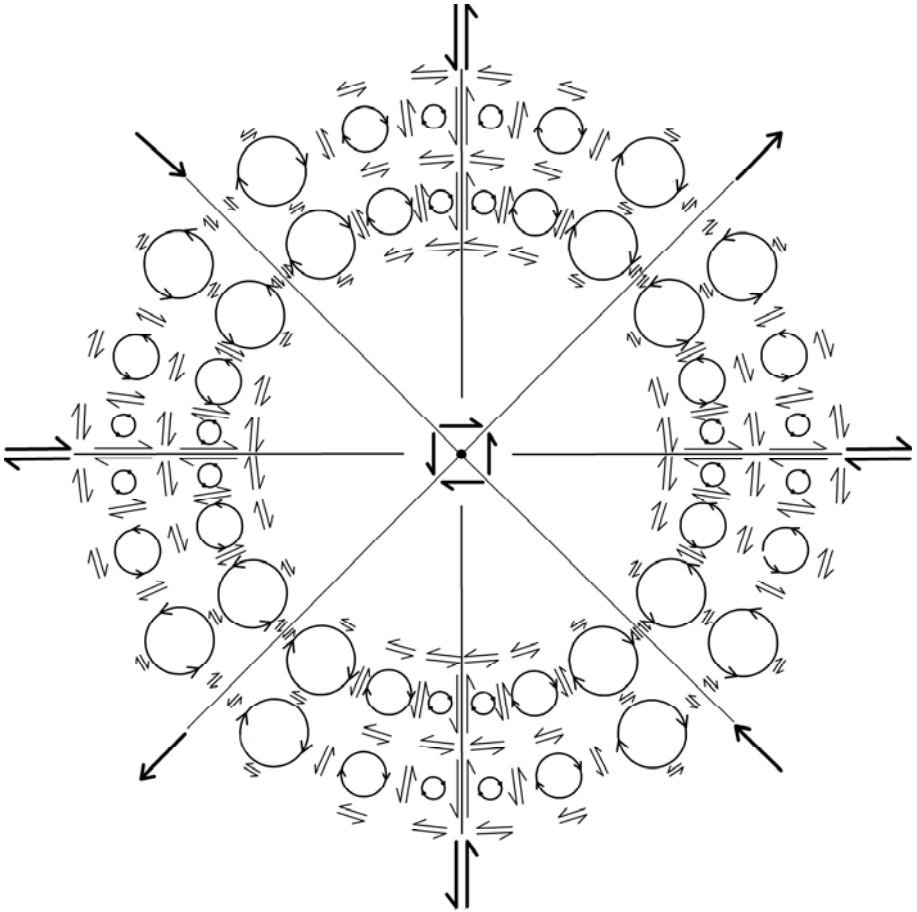
Fig. 3. Radial wave propagation for the shear and rotation strains (the rotation strains decrease near the shear axes to zero while the shear strains decrease between those axes).

$$w_n \to \left(w_\varphi, w_r\right) \quad . \tag{11a}$$

The molecular radial rotation component, $\ddot{\omega}_z$, and these molecular transport components, w_r and w_φ, should be subjected to release-rebound fracture equations with specific velocity-like term v (*cf.*, Eqs. 9a)

$$\varepsilon_{spq} \frac{\partial}{\partial x_p} \ddot{\omega}_q - \frac{\partial}{v \, \partial t} w_s = 0 \, , \qquad \varepsilon_{spq} \frac{\partial}{\partial x_p} w_q + \frac{\partial}{v \, \partial t} \ddot{\omega}_s = 0 \, , \tag{11b}$$

where

$$\varepsilon_{spq}\frac{\partial}{\partial x_p}\ddot{\omega}_q \to \left\{0,\frac{\partial\ddot{\omega}_R}{\partial z},-\frac{\partial\ddot{\omega}_R}{R\partial\varphi}\right\} \quad \text{at} \quad \ddot{\omega}_\varphi = 0 \ , \ \ddot{\omega}_z = 0 \ ,$$

$$\varepsilon_{spq}\frac{\partial}{\partial x_p}w_q \to \left\{\frac{\partial w_z}{R\partial\varphi}-\frac{\partial w_\varphi}{\partial z},-\frac{\partial w_z}{\partial R},\frac{\partial w_\varphi}{\partial R}\right\} \quad \text{at} \quad w_R = 0 \ . \tag{11c}$$

We obtain three relations for w_φ, w_z, and $\ddot{\omega}_R$:

$$\frac{\partial\ddot{\omega}_R}{\partial z}-\frac{\partial w_\varphi}{v\partial t}=0 \ , \qquad \frac{\partial\ddot{\omega}_R}{r\partial\varphi}+\frac{\partial w_z}{v\partial t}=0 \ , \tag{12a}$$

and

$$\frac{\partial w_z}{R\partial\varphi}-\frac{\partial w_\varphi}{\partial z}+\frac{\partial\ddot{\omega}_R}{v\,\partial t}=0 \ , \quad \text{for a given } R \text{ value} \tag{12b}$$

We have obtained finally a quite new system of relations for w_φ, w_z, and $\ddot{\omega}_R$, with a special velocity-like term v; eliminating from these relations the terms related to molecular rotations we obtain the following equations for the molecular displacements:

$$\frac{\partial^2 w_\varphi}{\partial z^2}-\frac{\partial^2 w_z}{R\partial\varphi\partial z}-\frac{\partial^2 w_\varphi}{v^2\partial t^2}=0 \ , \qquad \frac{\partial^2 w_z}{R^2\partial\varphi^2}-\frac{\partial^2 w_\varphi}{R\partial\varphi\partial z}-\frac{\partial^2 w_z}{v^2\partial t^2}=0 \ , \tag{13a}$$

and for the molecular rotation (molecular vortex structure):

$$\frac{\partial^2\ddot{\omega}_R}{R^2\partial\varphi^2}+\frac{\partial^2\ddot{\omega}_R}{\partial z^2}-\frac{\partial^2\ddot{\omega}_R}{v^2\,\partial t^2}=0 \ . \tag{13b}$$

These last Eqs. 13a, b are related to the fracture processes at the source with the release-rebound processes associated with the molecular fields.

According to K.P. Teisseyre (private communication 2012) we may relate this motion to a possible vortex structure in solids, defined as the molecular vorticity:

$$\varsigma_k = \varepsilon_{kpi}\frac{\partial\upsilon_i}{\partial x_p} \ \to \ \Xi_k = \varepsilon_{kpi}\frac{\partial w_i}{\partial x_p} \ , \tag{14a}$$

where for the radial propagation we have:

$$\Xi_R = \frac{\partial w_z}{R\partial\varphi}-\frac{\partial w_\varphi}{\partial z} \ , \quad \Xi_\varphi = -\frac{\partial w_z}{\partial R} \ , \quad \Xi_z = \frac{\partial w_\varphi}{\partial R} \ . \tag{14b}$$

Such a molecular vorticity we may be related to a gradient of pressure appearing in the P-wave propagation

$$\varepsilon_{spq}\frac{\partial}{\partial x_p}\Xi_q - \frac{\partial}{V^P\partial t}\frac{\partial p}{\partial x_s} = 0 \ . \tag{15a}$$

For the radial wave propagation, under the conditions $\partial p/\partial\varphi = 0$ and $\partial p/\partial z = 0$, we obtain:

$$\frac{\partial\Xi_z}{R\partial\varphi} - \frac{\partial\Xi_\varphi}{\partial z} - \frac{\partial}{V^P\partial t}\frac{\partial p}{\partial R} = 0 \ , \quad \frac{\partial}{\partial z}\Xi_r - \frac{\partial}{\partial R}\Xi_z = 0 \ , \quad \frac{\partial}{\partial R}\Xi_\varphi - \frac{\partial}{R\partial\varphi}\Xi_R = 0 \ . \tag{15b}$$

A similar problem may be applied to the plane waves along the Z-axis:

$$\varepsilon_{spq}\frac{\partial}{\partial x_p}\Xi_q - \frac{\partial}{V^P\partial t}\frac{\partial p}{\partial x_s} = 0 \ , \qquad \frac{\partial}{\partial r}\Xi_\varphi - \frac{\partial}{r\partial\varphi}\Xi_r - \frac{\partial}{V^P\partial t}\frac{\partial p}{\partial Z} = 0 \ , \tag{16a}$$

and under the condition $\partial p/\partial r = 0$ and $\partial p/\partial\varphi = 0$ we obtain

$$\frac{\partial}{r\partial\varphi}\Xi_Z - \frac{\partial}{\partial Z}\Xi_\varphi = 0 \ , \qquad \frac{\partial}{\partial Z}\Xi_r - \frac{\partial}{\partial r}\Xi_Z = 0 \ . \tag{16b}$$

In such a way, in a solid continuum there might appear the molecular rotations with the pressure propagation waves and related events. In this way, it is possible to explain an abundance of the rotation motions which occur together and just after the P-wave arrivals.

Our considerations indicate that in the presented system of the shear and rotation wave propagation we have the correlations regarding the space and time derivatives; an interaction between these fields could be direct or shifted by $\pi/2$. We should note that the asymmetric continuum theory includes a transition to the states close to fracture processes and, even further, to the fluid state. These fracture processes confirm the wave synchronization inside the molecular fields and the invariant 4D release-rebound system.

3. CONCLUSIONS

The presented 4D space-time system has been presented in a form of the rotation and shear strain vectors; their combined propagation is explained by the release-rebound interaction processes. At a seismic source, the molecular transport motions become transformed creating the real transport displacements. However, we cannot directly describe a fracture process in the frame of a continuum theory; we may only state that such a fracture is caused by an interaction of the real shear and rotation strains and related molecular transport displacements.

Following the author's theory, the molecular transport motions may exist in a whole space, although we do not assume the independent propagation of the molecular fields; thus, the presented considerations may help us to understand an influence of these molecular wave motions on the appearance of another fracture, even at a remote distance. Of course, any molecular propagation is only apparent: it follows from the real propagation of the strains in solids or real transport in fluids (see also Teisseyre 2013).

R e f e r e n c e s

Eshelby, J.D., F.C. Frank, and F.R.N. Nabarro (1951), The equilibrium of linear arrays of dislocations, *Philos. Mag. Ser. 7* **42**, 327, 351-364, DOI: 10.1080/14786445108561060.

Kossecka, E., and R. DeWitt (1977), Disclination kinematic, *Arch. Mech.* **29**, 633-651.

Majewski, E., and R. Teisseyre (2013), Thermodynamics with rotation motions: fragmentation and slip, *Acta Geophys.* **61**, 2, 281-310, DOI: 10.2478/s11600-012-0072-2.

Rybicki, K. (1986), Dislocations and their geophysical applications. **In:** R. Teisseyre (ed.), *Continuum Theories in Solid Earth Physics*, Elsevier–PWN, Warszawa, 18-186.

Teisseyre, R. (1985), Creep-flow and earthquake rebound: System of the internal stress evolution, *Acta Geophys. Pol.* **33**, 1, 11-23.

Teisseyre, R. (2011), Why rotation seismology: Confrontation between classic and asymmetric theories, *Bull. Seismol. Soc. Am.* **101**, 4, 1683-1691, DOI: 10.1785/0120100078.

Teisseyre, R. (2013), Molecular transport in fracture processes, *Acta Geophys.* **61**, 1, 18-25, DOI: 10.2478/s11600-012-0057-1.

Teisseyre, R., and M. Górski (2012), Induced strains and defect continuum theory: Internal reorganization of load, *Acta Geophys.* **60**, 1, 24-42, DOI: 10.2478/s11600-011-0046-9.

Varotsos, P.A. (2007), Comparison of models that interconnect point defect parameters in solids with bulk properties, *J. Appl. Phys.* **101**, 12, 123503, DOI: 10.1063/1.2745359.

Induction Sounding of the Earth's Mantle at a New Russian Geophysical Observatory

John MOILANEN and Pavel Yu. PUSHKAREV

Lomonosov Moscow State University, Moscow, Russia
e-mails: moilanen@mail.ru, pavel_pushkarev@list.ru

A b s t r a c t

Deep magnetotelluric (MT) sounding data were collected and processed in the western part of the East European Craton (EEC). The MT sounding results correspond well with impedances obtained by magnetovariation (MV) sounding on the new geophysical observatory situated not far from the western border of Russia. Inversion based on combined data of both induction soundings let us evaluate geoelectrical structure of the Earth's crust and upper and mid-mantle at depths up to 2000 km, taking into account the harmonics of 11-year variations. Results obtained by different authors and methods are compared with similar investigations on the EEC such as international projects CEMES in Central Europe and BEAR in Fennoscandia.

Key words: mantle, geoelectrical structure, East European Craton.

1. INTRODUCTION

By the late 2011, 35 European geomagnetic observatories have been registered in the International Real Time Magnetic Observatory Network (INTERMAGNET). Only one of these, Borok (58°04′ N, 38°14′ E), is operating in the European part of Russia (Anisimov *et al.* 2008). It is evident that the development of geophysical observatories in Russia is important for the international community to investigate the Earth's crust and mantle, for ex-

ample, by monitoring the physical fields. The Aleksandrovka Observatory belongs to the geophysical base of the Moscow University (250 km to the south-west of Moscow) which has been under construction since 2006. The place is poorly populated; there are no other settlements within a radius of 8 km and the distance to the nearest DC electrified railway is over 50 km, thus ensuring a low level of industrial noise. The settlement is electrified and the base is provided with diesel-generator emergency electric power. Internet communication is supported by two-way (GPS/GLONASS) satellite systems. The non-magnetic pavilion construction has been completed (54°53′79″ N and 35°00′87″ E) and geophysical data have been collected since May 2011. The purpose of the Observatory is to measure the full vector of the Earth's magnetic field (LEMI 025 fluxgate magnetometer; Korepanov *et al.* 1998), and to carry out uninterrupted recording of variations in horizontal components of the electric field (Shustov *et al.* 2012). Seismological three-component measurement stations were also installed. We hope to install proton magnetometer for absolute values of magnetic field measurement in the nearest future.

2. THEORETICAL BASE

The impedance is the basis of induction soundings. It was introduced by Leontovich in Russia at the beginning of 1930s (Rytov 1940). The strict theory of radio-waves spreading in the medium was developed by Rytov, who was Leontovich's follower (see Senior and Volakis 1995). Induction soundings were put into practice using equations for the first term of series. They were obtained for the boundary between isolated and conductive media. The impedance was considered as a scalar, or strictly speaking the functional of conductivity distribution via skin depth (Rikitake 1948, Tikhonov 1950, Cagniard 1953). Today, the impedance is a matrix that is formed by anisotropy or heterogeneity of the medium. It was introduced by Berdichevsky and Cantwell (see Berdichevsky and Dmitriev 2008). Equations for magnetotelluric (MT) and generalized magnetovariation (GMV) soundings are as follows:

$$
\begin{aligned}
E_x(\omega) &= Z_{xx}(\omega)B_x(\omega) + Z_{xy}(\omega)B_y(\omega), \\
E_y(\omega) &= Z_{yx}(\omega)B_x(\omega) + Z_{yy}(\omega)B_y(\omega), \\
B_x(\omega) &= Y_{xx}(\omega)E_x(\omega) + Y_{xy}(\omega)E_y(\omega), \\
B_y(\omega) &= Y_{yx}(\omega)E_x(\omega) + Y_{yy}(\omega)E_y(\omega),
\end{aligned}
\tag{1}
$$

$$
B_z(\omega) \approx C(\omega,r)\operatorname{div}B_\tau(\omega) + \operatorname{grad}C(\omega,r)B_\tau(\omega) .
\tag{1a}
$$

Here $C(\omega,r)$ is a response function in GMV sounding method, where r is a radius vector. It is transformed into apparent resistivity in SI with the for-

mula: $\rho^* = i\omega\mu_0 C^2$ where i is the imaginary unit, $\mu_0 = 4\pi 10^{-7}$ [Henry/m] the permeability of the free space, $\omega = 2\pi/T$ the circular frequency and T-the period. The laconic Eq. 1a was formulated by Guglielmi and Gokhberg (1987), and Schmucker (2003). A more precise theory was presented by Shuman and Kulik (2002), and Shuman (2007) (see Semenov and Shuman 2010). The complex Fourier amplitudes for the corresponding components of MT field, B_r, E_x, E_y, B_x, B_y, B_z, are connected through matrices of impedance $\|Z\|$ and admittance $\|Y\|$:

$$Z = \begin{pmatrix} Z_{xx} & Z_{xy} \\ Z_{yx} & Z_{yy} \end{pmatrix}, \quad Y = \begin{pmatrix} Y_{xx} & Y_{xy} \\ Y_{yx} & Y_{yy} \end{pmatrix}. \tag{2}$$

For an isotropic horizontal uniform medium, grad $C = 0$ (components of this vector on the x and y axis are different from tippers). The Eq. 1a can be transformed to the equality $B_z(\omega) \approx C(\omega, r)$ div $B_r(\omega)$, which was introduced for MV soundings by Berdichevsky et $al.$ (1969) and Schmucker (1970). It can be simplified for Dst variations, because the field induced by ring currents in the magnetosphere is linearly polarized in the Earth (Olsen 1998). This field is sufficient for defining the Earth response function (Banks 1969):

$$i\omega B_r(\omega) = C(\omega, r)\frac{2B_\theta(\omega)}{R\,\mu_0 tg(\theta_0)}. \tag{3}$$

Here R is the Earth's radius, θ_0 the geomagnetic colatitude of observation place, and B_r, B_θ the spectra of radial and colatitude components of the magnetic field. Equation 3 is valid in geomagnetic coordinate system on the sphere.

If all the elements of impedance matrix are found and $Z_{xx} \equiv Z_{yy} = 0$, $Z_{xy} = Z_{yx} = Z$ regardless of the direction (uncommon case in practice), the conductive medium can be presented as uniform, isotropic halfspace with one scalar impedance $Z(\omega)$. It is transformed to the scalar resistivity as follows

$$\rho = \frac{\mu_0 Z^2}{\omega}. \tag{4}$$

In a more common case, $Z_{xy} \neq Z_{yx}$ and additional impedances are nonzero. Then we have a non isotropic medium. In the best case it can be described by tangential anisotropic uniform halfspace. Its resistivity is defined by plate tensor. Components of apparent resistivity tensor can be found by the following expressions (Eq. 5), which were obtained theoretically from Maxwell equations by Reilly (1979) (see Weckmann et $al.$ 2003) and Semenov (2000):

$$\rho_{xx} = \mu_0 \left(Z_{xy}^2 - Z_{xx} Z_{yy} \right) \big/ \omega, \quad \rho_{xy} = \mu_0 Z_{xx} \left(Z_{yx} - Z_{xy} \right) \big/ \omega,$$
$$\rho_{yx} = \mu_0 Z_{yy} \left(Z_{xy} - Z_{yx} \right) \big/ \omega, \quad \rho_{yy} = \mu_0 \left(Z_{yx}^2 - Z_{xx} Z_{yy} \right) \big/ \omega. \tag{5}$$

Effective apparent resistivity is used for heterogeneous media. It is usually defined from polar diagrams of impedance or admittance matrixes. However, you can quite easily localize the heterogeneity place by examining the effective apparent resistivity. It can be calculated by using the generally accepted formula, without guidance of possible heterogeneity degree:

$$\rho_{\text{eff}} = \mu_0 \left(Z_{xx} Z_{yy} - Z_{xy} Z_{yx} \right) \big/ \omega. \tag{6}$$

This approach is well applicable during anomalous conductivity zones searching in exploration geophysics. But it is inadmissible during deep Earth's research because of the lack of proven data (drilling is well proven data for exploration geophysics).

Model with horizontal anisotropy is worth to be considered for conductive structure definition of the Earth's mantle. Its anomalous zones are local. It can be modeled by stratified, anisotropic structure in regional scale.

3.   DATA PROCESSING

Processing of MT data leads to defining the elements of MT matrix (Berdichevsky and Dmitriev 2008), *i.e.*, two unknowns from one of Eqs. 1.

In the GMV sounding case (Eq. 1a) the number of unknowns increases to three at least: scalar impedance and two tippers. Such a decision can be evaluated just under the assumption that the number of process realizations is big enough, for example, in the context of the theory of stochastic processes.

The obtained values have a stochastic character and so are characterized by confidence intervals under the assumption that displacement errors (bias), which depend on noises, are small. This is indicated by the corresponding coherencies (Reddy and Rankin 1974). There is a method to avoid bias by correlation of all observed data with similar data in a remote point (Gamble *et al.* 1979).

There are a lot of algorithms for defining $\|Z\|$ and $\|Y\|$ matrices by records of MT field. For example, one method is based on narrow band digital filtration (Narsky 1994, Berdichevsky and Dmitriev 2002), another one is based on classical method of spectral analysis (Sims *et al.* 1971), and a third one analyzes data in time domain (Nowożyński 2004). There is an interesting fact. The theory is constructed for a single frequency but impedances often characterize a certain band. It is also a set of realizations but for neighbouring frequencies.
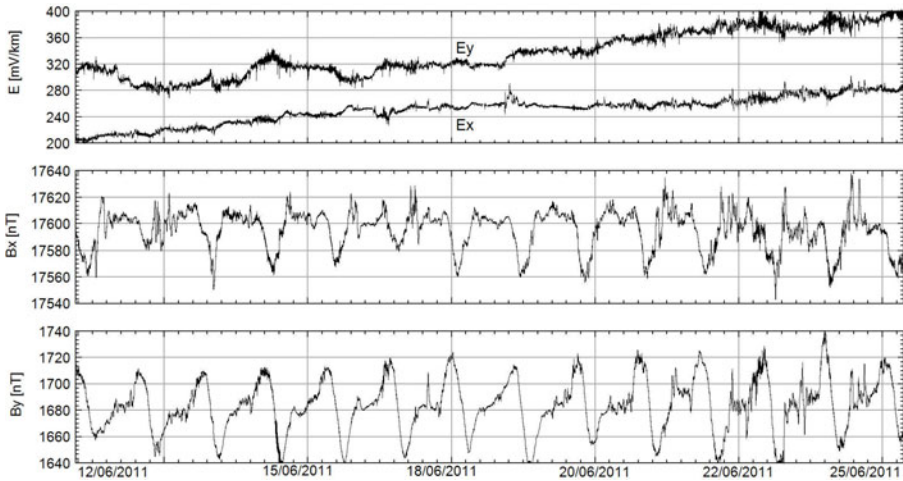
Fig. 1. An example of 15-day interval of MT field registered on the Aleksandrovka Observatory in June 2011. B_x and B_y are the components of magnetic induction vector, x the direction to the geographical north, y to the east, E_x and E_y are the components of electric field; the length of receiving lines is 100 m.
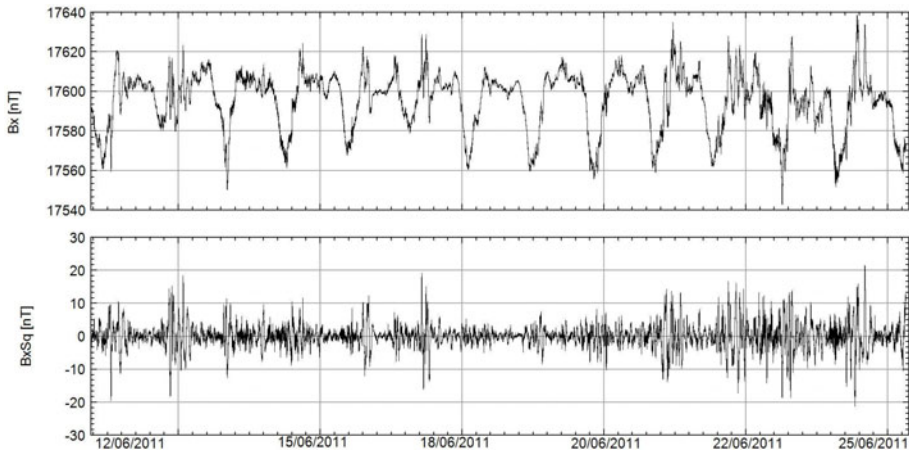
Fig. 2. Removing diurnal variations on the example of 15-days record on Alexandrovka base in June 2011 (B_x – component of the magnetic field, B_{xSq} – filtration result).

The MT data chosen for analysis were obtained in summer 2011. Some part of these data is presented in Fig. 1. While defining impedances it is necessary to switch from Eq. 1 for monochromatic electromagnetic wave to a continuous distribution of spectral frequency. Spectral distributions of field

components let the required functions values with their casual errors and bias (it is often higher than casual errors). The coherence defines reliability of implementation of Eqs. 1, 1a for real data (Semenov 2000).

For a long period, the data processing was performed by using the Petrischev/Tkachev program (Method 1, Semenov 1985). Processing scheme of MT-data in this method includes removing short-term spikes with a nonlinear filter (Naudy and Dreyer 1968), removing diurnal harmonics of observed field (Fig. 2) (Parkinson 1983) and, finally, selection of data with high coherency, *i.e.*, less distorted sections with noises.

It is worth noting that the main directions for different periods are different. Three sections were selected. The main direction was selected to be 150 degrees for the first section, $T \in [30, 2500]$ s; 170 degrees for the second section, $T \in [2500, 8000]$ s; and 150 degrees (the same as at the first one) for the third section, $T \in [8000, 25\,000]$ s (Fig. 3). The second processing stage includes analyzing the whole record with the selected main direction. Different displacement windows are chosen. This window is charged with the length of data for the fast Fourier transform.
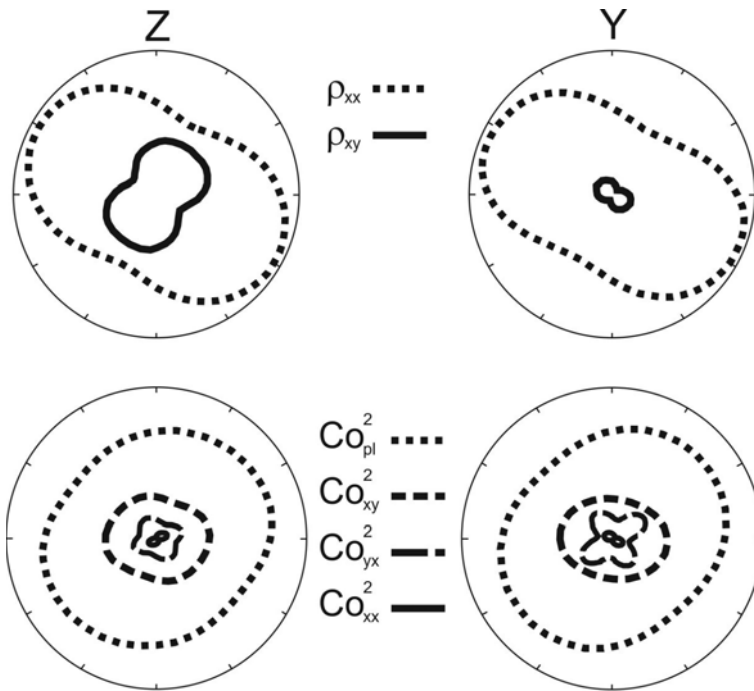
Fig. 3. Polar diagrams of primary and additional components of apparent resistivity tensor [Ohm·m] (upper image) and square coherencies: Co_{pl}^2 – plural, Co_{xy}^2, Co_{yx}^2 – singular, Co_{xx}^2 – input signals (bottom), received for impedance Z (left) and admittance Y (right) estimations for $T = 10\,000$ s.

The observed data were processed by different algorithms. The first method was used for processing just on the Aleksandrovka Observatory. Varentsov's algorithm (Method 2) was used in multi reference scheme. It is formed on robust averaging of estimates for several base points. Also, there were some boundary conditions on horizontal MV response changes between investigation point and remote bases (Varentsov 2007).

Additional MT and Audio-MT three-day's data were added to the joint process by Varentsov's and Larsen's algorithms (Larsen *et al.* 1996) (Method 3).
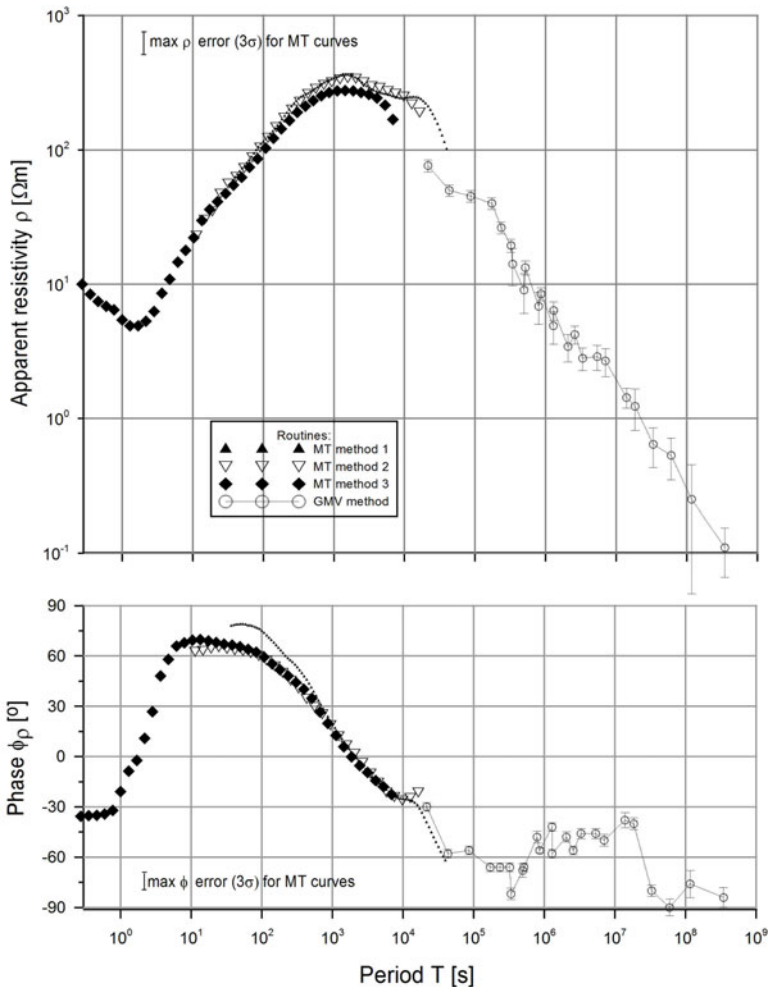
Fig. 4. Complex apparent resistivities of central part of the EEC estimated by three independent MT routines at the Aleksandrovka Observatory and GMV method at the Moscow Observatory.

4. COMPARISON OF PROCESSING RESULTS

Method 1 does not take into account frequency characteristics of the equipment. Results obtained by this algorithm differ from the Varentsov's and Larsen's algorithms in low period area (Fig. 4). It is necessary to consider frequency characteristics of the equipment for the period range lower than 600 s ($T < 6000$ s). Reliability of acquired data was improved using several base stations and robust averaging procedure. This is well seen especially in the long period area ($T > 6000$ s). The results of different processing are well corresponding up to the period of $T \approx 3$ hours ($T = 10^4$ s). Then serious discrepancies occur (Semenov and Shuman 2010, Shimizu *et al.* 2011). Three kinds of source fields are dominant at the period range 10^4-10^5 s daily oscillations (*Sq* variations) connected with the Earth's rotation, bays caused by polar cusp currents, and *Dst* variations caused by the magnetospheric ring current. The bays can be considered as a plane wave (Vanyan *et al.* 2002) with the vertical magnetic component in the middle latitudes, while the *Dst* variations contain the stable magnetic field which is collinear with geomagnetic axis (Banks 1981, Fujii and Shultz 2002). However, the MT soundings can be replaced by GMV soundings for the periods of 3-30 hours (Semenov *et al.* 2011). The data of Moscow Observatory (CEMES project, Semenov *et al.* 2008) were used for those purposes (Fig. 4).

5. THE INVERSION

First of all, the phase inversion of the impedance was accomplished by OCCAM algorithm (Constable *et al.* 1987). Thus, the module of apparent resistivity tensor was corrected for surface heterogeneity obtained by MT data (Fig. 5).

   Inversions by the well-known OCCAM and D+ algorithms were held for corrected responses of MT and GMV methods. D+ let us create geoelectrical horizontal layered media. Conductive layers approximate by thin layers with finite conductance (Parker 1980).

   Sediments, crust, asthenosphere, and mid-mantle layer were picked out during 1D interpretation (Fig. 6a). The sediments of Moscow syncline have a thickness of 800 m.

   The upper boundary of asthenosphere was picked out at a depth of 280 km. The upper boundary of the mid-mantle layer is at a depth of 660 km. The D+ results are quite representative. The upper boundary of the asthenosphere has a depth of 230 km. The mid-mantle layer has its edges at depths of 670 and 900 km. Additional conductive layer was picked out by the D+ algorithm. It has the conductance of 160 S at a depth of 65 km. We think this layer can be the appearance of the conductive layers of the lower crust. They should be shallower. We used previous investigations for evalua-
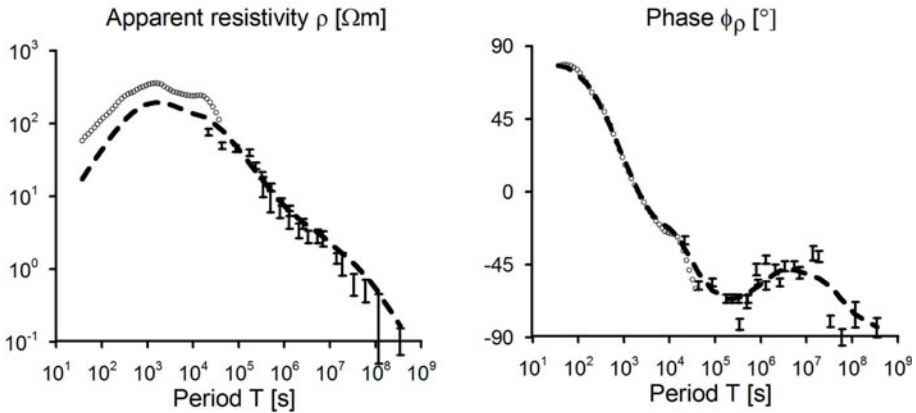
Fig. 5. The correction result obtained by OCCAM (dashed black) algorithm. Complex apparent resistivities estimated by the MT method for Aleksandrovka are shown as empty points and the MV method for the Moscow Observatory are shown as error bars.
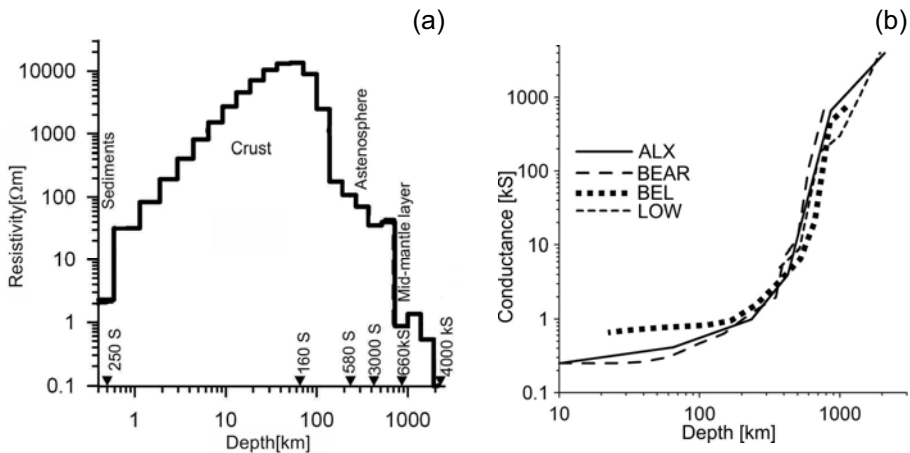
Fig. 6. Comparison of inversion results of MT and MV data for Aleksandrovka and Moscow Observatories; solid line is the OCCAM decision, arrows show D+ thin layers with their conductance (left panel) and conductance comparison obtained on different observatories. ALX – Aleksandrovka Observatory, BEAR – Fennoscandian experiment (Varentsov *et al.* 2002, Sokolova and Varentsov 2007), BEL – observatory in Poland, LOW – observatory in Sweden. Data from two last observatories were obtained during CEMES project (Semenov and Jozwiak 2006, Semenov *et al.* 2008).

tion of the obtained results. These were the BEAR project, based on data obtained in Fennoscandian region, and CEMES in the central part of Europe.

The comparison was carried out by the curves of integral conductance (Fig. 6b). The choice was made of the S curves. Each conductivity distribution corresponds with the one conductance value. So the uniqueness theorem for 1D inversion is proved only for infinite frequency range.

6. RESULTS AND DISCUSSION

Aleksandrovka is a geophysical observatory of Moscow University. Long-period registration of MT-field is held here. The obtained data were processed by several authors with different algorithms, such as the only observation point processing or using several base stations. The robust estimation was used. It has become a standard procedure during MT data processing. We have got close results for $T < 10^4$ s for all researching algorithms.

For longer periods it is hard to get truthful estimations due to several reasons. Estimation procedures of transfer operators for several base stations improve processing result. The obtained results correspond well with deep geoelectrical mantle properties. Moreover, they figure on a relative homogeneity of the mantle on the EEC. The sharp increase of mantle conductivity begins with depths of 300-400 km. We think it is connected with the asthenosphere.

References

Anisimov, S.V., E.M. Dmitriev, N.K. Sycheva, A.N. Sychev, V.P. Shcherbakov, and Y.K. Vinogradov (2008), Information technology in geomagnetic measurements perfomed in the Borok Geophysical Observatory, *Geofiz. Icc.* **9**, 3, 62-76 (in Russian).

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

Banks, R.J. (1981), Strategies for improved global electromagnetic response estimates, *J. Geomagn. Geoelectr.* **33**, 11, 569-585, DOI: 10.5636/jgg.33.569.

Berdichevsky, M.N., and V.I. Dmitriev (2002), *Magnetotellurics in the Context of the Theory of Ill-Posed Problems*, Investigations in Geophysics, No. 11, Society of Exploration Geophysicists, Tulsa, 215 pp.

Berdichevsky, M.N.,and V.I. Dmitriev (2008), *Models and Methods of Magnetotellurics*, Springer, Berlin, 563 pp.

Berdichevsky, M.N., L.L. Vanyan, and E.B. Fainberg (1969), The frequency sounding of the Earth by spherical analysis of electromagnetic variations, *Geomagn. Aeronomy* **9**, 2, 372-374 (in Russian).

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

Constable, S.C., R.L. Parker, and C.G. Constable (1987), Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data, *Geophysics* **52**, 3, 289-300, DOI: 10.1190/1.1442303.

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

Gamble, T.D., W.M. Goubau, and J. Clarke (1979), Magnetotellurics with a remote magnetic reference, *Geophysics* **44**, 1, 53-68, DOI: 10.1190/1.1440923.

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

Korepanov, V., A. Best, B. Bondaruk, H.-J. Linthe, J. Marianiuk, K. Pajunpää, L. Rakhlin, and J. Reda (1998), Experience of observatory practice with LEMI-004 magnetometers, *Rev. Geofis.* **48**, 31-40.

Larsen, J.C., R.L. Mackie, A. Manzella, A. Fiordelisi, and S. Rieven (1996), Robust smooth magnetotelluric transfer functions, *Geophys. J. Int.* **124**, 3, 801-819, DOI: 10.1111/j.1365-246X.1996.tb05639.x.

Narsky, N.V. (1994), Robust methods of processing magnetotelluric variation, Ph.D. Thesis, Moscow State University, Moscow, Russia, 130 pp. (in Russian).

Naudy, H., and H. Dreyer (1968), Essai de filtrage non-linéaire appliqué aux profils aéromagnétiques, *Geophys. Prospect.* **16**, 2, 171-178, DOI: 10.1111/j.1365-2478.1968.tb01969.x (in French).

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

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

Parker, R.L. (1980), The inverse problem of electromagnetic induction: Existence and construction of solutions based on incomplete data, *J. Geophys. Res.* **85**, B8, 4421-4428, DOI: 10.1029/JB085iB08p04421.

Parkinson, W.D. (1983), *Introduction to Geomagnetism*, Scottish Academic Press, Edinburgh, 433 pp.

Reddy, I.K., and D. Rankin (1974), Coherence functions for magnetotelluric analysis, *Geophysics* **39**, 3, 312-320, DOI: 10.1190/1.1440430.

Reilly, W.I. (1979), Anisotropy tensors in magnetotelluric application, Tech. Rep., Department of Scientific and Industrial Research, Wellington, New Zealand.

Rikitake, T. (1948), Notes on electromagnetic induction within the Earth, *Bull. Earthq. Res. Inst.* **24**, 1/4, 1-9.

Rytov, S.E. (1940), Calculation of the skin-effect by the perturbation method, *J. Exp. Theor. Phys.* **10**, 2, 180-189 (in Russian).

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

Schmucker, U. (2003), Horizontal spatial gradient sounding and geomagnetic depth sounding in the period range of daily variations. **In:** A. Hördt and J.B. Stoll (eds.), *Protokoll über das Kolloquium Elektromagnetische Tiefenforschung, 29.09.-3.10.2003, Königstein, Deutschland*, 306-317.

Semenov, V.Yu. (1985), *Processing of Magnetotelluric Sounding Data*, Nedra, Moscow, 133 pp. (in Russian).

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

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

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

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

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

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

Shimizu, H., A. Yoneda, K. Baba, H. Utada, and N.A. Palshin (2011), Sq effect on the electromagnetic response functions in the period range between 10^4 and 10^5 s, *Geophys. J. Int.* **186**, 1, 193-206, DOI: 10.1111/j.1365-246X.2011. 05036.x.

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

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

Shustov, N.L., V.A. Kulikov, E.V. Moilanen, A.Yu. Polenov, P.Yu. Pushkarev, V.K. Khmelevskoi, and A.G. Yakovlev (2012), The creation of a geophysical observatory at the Aleksandrovka research base of the Geological Faculty of Moscow State University in the Kaluga Region, *Moscow Univ. Geol. Bull.* **67**, 4, 253-258, DOI: 10.3103/S0145875212040096.

Sims, W.E., F.X. Bostick, and H.W. Smith (1971), The estimation of magnetotelluric impedance tensor elements from measured data, *Geophysics* **36**, 5, 938-942, DOI: 10.1190/1.1440225.

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

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

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

Varentsov, I.M. (2007), Arrays of simultaneous electromagnetic soundings: design, data processing and analysis. **In:** V.V. Spichak (ed.), *Electromagnetic Sounding of the Earth's Interior*, Methods in Geochemistry and Geophysics, Vol. 40, Elsevier, Amsterdam, 259-275.

Varentsov, I.M., M. Engels, T. Korja, M.Yu. Smirnov, and the BEAR Working Group (2002), A generalised geoelectric model of Fennoscandia: A challenging database for long-period 3D modelling studies within the Baltic Electromagnetic Array Research (BEAR) Project, *Izv. – Phys. Solid Earth* **38**, 11, 855-896.

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

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

Effect of Time on Dynamic Shear Modulus of Selected Cohesive Soil of One Section of Express Way No. S2 in Warsaw

Wojciech SAS, Katarzyna GABRYŚ, and Alojzy SZYMAŃSKI

Faculty of Civil- and Environmental Engineering,
Warsaw University of Life Sciences (SGGW), Warsaw, Poland;
e-mails: wojciech_sas@sggw.pl,
katarzyna_gabrys@sggw.pl (corresponding author), ajozy_szymanski@sggw.pl

A b s t r a c t

Several researches published comprehensive reports on dynamic soil properties of cohesive soils, in which many of them outlined, *i.e.*, key factors affecting the dynamic shear modulus. For cohesive soils, the modulus at small strains ($\gamma < 10^{-3}$ %) is, first of all, a function of void ratio and effective confining stress. For clays, however, secondary time effects and clay mineralogy (fabric and structure) also appear to be important. The influence of confinement of laboratory-prepared as well as naturally deposited clays consists in an increase of shear modulus logarithmically as a function of time. In this paper, the effect of duration of the various confining pressures on dynamic shear modulus (G) of selected cohesive soils from Warsaw area was evaluated. Shear modulus was determined on the basis of resonant column tests, at low and high shearing strain amplitudes. It is shown that the calculated shear modulus is time-dependent; during approximately first 1000 minutes of consolidation, the moduli increased by almost 50%. Moreover, it is characterized by two phases: an initial one results from primary consolidation and a second one, which occurs after the end of primary consolidation, herein about 16-17 hours, and is called "long-term time effect". This effect was found also for modulus at higher shearing strains ($\gamma > 10^{-3}$ %, *e.g.*, 3×10^{-3} %, 5×10^{-3} %, 8×10^{-3} %, 2×10^{-2} %).

Key words: dynamic shear modulus, time-dependency, cohesive soil.

1.   INTRODUCTION

The dynamic shear modulus is an important property of soil, especially from the viewpoint of evaluating wave propagation in soil and, on the basis of it, receiving access to dynamic response of embankments, pavements, foundations, *etc*. (Tezcan *et al.* 2009). Road subgrades or shallow foundations are continuously loaded cyclically by various machines, vehicles (cars, trains), and other factors (Khosravi *et al.* 2010, Sidorova and Voznesensky 2010). Earthquakes, pile driving, explosions also put the load to water-saturated or unsaturated, fine- and coarse-grained or cohesive soil deposits (Vucetic and Dobry 1991).

The dynamic shear modulus of soil in the range of strain amplitude less than 10^{-4} is referred to as the maximum or small-strain dynamic shear modulus, denoted by the symbol G_{max}. Many studies have been conducted to find and properly characterized factors which affect this parameter. The main elements that have an effect on G_{max} include: shear strain amplitude (γ), mean effective confining stress (σ'_m, p'), soil type and plasticity index (PI). Other factors which influence the small-strain dynamic shear modulus, but appear to be less important, include: frequency of loading, number of loading cycles, void ratio, degree of saturation, overconsolidation ratio, grain characteristics, and many others. Additionally, it has been suggested that through grouping the test data by geology some less essential component may be also taken into account (Zhang *et al.* 2005, Markowska-Lech *et al.* 2013).

One other factor, *i.e.*, duration of the confining pressure, is of fundamental importance in laboratory determination of shear modulus as well. The aspect of effect of time on soil stiffness is, unfortunately, often either misunderstood or neglected. This overlooking can have negative impact on the interpretation of the test data and results, which may lead to some failure in deriving maximum benefit from a research program (Anderson and Stokoe 1978).

In this paper, some key information about the effect of time, under fix pressure conditions, on shear modulus of selected cohesive soils from area of Warsaw is presented. Previous studies on this subject are briefly reviewed here. Testing procedures are shortly described as well. Finally, examples of the results from laboratory tests are given, which illustrate the significance of the effect of time on soil stiffness.

2.   SECONDARY TIME EFFECTS

Several investigators have shown that time-dependency is a significant parameter affecting the dynamic properties of soils (Zavoral 1990). The effect of confinement duration at the unchanging pressure conditions was first observed in 1961 (Richart 1962). In the next years, other researchers reported

the followings: during experiments at a constant confining pressure, shear modulus at the low strain amplitude (below 0.001%) increased with time of specimen confinement (*e.g.*, Lawrence 1965, Afifi and Richart 1973, Anderson and Woods 1976). These general observations led to more detailed studies of artificially prepared soil samples and the growth of their stiffness with time. The last tests at a long-lasting pressure on undisturbed specimens of cohesive and loose soils showed similarities. Time-dependent behaviour is a characteristic feature of natural soils too (Anderson and Stokoe 1978).

For most soils, the time-dependency at low strain level can be described by an initial phase, while the shear modulus changes instantly with the logarithm of time, as presented in Fig. 1. This early stage results from changes in the void ratio during the primary consolidation process. Therefore, this phase is related to primary consolidation. During the second part, the value of modulus rises nearly linearly with the logarithm of time, which is caused probably by strengthening of physical-chemical bonds in cohesive soils and by an increase in particle-particle contacts in cohesionless soils. The latter stage refers to the so-called "long-term time effect" that occurs only after the primary consolidation course has been completed.

There are two methods to explain the long-term time effect. The result of time-dependency can be, on the one hand, expressed in an absolute sense, *i.e.*, as a coefficient of shear modulus increase with time (I_G)
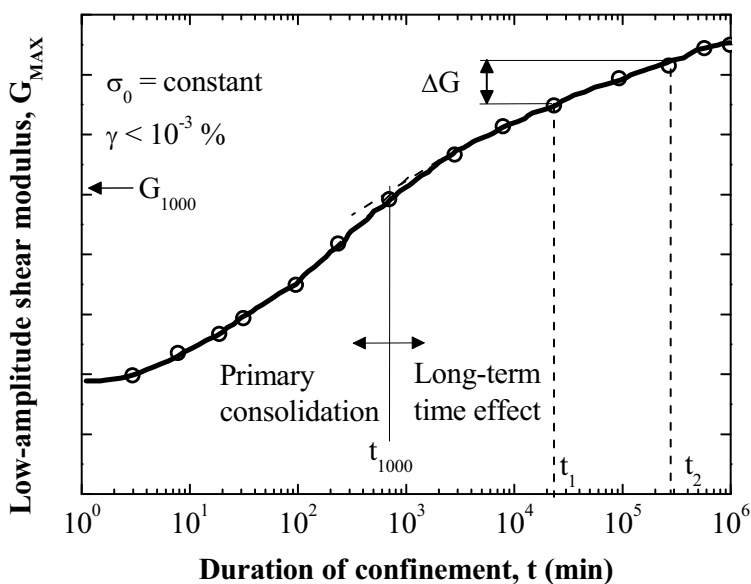
Fig. 1. Division of stages in low-amplitude shear modulus-time response (after Anderson and Stokoe 1978).

$$I_G = \frac{\Delta G}{\log_{10}\left(\dfrac{t_2}{t_1}\right)} \quad , \qquad (1)$$

where ΔG is the increase in low-amplitude shear modulus in logarithmic cycle of time, and t_1, t_2, are times after primary consolidation (Fig. 1.)

In a numerical manner, for one logarithmic cycle of time, I_G equals ΔG.

The other way of interpretation of the long-term time effect is, in relative terms, by using the normalized shear modulus with time (N_G) according to the formula

$$N_G = \frac{I_G}{G_{1000}} \quad , \qquad (2)$$

where G_{1000} is the modulus measured after 1000 minutes of constant confining pressure, which must be after completion of primary consolidation.

In general, parameter N_G becomes greater when the mean grain size decreases (Afifi and Richart 1973). However, more logical relationship between the rate of secondary modulus (N_G) and plasticity index (PI) was proposed by Kokusho *et al.* (1982). The plasticity index (PI) is a measure of the plasticity of a soil; it is the size of the range of water contents for which the soil exhibits plastic properties. PI is the difference between the liquid limit and the plastic limit (PI = LL-PL). PI reflects the chemical activity of clay minerals; therefore, with growing plasticity index there is an increase of N_G, amounting to as much as 25% per log cycle of time.

In most cases, the measured N_G value for clays ranges between 0.05 and 0.25. For sands, parameter N_G varies between 0.01 and 0.03 and rises when the soil gets finer, as presented by Mesri *et al.* (1990). Jamiolkowski and Manassero (1996) reported that silica sand samples reached the value of N_G from 0.001 to 0.03, sand samples with 50% of mica 0.039, whereas carbonate sand samples 0.05-0.12. On the basis of the experiments of Howie *et al.* (2002), the rate of increase of the shear modulus with time for very loose carbonate sands grows along with the stress level. Isotropic stress state arrived at a slower rate of increase in stiffness (Soga 2005).

3. LABORATORY TESTS

To study the time dependent increase in dynamic shear modulus, a set of resonant column tests was performed. The material investigated belongs to a natural cohesive soil formation in the Warsaw area, originally taken from the region of the express way No. S2 (around the road embankment No. WD-18). It is a clayey sand, clSa (Eurocode 7), with index properties listed in Table 1. The samples were taken from a depth of around 2.0 m selected carefully

Index parameters of specimens

Parameter	Value	
	Clayey sand (1)	Clayey sand (2)
w_L [%]	36.80	35.21
w_P [%]	13.76	11.43
I_P [%]	23.04	23.78
G_S [–]	2.68	2.68
w [%]	14.43	14.20
ρ [kg/m³]	2.12	2.16

Explanations: w_L is the liquid limit, w_P the plastic limit, I_P the plastic index, G_S the specific gravity, w the water content, and ρ is the mass density.

considering the uniformity of the soils structure, its physical properties, and its double-phase. Only one type of undisturbed soil samples was obtained, namely tube samples, with the inner diameter of the tube of about 80 mm. Tubes were carefully and gently pressed into the pre-drilled holes. All the samples were sealed and stored in a humidity room until needed.

The resonant column apparatus successfully used in this research was manufactured by British company GDS Instruments Ltd (GDS 2010). The device is presented in detail by Sas and Gabryś (2012), and Gabryś *et al*. (2013). Just to sum up briefly, the object of the resonant column technique is to vibrate the soil-top-cap system at first-mode resonance. This resonance is defined as the frequency at which the maximum top-cap motion is obtained during a sweep of frequencies, here mostly starting around 30 Hz. At the first-mode resonance, material in a cross-section at every elevation vibrates in phase with the top of the sample. Shear wave velocity, and then shear modulus are determined based on the system's constants, size, weight, and shape of the tested specimen (GDS 2010).

Table 2 summarized the authors' experiments during which the effect of long-term confinement on low-amplitude dynamic shear modulus was eval-uated. Time effect was recorded for clayey sand samples over a wide range of confining pressures, σ_0 from 335 up to 685 kPa, as well as a broad variety in shear moduli on low amplitude, $G_{\max}$ from 15 up to 175 MPa. The shear modulus for investigated samples was monitored at various times during the consolidation period. This is generally involved taking measurements well beyond the time required for the end of primary consolidation in order to find the time dependent increase at a constant effective pressure.

The tests conditions

Soil type	Specimen type	Mean effective stress [kPa]	Confining pressure [kPa]	Shear modulus on low amplitude [MPa]	I_G [MPa]	N_G [%]
Clayey sand (1)	undisturbed	45	335	15-30*	–	–
		90	380	30-55*	–	–
		135	425	40-90**	20.44	27.25
		180	470	80-95*	–	–
		225	515	80-110*	–	–
Clayey sand (2)	undisturbed	180	550	80-170**	6.81	4.72
		270	640	155-175**	30.51	21.83
		315	685	170-175**	6.42	5.14

*)measurements conducted during primary consolidation,
**)measurements conducted during and after the end of primary consolidation.

4. EXPERIMENTAL RESULTS

4.1 Low-amplitude test procedure

In time-dependent resonant column tests, dynamic shear modulus was defined at various times after application of different confining pressures. The recording time was usually: 1, 10, 20, 30, 45, 60, and 90 minutes after the selected pressure was employed, as well as just after the completion of primary consolidation. In this study, strains reached the level below 0.001%, and therefore it was a low-amplitude test. Figure 2 presents some examples of the results for the variation of shear modulus at changing mean effective stress. The duration of vibration at shearing strain levels of this kind is negligible (Anderson and Stokoe 1978). It should not be essential if the specimen is continuously or only seconds under vibration. The effect of time in Fig. 2 is noticeable for all mean effective stresses (p'). An increase in values of dynamic shear modulus is about 4-6%, at each time step when the measurements were performed. Generally, from the beginning till the end of primary consolidation, the rise in shear modulus may be as high as 50%. Figure 3 is also an illustration showing low-amplitude shear modulus growth in time at various stress level. On the basis of this diagram it can be seen in detail how the parameter G_{max} grows over time; we see at which point of time the increase in values of G_{max} was recorded to be the largest. In the first 10-20 minutes, the rise of G_{max} is quite rapid and relatively high in comparison with the next measurements; then this growth gets stabilized until the final phase of the test, when it achieves the biggest values, up to even 25 MPa for $p' = 90$ kPa and $p' = 225$ kPa.
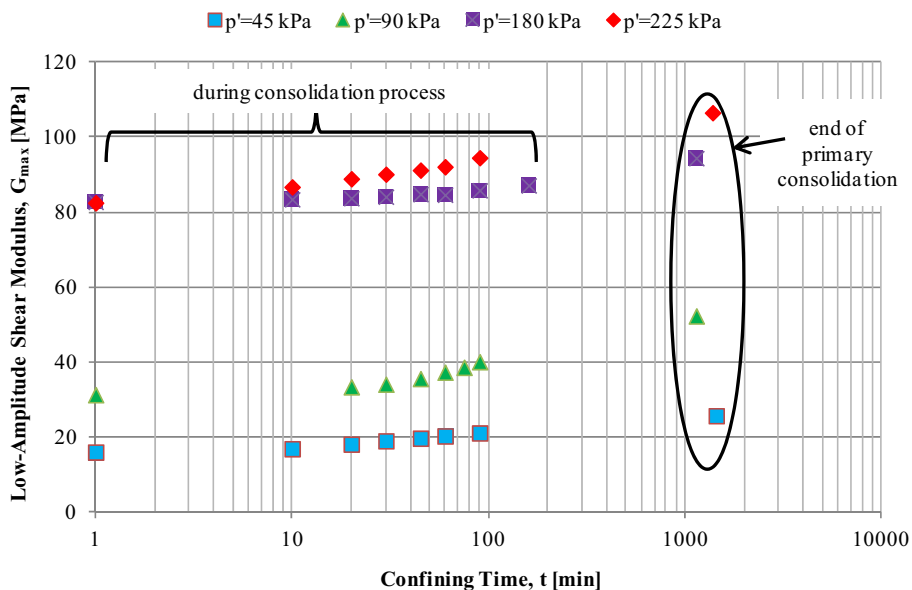
Fig. 2. Time-dependent shear modulus of clayey sand samples from Warsaw area under different stress conditions.
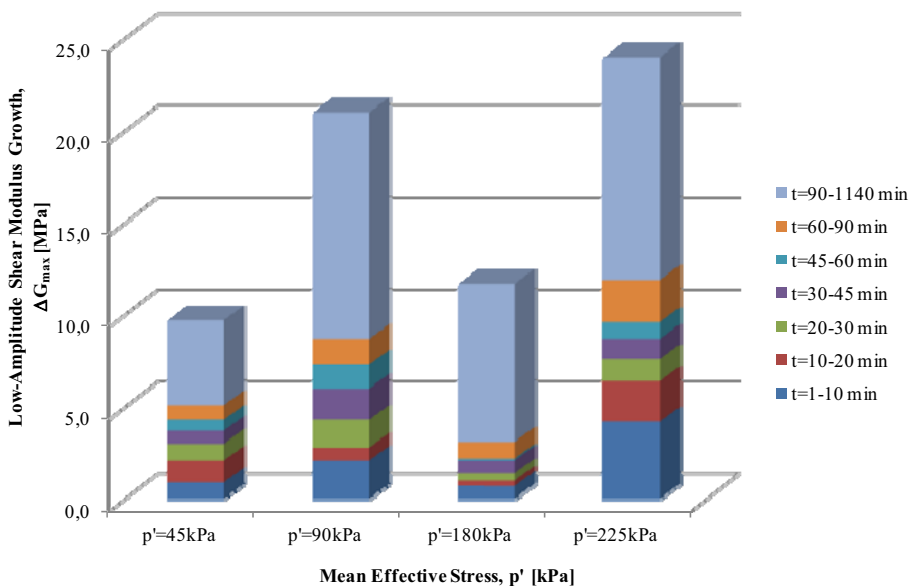
Fig. 3. Low-amplitude shear modulus growth in time at various stress levels.
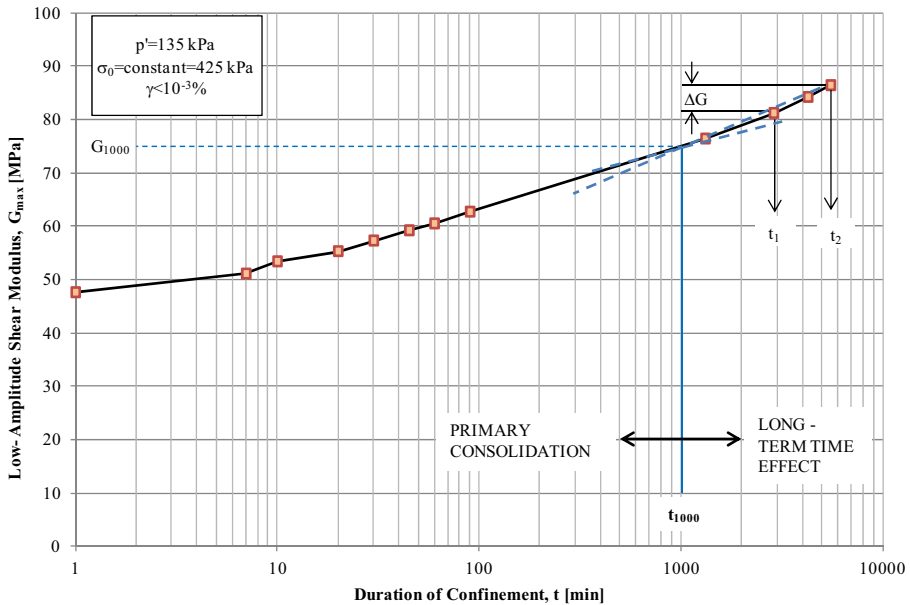
Fig. 4. Low-amplitude test sequence for a clayey sand sample and an example of the determination of I_G parameter.

During the whole low-amplitude test, the conditions with a possibility of drainage were maintained. It was not necessary to close the valves due to dynamic oscillations, because the level of deformation was too low to generate the excess pore-water pressure.

The measurements at low-amplitude for certain samples were performed after the end of primary consolidation was detected. On the basis of this method, the parameter I_G could be well evaluated. The procedure of determining an increase in low-amplitude shear modulus in time for a clayey sand sample is illustrated in Fig. 4.

Two distinct phases of modulus-time response are evident in Fig. 5. In order to better demonstrate this division into two stages of changes in G_{max}, the authors propose to introduce a new variable, called t_{Gstab}, which will indicate the end of primary consolidation and the beginning of the long-term time effect. The parameter t_{Gstab} will be the moment when changes in the value of low-amplitude of shear modulus start to be smaller, the values themselves are more stable, and their increase is low and nearly linear. This new parameter does not necessarily coincide with t_{1000}, which means the time at 1000 minutes of the constant confining pressure and, basically, in most cases, the end of primary consolidation.

In order to obtain dynamic shear modulus, resonant frequencies are first converted to shear wave velocities by using elastic theory formulas. The cor-
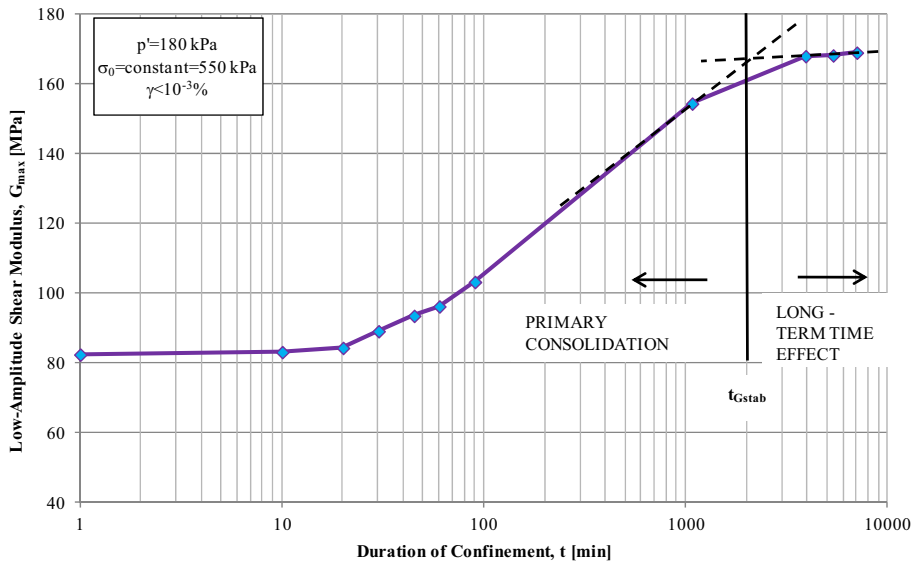
Fig. 5. Another illustration of the two phases of modulus-time response curve with the proposal to introduce a new parameter $t_{G\text{stab}}$.
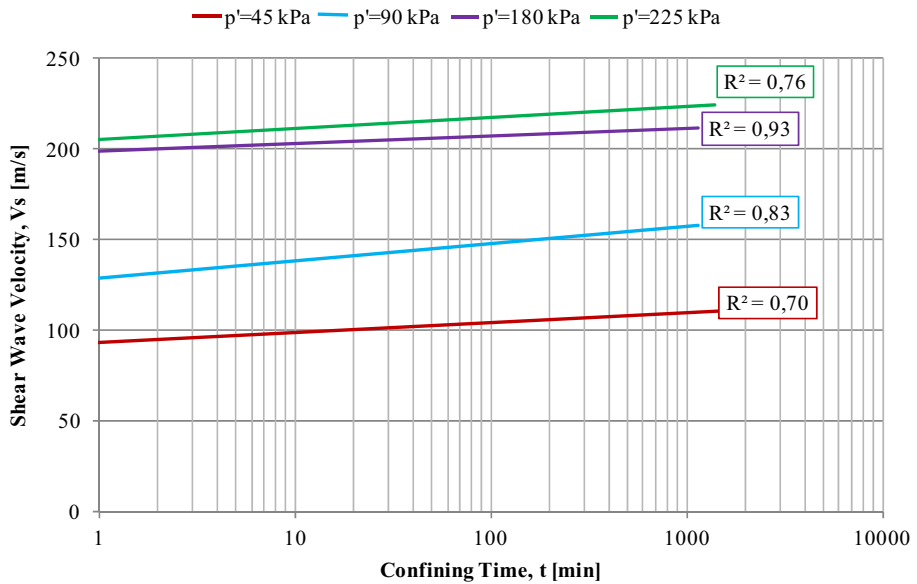
Fig. 6. Variation in shear wave velocity with time for clayey sand samples from Warsaw area under different stress conditions, ranging from 45 up to 225 kPa.

Duration of Confinement, t [min]
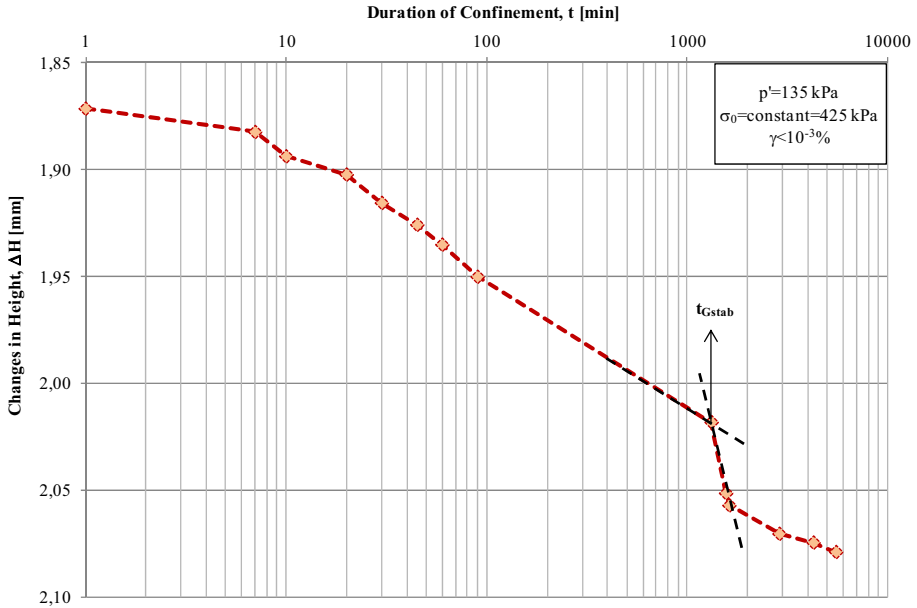
Fig. 7. Height changes with time for a clayey sand sample under constant confining pressure, *i.e.*, $p' = 135$ kPa.

rect calculations should consider the proper values of specimen height, weight, and volume at the measurement time. Figure 6 presents the variation of shear wave velocity as a function of logarithm of time at a fix pressure. With increasing time and increasing mean effective stress, the shear wave velocity rises as well, mainly from 1 until 3 m/s in a particular time period. The increase of V_S can be assumed linear (Fig. 6). The regression line, however, only for mean effective stress (p') of 90 and 180 kPa approximates well the real data points; the coefficient of determination (R^2) amounts to more than 0.8 (strong correlation). In the other two cases, assuming fitting by linear function as well, the authors received some positive correlations, but not a perfect one (R^2 is at the level of 0.7). In Figure 7 changes of specimen height are shown. As an example of the change in sample height during resonant column test, the measurements at the confining pressure $\sigma_0 = 425$ kPa and mean effective stress $p' = 135$ kPa were chosen.

4.2 High-amplitude test procedure

After the process of primary consolidation was finished, the measurements of dynamic shear modulus at shearing strains bigger than 0.001% were performed. During primary consolidation it is not recommended to vibrate specimens at higher stain amplitude due to the indefinite effective stress

conditions, as well as the possibility of variation in pore-water pressure and soil structure in the meantime. In the authors' researches, high-amplitude tests were conducted straight away after primary consolidation, but also some time later, up to 1 or 2 days. It was important to have the long-term time effect well defined at low-amplitude of shearing strain.

In this study, high-amplitude measurements were performed in several steps by increasing the amplitude of vibration in torsional mode. During these tests, in order to excite the electromagnetic field and induce a wave propagating through the examined material, the corresponding coil voltage values were applied, from the value of 0.1 up to 1.0 V, with a step of 0.1 V (Table 3). Figure 8 portrays the results from high-amplitude test sequence. Between all measurements there was always some rest time, about one minute.

Table 3

The tests conditions during high-amplitude measurements

Soil type	Speci-men type	Mean effective stress [kPa]	Time of the end of primary consolidation [min]	High shearing strain [%]	Shear modulus on high ampli-tude [MPa]	Corre-sponding coil voltage [V]
Clayey sand (1)	undis-turbed	45	1440	2.5E-02 – 3.7E-01	20-5	0.2-1.0
		90	1140	1.3E-02 – 1.4E-01	42-14	
		135	1320	8.9E-03 – 8.9E-02	76-24	
		180	1140	7.5E-03 – 6.6E-02	84-35	
		225	1380	6.7E-03 – 5.4E-02	97-44	
Clayey sand (2)	undis-turbed	180	1070	4.8E-03 – 5.4E-02	115-51	
		270	1220	2.0E-03 – 2.9E-02	171-102	
		315	1380	2.0E-03 – 2.7E-02	173-119	

After the studies on high-amplitude dynamic shear modulus were completed, the specimens remained under constant confining pressure, with drainage and low-amplitude modulus monitored with time (Fig. 9). It is visible in Fig. 9 that low-amplitude shear modulus had initially a different value, *i.e.*, smaller than that reached before the high-amplitude sequence. The observed decrease in dynamic shear modulus is temporary and lasts only until the low-amplitude modulus regains with time the value forecasted by the long-term time effect. If this does not happen, it may mean a permanent alteration of the specimen, which cannot be longer used in the tests.

A number of studies were carried out to show that long-term modulus increase occurs not only at low but also at intermediate strain level (Fig. 10).
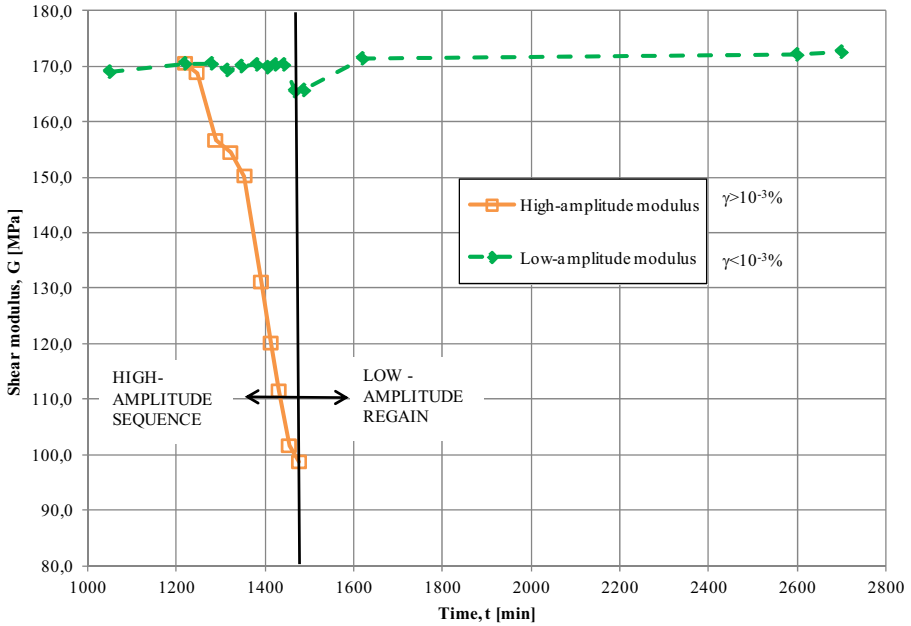
Fig. 8. High-amplitude test sequence (when shearing strain level $\gamma > 10^{-3}\%$) for a clayey sand sample obtained on the basis of laboratory tests.
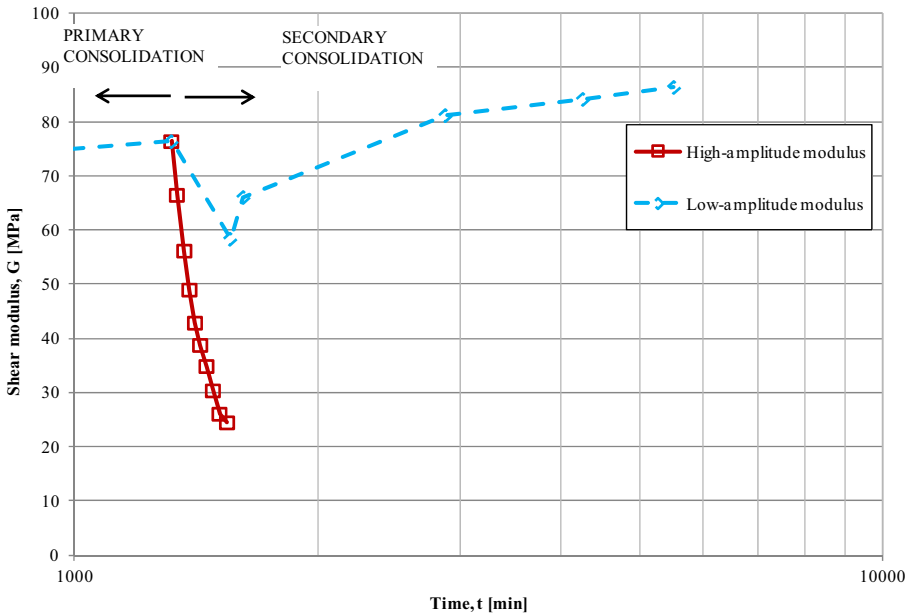
Fig. 9. Changes of high-amplitude and low-amplitude modulus with time and a temporary decrease in dynamic shear modulus.
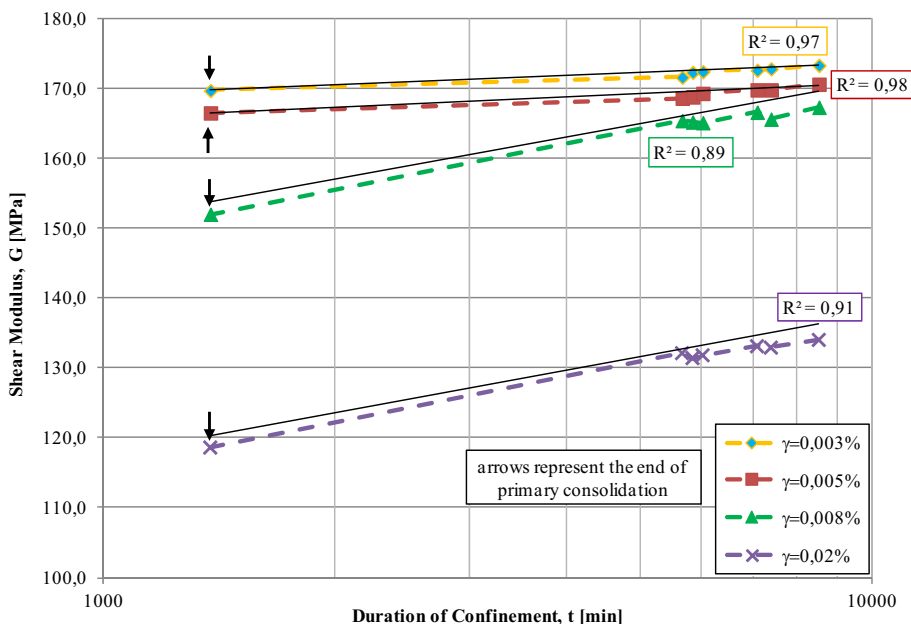
Fig. 10. Time-dependent shear modulus at high shearing strain amplitude (shearing strain values γ ranging from 3×10^{-3} % up to 2×10^{-2} %).

The authors' researches confirm tests of other scientists (Anderson and Stokoe 1978) that at shearing strain level between 0.001 and 0.1% the dynamic shear modulus grows with time, and this increase is linear; the linear regression line represents well the real data (R^2 at least equal to 0.9). The increase in values of dynamic shear modulus measured at high shearing strain amplitude is about 4 MPa for strain level $\gamma = 0.003$-0.005% and around 15 MPa for strain level $\gamma = 0.008$-0.02% during the studies. In the main, from the end of primary consolidation until the end of measurement time, the rise in shear modulus may be as high as 6-23%. Some general similarity between the growth with time in low- and high-amplitude shear modulus can be observed. Thus, it might be concluded that factors which have an influence on low-amplitude modulus time response affect the high-amplitude modulus time response as well.

5.   SUMMARY AND CONCLUSIONS

The behaviour of soil stiffness with time is of fundamental importance both in planning as well as in executing a laboratory testing program. Determination of shear modulus-time response and its proper understanding can be essential for prediction of *in situ* shear modulus based on the laboratory tests.

The authors of this paper tried to show this significance of the effect of confinement duration at a constant pressure on the magnitude of shear modulus of a selected cohesive soil from the capital of Poland while performing experiments in resonant column apparatus. This attempt is particularly important due to the lack of such information in the respect of Polish soils.

In summary, according to the literature and data presented in this article, the following conclusions can be drawn:

❑ Dynamic shear modulus of selected cohesive soil from Warsaw area varies with time of confinement at a fixed level of pressure.

❑ There are two phases which characterize the time-dependent response of shear modulus at low shearing strains (less than 0.001%). These are: an initial phase which occurs because of primary consolidation, and a second phase during which the modulus increases nearly linearly with the logarithm of time. The latter stage is referred to as "long-term time effect".

❑ The authors suggest the introduction of a new variable, $t_{G\text{stab}}$, as a moment which will be a boundary between the two phases, as described above (Fig. 5).

❑ At intermediate strain level (from 0.001 to 0.1%) a rise in shear modulus with time is noticeable as well. This increase in high-amplitude modulus is rather slightly less than that which occurs at low-amplitude shearing strains. This growth is exactly around 2% for shearing strain equal to $3 \times 10^{-3}\%$ and $5 \times 10^{-3}\%$, as well as approximately 10% for shearing strain equal to $8 \times 10^{-3}\%$ and $2 \times 10^{-2}\%$.

❑ The measurements of shear modulus at shearing strain above 0.001% should be performed as described by the authors in Subsection 4.2. Low-amplitude modulus can be used as a reference value in order to define some possible permanent changes in specimens when high-amplitude tests are performed.

The authors of this paper would like to emphasize that the role of the researchers is to point out the values and the variability of soil stiffness, whereas the role of the designers is to choose the right parameters for the proper designing and planning of engineering constructions. It is necessary therefore to keep effective cooperation between the engineers and the designers, from the beginning of the construction design and while considering various construction work schemes. Although geotechnical activities are defined by Eurocode 7 (EN 1997), there is no detailed description of laboratory experiments of the type presented in this article. Perhaps, this is due to some difficulties relating to the determination of the dynamic shear modulus and its further application by the designers.

R e f e r e n c e s

Afifi, S.S., and F.E. Richart Jr. (1973), Stress-history effects on shear modulus of soils, *Soils Found.* **13**, 1, 77-95, DOI: 10.3208/sandf1972.13.77.

Anderson, D.G., and K.H. Stokoe II (1978), Shear modulus: A time-dependent soil property. **In:** *Dynamic Geotechnical Testing*, ASTM Special Tech. Publ. No. 654, American Society for Testing and Materials, Baltimore, 66-90.

Anderson, D.G., and R.D. Woods (1976), Time-dependent increase in shear modulus of clay, *J. Geotech. Eng. Div. ASCE* **102**, 5, 525-537.

EN (1997), *Eurocode 7: Geotechnical Design – General Rules*, European Prestandard ENV.

Gabryś, K., W. Sas, and A. Szymański (2013), Resonant column apparatus as a device for dynamic testing of cohesive soils, *Prz. Nauk. Inż. Kształt. Środ.* **59**, 3-13 (in Polish).

GDS (2010), *The GDS Resonant Column System Handbook (Version 2.2.2010)*, Geotechnical Digital Systems Instruments Co., London.

Howie, J.A., T. Shozen, and Y. P. Vaid (2002), Effect of ageing on stiffness of very loose sand, *Can. Geotech. J.* **39**, 1, 149-156, DOI: 10.1139/T01-085.

Jamiolkowski, M., and N. Manassero (1996), The role of in situ testing in geotechnical engineering – thoughts about the future. **In:** C. Craig (ed.), *Advances in Site Investigation Practice*, Thomas Telford Publ., London, 929-951.

Khosravi, A., M. Ghayoomi, J. McCartney, and H.-Y. Ko (2010), Impact of effective stress on the dynamic shear modulus of unsaturated sand. **In:** D.O. Fratta, A.J. Puppala, and B. Muhunthan (eds.), *Proc. GeoFlorida 2010: Advances in Analysis, Modelling & Design*, American Society of Civil Engineers, 410-419, DOI: 10.1061/41095(365)38.

Kokusho, T., Y. Yoshida, and Y. Esashi (1982), Dynamic properties of soft clay for wide strain range, *Soils Found.* **22**, 4, 1-18, DOI: 10.3208/sandf1972. 22.4_1.

Lawrence, F.V. Jr. (1965), Ultrasonic shear wave velocities in sand and clay, Res. Rep. No. 23, Massachusetts Institute of Technology, Cambridge, USA.

Markowska-Lech, K., M. Lech, M. Bajda, and A. Szymański (2013), Small strain stiffness in overconsolidated Pliocene clays, *Ann. Warsaw Univ. Life Sci. SGGW* **45**, 2, 169-181, DOI: 10.2478/sggw-2013-0014.

Mesri, G., T.W. Feng, and J.M. Benak (1990), Postdensification penetration resistance of clean sands, *J. Geotech. Eng.* **116**, 7, 1095-1115, DOI: 10.1061/ (ASCE)0733-9410(1990)116:7(1095).

Richart, F.E. Jr. (1962), Closure of "Foundation Vibrations", *Trans. ASCE* **127**, 1, 918-926.

Sas, W., and K. Gabryś (2012), Laboratory measurement of shear stiffness in resonant column apparatus, *Acta Sci. Pol. – Architectura* **11**, 4, 29-39.

Sidorova, A.I., and E.A. Voznesensky (2010), Estimation of additional foundation settlements caused by dynamic loading in urban areas, *Acta Geophys*. **58**, 1, 126-142, DOI: 10.2478/s11600-009-0040-7.

Soga, K. (2005), Lecture 3: Time effects observed in granular materials, Socio-Environmental Engineering, COE Hokkaido University, Hokkaido, Japan, http://www.eng.hokudai.ac.jp/COE-area/workshop/pdf/05feb_lec_soga3.pdf.

Tezcan, S.S., Z. Ozdemir, and A. Keceli (2009), Seismic technique to determine the allowable bearing pressure for shallow foundations in soils and rocks, *Acta Geophys*. **57**, 2, 400-412, DOI: 10.2478/s11600-008-0077-z.

Vucetic, M., and R. Dobry (1991), Effect of soil plasticity on cyclic response, *J. Geotech. Eng. ASCE* **117**, 1, 89-107, DOI: 10.1061/(ASCE)0733-9410 (1991)117:1(89).

Zavoral, D. (1990), Dynamic properties of an undisturbed clay from resonant column tests, M.Sc. Thesis, Department of Civil Engineering, The University of British Columbia, Vancouver, Canada.

Zhang, J., R.D. Andrus, and H. Juang (2005), Normalized shear modulus and material damping ratio relationships, *J. Geotech. Geoenviron. Eng. ASCE* **131**, 4, 453-464, DOI: 10.1061/(ASCE)1090-0241(2005)131:4(453).

Redatuming Operators Analysis
in Homogeneous Media

Fransisco de Souza OLIVEIRA[1,4],
Jose J.S. de FIGUEIREDO[1,2], and Lucas FREITAS[3]

[1]Faculty of Geophysics, Federal University of Para,
Belem, Brazil; e-mail: frasol@ufpa.br
[2]National Institute of Petroleum Geophysics (INCT-GP),
Salvador, Brazil; e-mail: jadsom@ufpa.br (corresponding author)
[3]Geoprocessados, Virlemosa, Mexico; e-mail: lucas.batista.freitas@gmail.com
[4]Faculty of Meteorology, Federal University of Para, Belem, Brazil

A b s t r a c t

A redatuming operation is used to simulate the acquisition of data in new levels, avoiding distortions produced by near-surface irregularities related to either geometric or material property heterogeneities. In this work, the application of the true-amplitude Kirchhoff redatuming (TAKR) operator on homogeneous media is compared with conventional Kirchhoff redatuming (KR) operator restricted to the zero-offset case. The TAKR and the KR operators are analytically and numerically compared in order to verify their impacts on the data at a new level. Analyses of amplitude and velocity sensitivity of the TAKR and KR were performed: one concerning the difference between the weight functions and the other related to the velocity variation. The comparisons between operators were performed using numerical examples. The feasibility of the KR and TAKR operators was demonstrated not only kinematically but also dynamically for their purposes. In other words, one preserves amplitude (KR), and the other corrects the amplitude (TAKR). In the end, we applied the operators to a GPR data set.

Key words: Kirchhoff redatuming, true-amplitude Kirchhoff redatuming, amplitude analysis, travel-time analysis.

1. INTRODUCTION

Wavefield redatuming is an operation that transforms seismic data based on the assumption of a new measurement surface. In other words, given a data set acquired on an initial surface, it generates a simulated data set as if it were measured on another surface. Among its applications, we have near-surface corrections (Cox 1999), OBC processing (Jin *et al.* 2011), dual-sensor streamer de-ghosting (Klüver 2009), and multiple attenuation (Wiggins 1988).

The wave-equation based redatuming operators are the most accurate ones. Over the years several approaches have been proposed: based on the Kirchhoff integral (Berryhill 1984), using the phase-shift method (Margrave and Ferguson 1999) and based on the common-focus point (CFP) technology.

The redatuming operators are especially costly in the pre-stack domain in which most applications occur. Though the Kirchhoff method is rather straightforward and efficient, it is still expensive compared to static correction and requires an accurate interval velocity field above the datum. In addition, the Kirchhoff scheme is applied to common source and receiver gathers; in other to relocate sources and receivers, respectively. An analytical Kirchhoff-type redatuming operator, based on straight ray approximation (SRD) (Alkhalifah and Bagaini 2006) fills the gaps between static correction and wave-equation redatuming. It uses the assumption of local homogeneity, potentially useful for most media. The small size of the operator and its analytical expression provides cost-effectiveness and little sensitivity to velocity errors.

Toward a true-amplitude Kirchhoff-type operator, particular cases of the migration to zero-offset (MZO) operator proposed by Tygel *et al.* (1998) were analytically formulated for zero-offset configuration on homogeneous models and compiled into a true-amplitude Kirchhoff redatuming (TAKR) operator (Oliveira *et al.* 2009, Pila *et al.* 2014). This operator provides correct kinematic and dynamic redatumed traces. The reader should notice that the term true-amplitude is used here on a more strict sense, beyond amplitude relativity preservation. More specifically, the amplitude is not only preserved within a given event for different offsets. In this case, the amplitude has its geometrical spreading component adjusted to honor the new measurement surface.

For seismic data processing, the restriction to zero-offset configuration limits the applicability of a redatuming operator to event repositioning (*e.g.*, in moving a stack from floating to final datum). However, as described by Liu *et al.* (2007), zero-offset redatuming plays a more important role in imaging GRP profiles. Moreover, an amplitude-friendly processing sequence is of great importance since amplitude analysis of GPR profiles has applications in shallow aquifers characterization (Bradford 1999), determination of subsurface contaminant (Schmalza and Lennartzb 2002) and soil water content variations (Cassidy 2007) in hydrological studies, and archaeological prospection (Khwanmuang and Udphuay 2012, Zhao *et al.* 2013).

In this work, we analyze the amplitude behavior of two Kirchhoff-based redatuming operators through homogeneous media for the zero-offset case: the operator described by Berryhill (1984) and the true-amplitude operator proposed by Pila *et al.* (2014). We illustrate the feasibility of our analysis for a couple of numerical examples and for the application in a GPR data set from Siple Dome-Antarctica.

2. THEORETICAL BACKGROUND

In this section, we show our mathematical analysis for amplitude and travel-time attributes at redatuming operation. Our analysis relies on the stationary phase method to analytically solve the Kirchhoff redatuming in 2.5 D homogeneous media restricted to a zero-offset configuration.

2.1 2.5 D Kirchhof-based redatuming

Kirchhoff redatuming is based on the integral formulation of Kirchhoff migration (Schneider 1978). Like its migration counterpart, the redatumed wavefield U_o is calculated by successive weighted summations of input wavefield U_i along proper trajectories. More specifically,

$$U_o(\xi_o, t_o) = \frac{1}{\sqrt{2\pi}} \int_{Z_i} W d^{1/2} U_i(\xi_i, t_o \pm \tau) d\xi_i \,, \tag{1}$$

where ξ_i and ξ_o are the horizontal coordinates of the input and output datums Z_i and Z_o, $d^{1/2}$ denotes the half-derivative of the input wavefield U_i, τ is the travel-time between the output location $(\xi_o, Z_o(\xi_o))$ and input location $(\xi_i, Z_i(\xi_i))$, W is a properly chosen weight. In frequency domain, $d^{1/2}$ corresponds to $i\omega$. Note that the $\pm$ sign is chosen appropriately whether the output datum is above $(-)$ or below $(+)$ the input datum.

In this work, we analyse the amplitude behavior of two Kirchhoff-based redatuming operators through homogeneous media for the zero-offset case: the preserving operator described by Berryhill (1984) and the true-amplitude operator proposed by Pila *et al.* (2014).

These operators are kinematically identical. As expected, when redatuming from datum Z_i to Z_o, through a homogeneous medium with velocity v, the travel-time τ is directly calculated from the distance between input and output locations $d(\xi_i, \xi_o)$, namely

$$d(\xi_i, \xi_o) = \sqrt{(\xi_o - \xi_i)^2 + (Z_o(\xi_o) - Z_i(\xi_i))^2]} \,. \tag{2}$$

However, while the former aims to preserve relative amplitudes by direct application of the Kirchhoff integral, the latter aims to adjust the geometrical spreading by direct application of the MZO operator. These objectives are reached by using different stacking weights. The amplitudes are preserved using the weight W_{pa} given by

$$W_{pa} = \frac{\phi(\xi_i, \xi_o)}{\sqrt{vd}} , \tag{3}$$

where

$$\phi(\xi_i, \xi_o) = d^{-1} \left[Z_o(\xi_o) - Z_i(\xi_i) - (\xi_o - \xi_i) Z_i'(\xi_i) \right] \tag{4}$$

accounts for the incidence correction. On the other hand, the true-amplitudes are achieved incorporating a geometrical spreading adjusting term

$$\mathcal{G}(\xi_i, \xi_o, t_o, v) = \frac{2d}{vt_o} \tag{5}$$

into the summation weights, generating the true-amplitude weight W_{ta} given by

$$W_{ta} = (1 \pm \mathcal{G}) W_{pa} = \left(1 \pm \frac{2d}{vt_o} \right) \frac{\phi(\xi_i, \xi_o)}{\sqrt{vd}} , \tag{6}$$

where t_0 is the output time and d is the distance between input and output locations.

2.2 Analytical analysis of velocity sensitiveness

A first understanding of the impact of velocity errors on both operators can be obtained by analytical analysis of a simple case. In order to achieve that, we derive the stationary phase approximation (Bleistein 1986) of Eq. 1 in the case of horizontal plane input and output datums and a horizontal reflector in a constant velocity medium.

In the case of a reflector located at depth D and with unitary reflectivity, the zero-offset input data set $U_i(\xi_i, t)$ is given by

$$U_i(\xi_i, \omega) = \frac{1}{4\pi D} e^{-i\omega 2D/v} , \tag{7}$$

where v is the medium velocity. Therefore, when using a redatuming velocity ν, Eq. 1 can be replaced by

$$U_o(\xi_o, \omega) = \frac{1}{\sqrt{2\pi}} \int_{Z_i} W(\xi_i; \xi_o) \left(\frac{\sqrt{i\omega}}{4\pi D} e^{-i\omega 2D/v} \right) e^{\pm i\omega 2d(\xi_i; \xi_o)/\nu} d\xi_i$$

$$= \frac{1}{\sqrt{2\pi}} \int_{Z_i} \left(W(\xi_i; \xi_o) \frac{\sqrt{i\omega}}{4\pi D} \right) e^{-i\omega \phi(\xi_i; \xi_o)} d\xi_i , \tag{8}$$

where $\phi(\xi_i; \xi_o) = 2D/v \pm 2d(\xi_i; \xi_o)/\nu$. For this simple case, it is easy to note that the phase $\phi(\xi_i; \xi_o)$ is stationary when $\xi_i = \xi_o$. The redatumed sample $U_o(\xi_o, \omega)$ can then be approximated by

$$
\begin{aligned}
U_o(\xi_o, \omega) &\approx \sqrt{\frac{1}{i\omega \phi''(\xi_o; \xi_o)}} \left(W(\xi_o; \xi_o) \frac{\sqrt{i\omega}}{4\pi D} \right) e^{-i\omega \phi(\xi_o; \xi_o)} \\
&\approx \sqrt{\frac{\nu z}{2i\omega}} \left(W(\xi_o; \xi_o) \frac{\sqrt{i\omega}}{4\pi D} \right) e^{-i\omega 2(D/v \pm z/\nu)} \\
&\approx W(\xi_o; \xi_o) \left(\frac{\sqrt{\nu z}}{4\pi D \sqrt{2}} \right) e^{-i\omega 2(D/v \pm z/\nu)} .
\end{aligned}
\tag{9}
$$

In the case of preserving amplitudes we use $W = W_{pa}$ (Eq. 5), after which we have

$$
\begin{aligned}
U_o(\xi_o, \omega) &\approx W_{pa}(\xi_o; \xi_o) \frac{\sqrt{\nu z}}{4\pi D \sqrt{2}} e^{-i\omega 2(D/v \pm z/\nu)} \\
&\approx \left(\frac{\sqrt{2}}{\sqrt{\nu z}} \right) \frac{\sqrt{\nu z}}{4\pi D \sqrt{2}} e^{-i\omega 2(D/v \pm z/\nu)} \\
&\approx \frac{1}{4\pi D} e^{-i\omega 2(D/v \pm z/\nu)} .
\end{aligned}
$$

As expected, the redatumed signal has the same amplitude as the input data (Eq. 7), it does not depend on the chosen redatuming velocity ν, and its phase is shifted by z/ν.

In the case of taking into account the geometrical spreading factor we use $W = W_{ta}$ into the Fourier inverse of Eq. 9:

$$
\begin{aligned}
U_o(\xi_o, t_o) &\approx W_{ta}(\xi_o; \xi_o) \frac{\sqrt{\nu z}}{4\pi D \sqrt{2}} \delta \left(t_o - 2(D/v \pm z/\nu) \right) \\
&\approx \left[\left(1 \pm \frac{2z}{\nu t_o} \right) \frac{\sqrt{2}}{\sqrt{\nu z}} \right] \frac{\sqrt{\nu z}}{4\pi D \sqrt{2}} \delta \left(t_o - 2(D/v \pm z/\nu) \right) \\
&\approx \left(1 \pm \frac{2z}{\nu t_o} \right) \frac{1}{4\pi D} \delta \left(t_o - 2(D/v \pm z/\nu) \right) .
\end{aligned}
\tag{10}
$$

Note that in order to keep consistency, the $\pm$ operator, chosen according to the redatuming direction, has an opposite meaning in the last equation. For example, when redatuming upwards, the first occurrence is $-$ and the other

is $+$. Namely, at the redatumed reflector we have

$$
\begin{aligned}
U_o\left(\xi_o, t_o = 2(D/v \pm z/\nu)\right) &\approx \left(1 - \frac{z}{D\nu/v + z},\right) \frac{1}{4\pi D} \\
&\approx \left(\frac{\nu}{v}\right) \frac{1}{4\pi(D\nu/v + z)}, \\
&\approx \frac{1}{4\pi(D + zv/\nu)} \ .
\end{aligned} \tag{11}
$$

The true-amplitude redatuming operator, in order to properly account for geometrical spreading, relies on a proper choice of the redatuming velocity ν. The redatumed amplitude is strongly dependent on the relation between the chosen redatuming velocity ν and the medium velocity v.

3. RESULTS

The validity of the analysis of amplitude and travel-time with a variation in the velocity model was confirmed by numerical and real data sets. In this section, we have performed our analysis on three homogeneous numerical examples and on a real GPR data set.

3.1 Model I: two horizontal layers

In this synthetic example, we apply redatuming operators in a horizontally layered model with two horizontal homogeneous layers, where the first acoustic wave velocity background was $v_1 = 2000$ m/s and the second was $v_2 = 2200$ m/s. The model that consists of two horizontal homogeneous layers is depicted in Fig. 1. The zero-offset data set (Fig. 2a), was simulated with a 25 Hz Ricker wavelet by Kirchhoff modeling and sampled at 4 ms every 10 m. Its measurement level is constant at level $z = 0$ m. Another data set, to be used as reference, was simulated with the measurement level equal to the target output
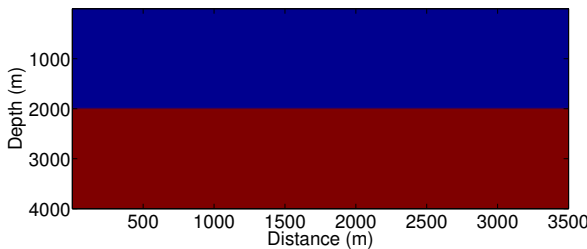
Fig. 1. Velocity model with two homogeneous layers (Model I). The first layer velocity is 2000 m/s and the second 2200 m/s.
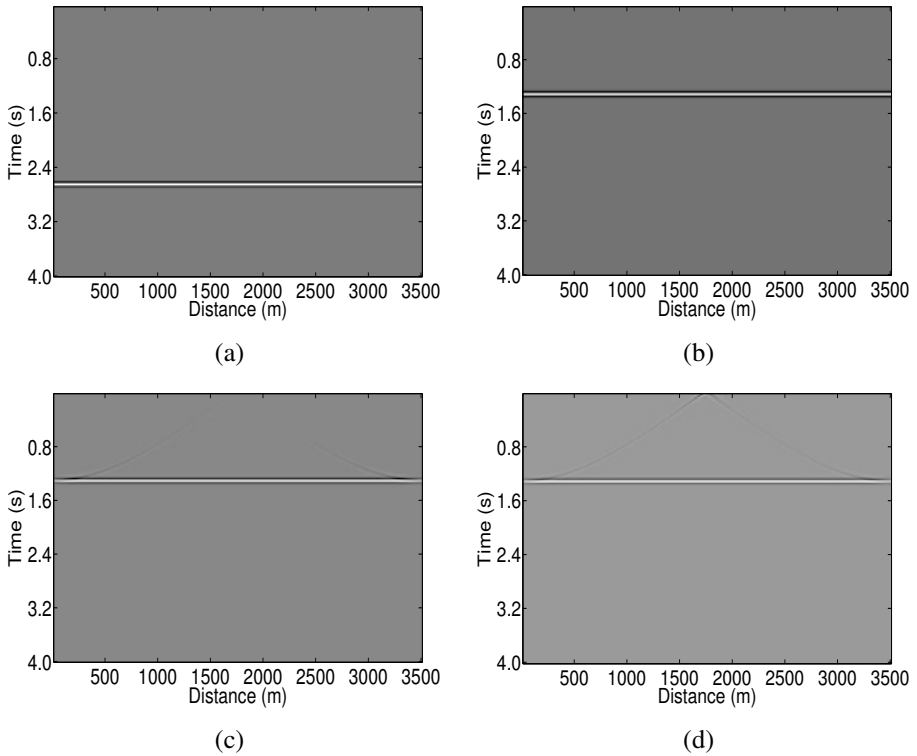
Fig. 2. Model I data sets: (a) input, (b) reference and outputs of (c) amplitude-preserving, and (d) true-amplitude redatuming operators.

datum ($z = 1000$ m) with identical parameters (see Fig. 2b). This datum level corresponds to 50% of the first layer depth.

The output of both true-amplitude and conventional redatuming operators is presented in Fig. 2c, d. A single-trace detailed comparison is presented in Fig. 3a. As expected, both operators result in traces kinematically equal to the reference trace. On the other hand, while the output from the conventional operator has the same amplitude as the input, the true-amplitude operator yields traces with the same amplitude as the reference trace. As we can notice in Fig. 3b, the amplitude erros are low and probably due to numerical issues.

A second experiment on the same model aimed to analyse the velocity sensitivity on both redatuming operators. Since the operators are kinematically the same, the errors in travel-time are identical (see Fig. 4a, b). On the other hand, as commented in Subsection 2.2, while the amplitude of the conventional operator is independent of the velocity error, the true-amplitude operator is considerably sensitive to it. The errors of the amplitude-preserving operator remain at 7%

(a)

(b)
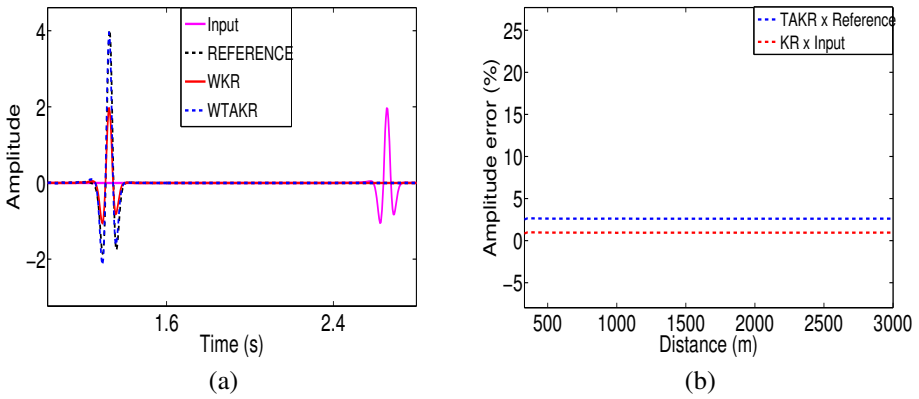
Fig. 3: (a) Seismic trace of the input (purple), reference (black dashed), KR (red), and TAKR (blue dashed). The KR and TAKR data were repositioned with the weight function (3) and (6), respectively, and the redatuming in this case recovered the correct amplitude; (b) The relative amplitude errors between the TAKR data and the reference (blue line) are around 3.5 %. The relative error between the KR and the reference data (red line) is around 1%.

(a)

(b)

(c)

(d)
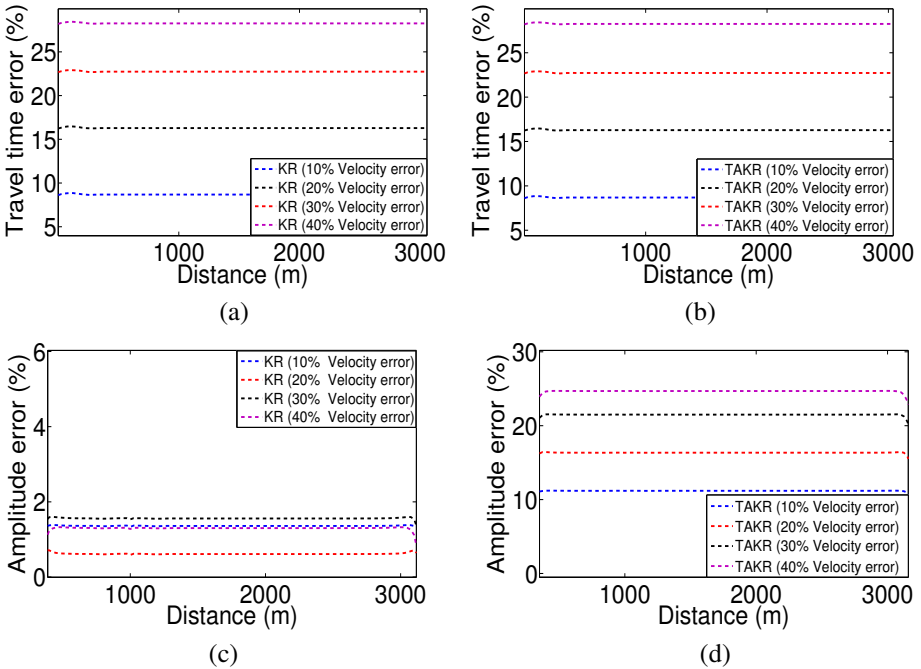
Fig. 4. Velocity error plots: traveltime error for (a) amplitude-preserving and (b) true-amplitude redatuming operators, and the respective amplitude errors shown in (c) and (d).

even for velocity errors of 40%, while the true-amplitude operator errors reach 27% in this case.

As it can be seen, both operators present accurate results, either preserving or adjusting the amplitude. However, in order to obtain an estimate of the geometrical spreading, the true-amplitude operator relies on a proper velocity model. Both analytic and numerical analyses show that the amplitude error for this operator is strongly affected by errors in the velocity field.

3.2 Model II: four horizontal layers

Here, in order to verify the feasibility of both operators when V_{RMS} velocities are taken into account, we generated a second model which consists of four horizontal homogeneous layers, as depicted in Fig. 5. The interval velocities of this model were 1500, 2000, 2500, and 3000 m/s.

The source-receiver pairs were positioned at every 25 m. The data was modelled by Trisies (from Seismic Unix) which uses a Gaussian beam operator to produce true-amplitude synthetic seismograms with a Ricker wavelet of 25 Hz. This seismogram was used as an input for KR operator in zero-offset. The zero-offset data set, shown in Fig. 6a has a sampling ratio of 4 ms and its measurement level is constant at $z = 0$ m. Another data set, to be used as reference, was simulated having the measurement level equal to the target output datum ($z = 1000$ m) with identical parameters (see Fig. 6b) also by Trisies.

The outputs of both true-amplitude and conventional redatuming operators are presented in Fig. 6c, d. One-trace detailed comparisons are presented in Figs. 7 and 8. Just like the first experiment, both operators presented good accuracy in their kinematic and dynamic purposes.

As we can see Fig. 8, even when the V_{RMS} velocity is used in the redatuming operation, the error between the KR and input data set is less than 1% for all layers, even deepest. However, in case of TAKR redatuming operation
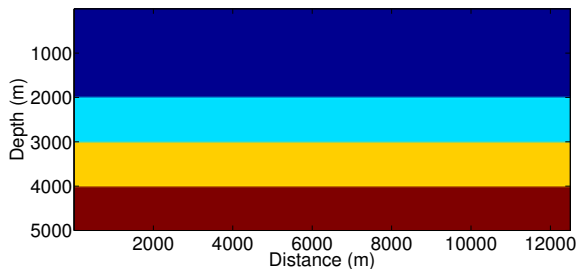
Fig. 5. Velocity model with four homogeneous layers with velocities 1500, 2000, 2500, and 3000 m/s (Model II).
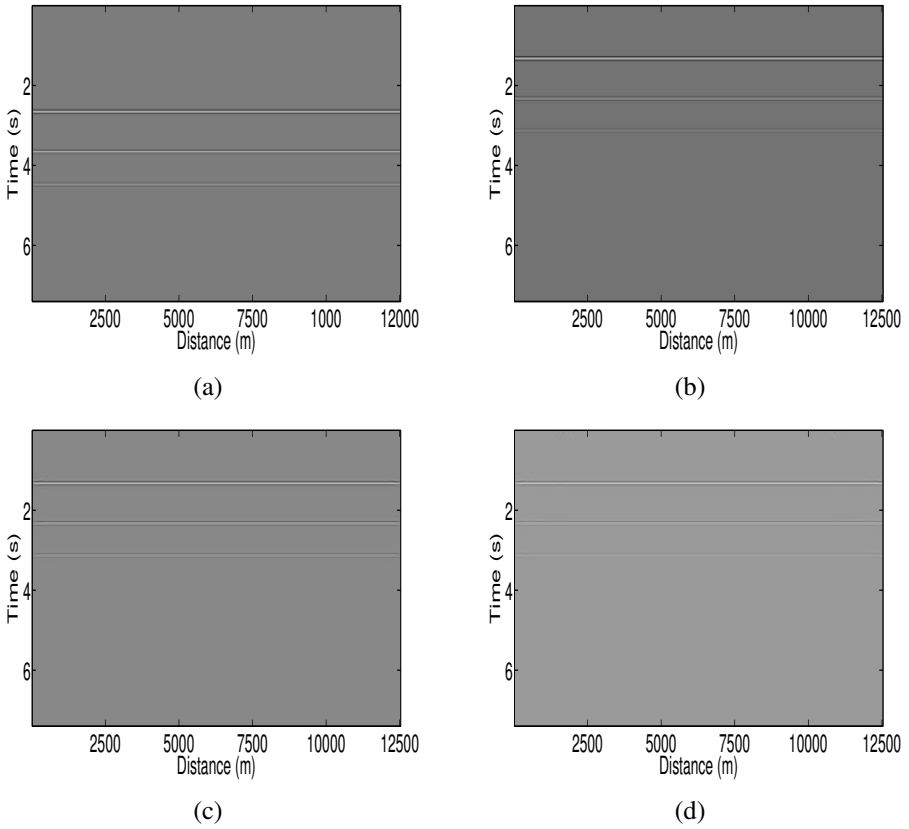
Fig. 6. Model II data sets: (a) input, (b) reference and outputs of (c) amplitude-preserving, and (d) true-amplitude redatuming operators.
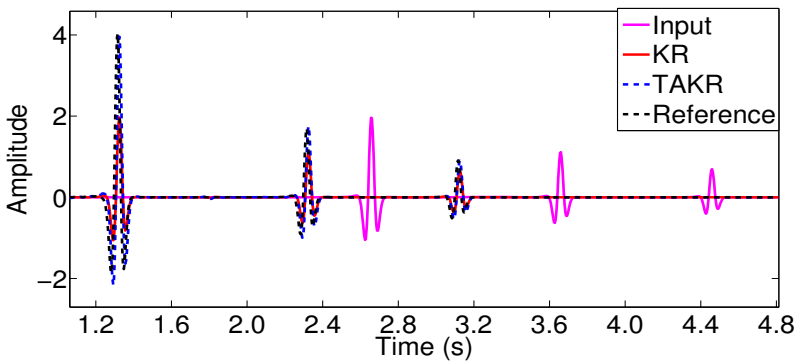
Fig. 7. Single-trace detailed comparison of Model II data sets.
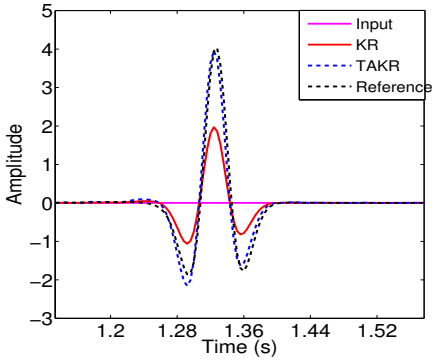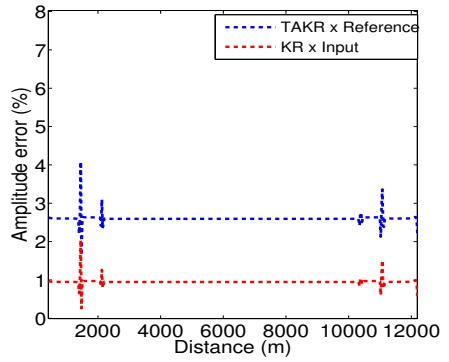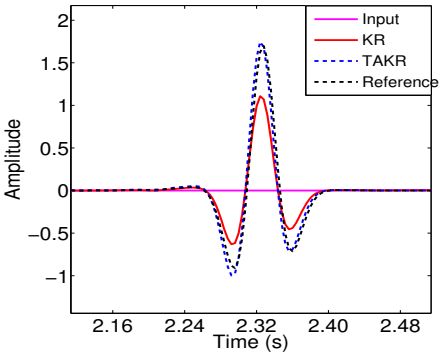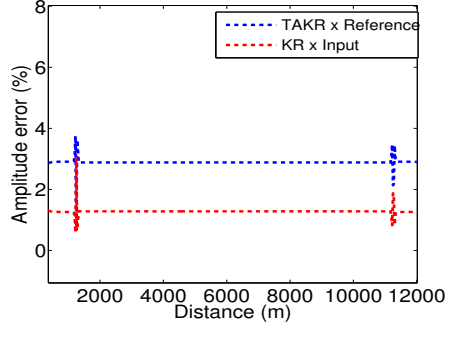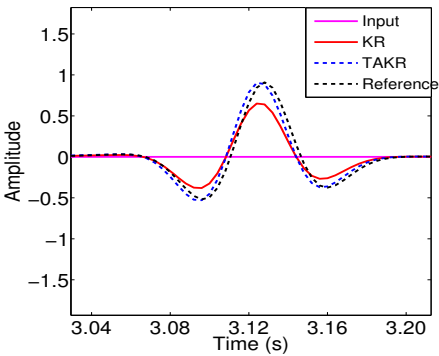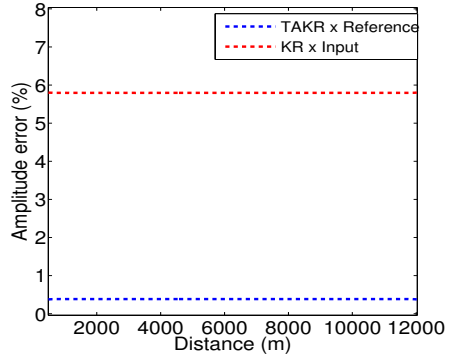
Fig. 8. Detailed comparison: (a), (c), and (e) single-trace zoom for first, second, and third events; (b), (d), and (f) multi-trace amplitude error plot for first, second, and third events.

the error is more sensitive to differences between true and V_{RMS} velocities. Based on what Subsection 2.2 shows, the TAKR is more sensitive to velocity error. The deeper the layer, the more V_{RMS} moves away from interval velocity. It is expected the that TAKR operator is more sensitive to velocities of the deeper layers.

3.3 Model III: four curved layers

The third model consists of four curved homogeneous layers as depicted in Fig. 9. The zero-offset data set (Fig. 10a) was simulated with a 25 Hz Ricker wavelet by Gaussian beam modelling, sampled at 4 ms and 10 m, and its measurement level is constant at $z = 0$ m. Another data set, to be used as reference, was simulated having the measurement level equal to the target output datum ($z = 200$ m) with identical parameters (see Fig. 10b).

The output of both true-amplitude and conventional redatuming operators are presented in Fig. 10c, d. One-trace detailed comparison is presented in Fig. 11.

Figure 11 compares the central trace of the KR and TAKR data with the central trace of the input and the reference sections. Just like in the model with four layers, this experiment showed that the difference in amplitude between the KR response and the input data is still the same with the increase of depth, and the TAKR response and reference data increased with the increase of depth. However, when it comes to amplitude recovering, the TAKR profile has the amplitude recovering factor.
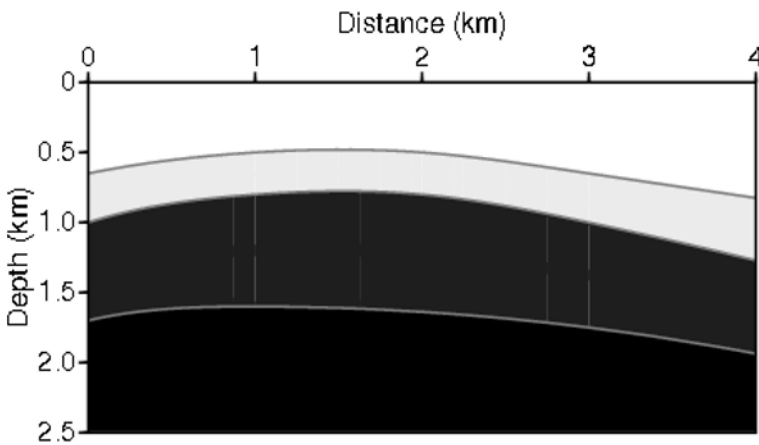
Fig. 9. Velocity model with four curved layers with velocities 1581, 1690, 1826, and 2000 m/s (Model III).
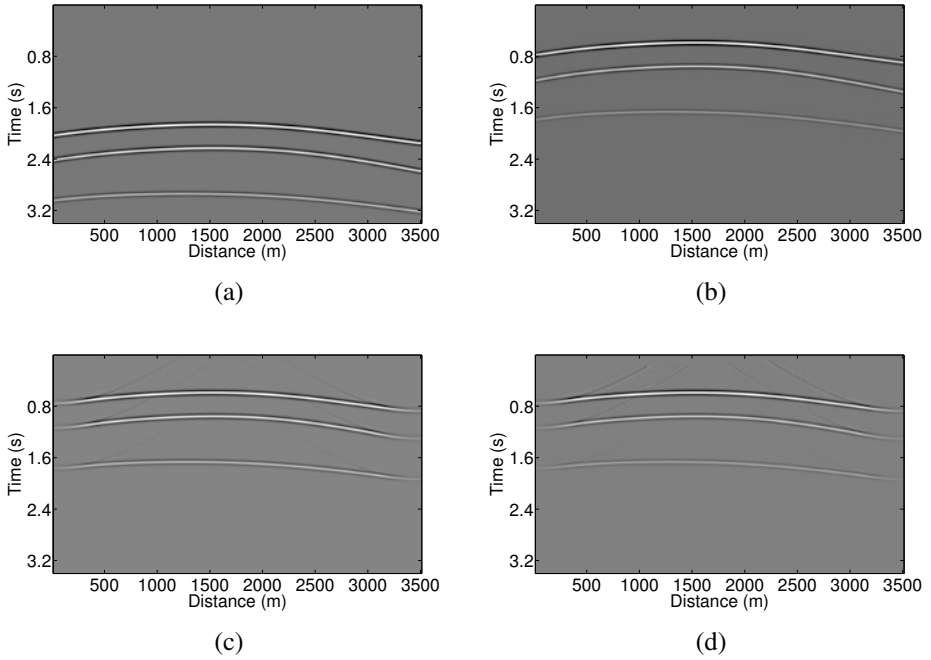
Fig. 10. Model III data sets: (a) input, (b) reference and outputs of (c) amplitude-preserving, and (d) true-amplitude redatuming operators.
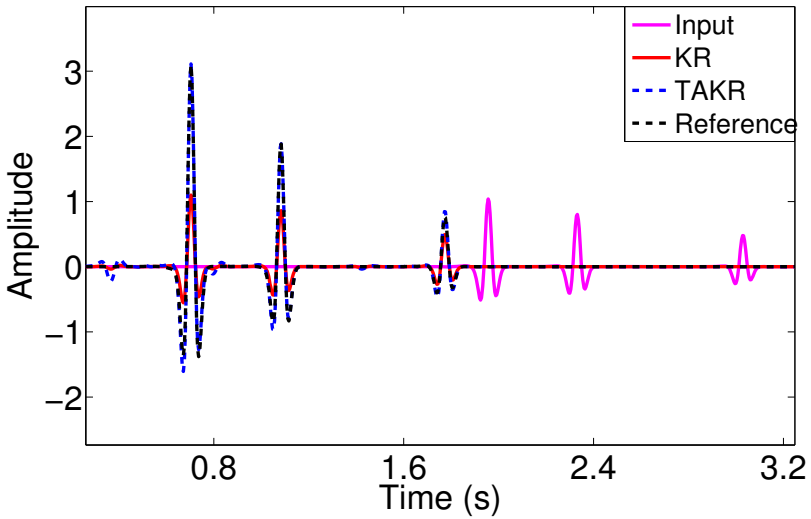
Fig. 11. Single-trace detailed comparison of Model III data sets.

4. FIELD DATA EXAMPLE:
GROUND PENETRATING RADAR (GPR) DATA SET

Siple Dome (81.65 S, 148.81 W) is an ice dome which is approximately 100 km wide and 100 km long, located 130 km east of Siple Coast in Antarctica. This place is of particularly important on determining the current mass balance of the West Antarctic ice sheet (WAIS).

The GPR survey was performed at a specific location indicated in Fig. 12a. Figure 12b shows a slice of the 100 MHz GPR profile acquired in Siple Dome. The distance between the transmitter and receiver antennas was 1.0 m (half-offset, $h = 0.50$ m), and the interval between traces was 0.75 m. A total of 8000 traces was collected along the 6000 m survey line. But here, for observation issues we are showing a spatial window of 3000 m (see Fig. 12b). The length of the time window was 913 ns and the number of samples per trace was 1870, resulting in a time sampling rate of 0.49 ns.

As an illustrative example, we applied both operators to the profile from Fig. 12b, redatuming it to a flat output datum located at 15 m depth. On a first example, we used ice velocity (0.16 m/ns) as redatuming velocity (Alley and Bentley 1988). As expected, the difference between the two results is purely dynamic (see Fig. 13). While the conventional amplitude-preserving operator maintains the relativeness of the input profile (Fig. 13a), the true-amplitude operator boosts up amplitudes on the first 300 ns of the section. This enhancement

(a)   (b)
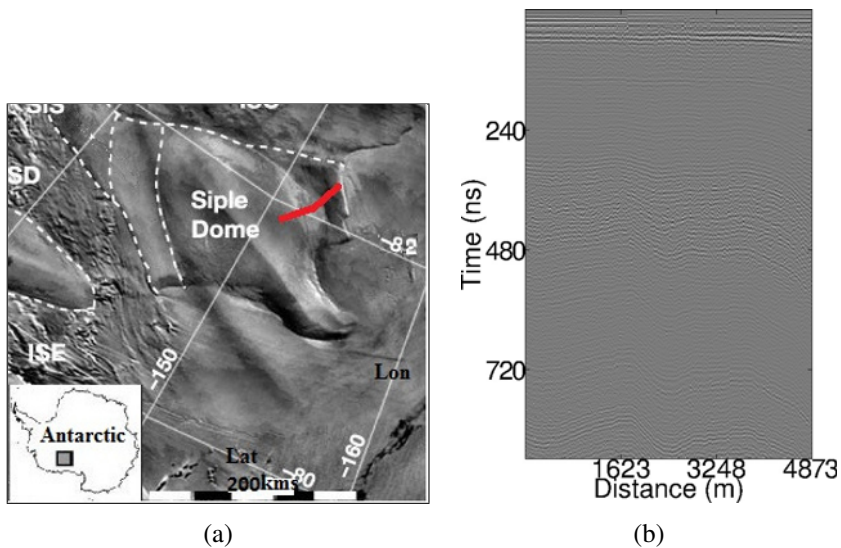
Fig. 12: (a) Map of location of Siple Dome in eastern Antarctica, Antarctic Explores; (b) Input GPR data 100 MHz acquired in Siple Dome.
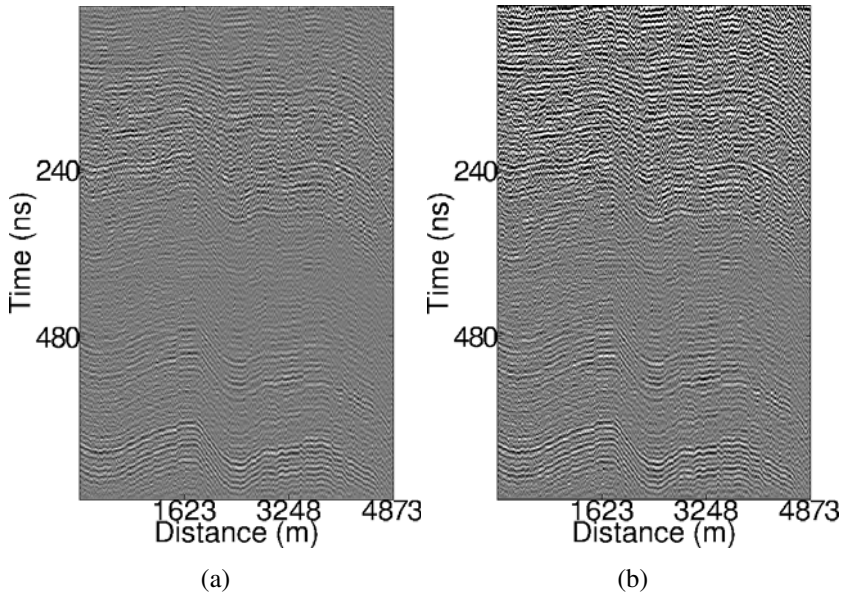
Fig. 13. Application of (a) amplitude-preserving, and (b) true-amplitude redatuming operator to profile of Fig. 12b.
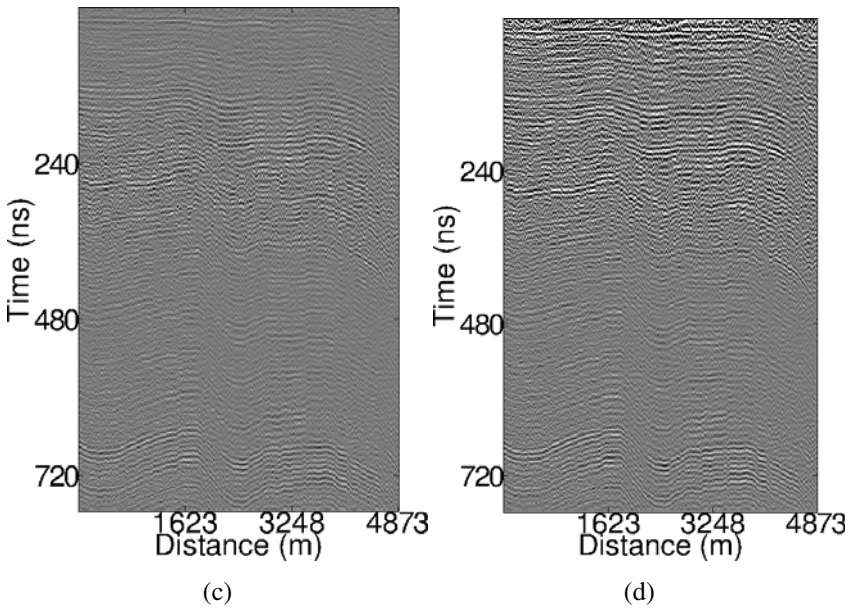
Fig. 14. Illustration of using an inacurate velocity field for (a) amplitude-preserving, and (b) true-amplitude redatuming.

is more evident on the shallow part due to the bigger ratio between the acquired and adjusted geometrical spreading factors.

On a second example, in order to simulate the use of a non accurate velocity field, we chose a 50% higher redatuming velocity (0.24 m/ns) when compared with ice velocity. As showed in the previous sections, Fig. 14 illustrates the impact of velocity error on both redatuming operators. Note that Figs. 13a and 14a present similar amplitude responses. On the other hand, the amplitude response of the true-amplitude operator is affected by the velocity error (see Figs. 13b and 14b).

As it can be seen, in both cases (with correct and wrong velocity) the amplitude at the deepest layers is highlighted for TAKR operation. In other words, the events presented better lateral continuity when compared with the TK result. We can observe these characteristics on the first events on the top of the data, on the left side around 300 ns, and at the end in 700 ns.

5. CONCLUSIONS

In this work, a comparative analysis was performed on two Kirchhoff redatuming operators: conventional and true-amplitude. When formulating both operators as variations of the same Kirchhoff integral operator, it become obvious that the difference between them is strictly dynamical and is due to a geometrical spreading correction factor at the stacking weight. This factor is responsible for replacing the input geometrical spreading by one adjusted to the new measurement surface. This difference is illustrated by numerical examples and one GPR field data set application. Using these examples, we illustrated that both operators fulfil their purposes, either preserving or adjusting the amplitudes.

Velocity sensitiveness analyses were also performed both analytically and numerically. We demonstrated that the true-amplitude operator is more sensitive to inaccuracies in the velocity field. While the conventional (amplitude-preserving) operator amplitude errors remain low, the true-amplitude operator amplitude errors increase (reaching 27% when in presence of 40% of velocity error).

The feasibility of our results were demonstrated through their application in GPR data acquired in a profile in Siple Dome-Antarctica. The TAKR and KR application to the GPR data showed that, in both cases, the reflectors were better delineated, presented better lateral continuity and the improved resolution of the main events, specially when the layers were deeper.

R e f e r e n c e s

Alkhalifah, T., and C. Bagaini (2006), Straight-rays redatuming: A fast and robust alternative to wave-equation-based datuming, *Geophysics* **71**, 3, U37–U46, DOI: 10.1190/1.2196032.

Alley, R., and C. Bentley (1988), Ice-core analysis on the Siple Coast of West Antarctica, *Ann. Glaciol.* **11**, 1-7.

Berryhill, J.R. (1984), Wave-equation datuming before stack, *Geophysics* **49**, 11, 2064-2066, DOI: 10.1190/1.1441620.

Bleistein, N. (1986), Two-and-one-half dimensional in-plane wave propagation, *Geophys. Prospect.* **34**, 5, 686-703, DOI: 10.1111/ j.1365-2478.1986.tb00488.x.

Bradford, J. (1999), Characterizing shallow aquifers with wave-propagation based geophysical methods: Imaging and attribute analysis, Ph.D. Thesis, Rice University, Houston, USA, http://hdl.handle.net/1911/18730.

Cassidy, N.J. (2007), Evaluating LNAPL contamination using GPR signal attenuation analysis and dielectric property measurements: Practical implications for hydrological studies, *J. Contam. Hydrol.* **94**, 1-2, 49-75, DOI: 10.1016/j.jconhyd.2007.05.002.

Cox, M. (1999), *Static Corrections for Seismic Reflection Surveys*, Geophysical References, Vol. 9, Society of Exploration Geophysicists, Tulsa.

Jin, D., P. Yan, Q.-S. Tang, H.-B. Zheng, Y.-L. Wang, Q.-J. Lin, and J. Chen (2011), Application of Kirchhoff integral wave field extrapolation to water layer datuming for OBC record, *J. Trop. Oceanogr.* **30**, 6, 84-89.

Khwanmuang, W., and S. Udphuay (2012), Ground-penetrating radar attribute analysis for visualization of subsurface archaeological structures, *Leading Edge* **31**, 8, 946-949, DOI: 10.1190/tle31080946.1.

Klüver, T. (2009), Wavefield separation of dual-sensor towed streamer data using Kirchhoff type datuming and migration operators. **In**: *79th SEG Annual Meeting, 25-30 October 2009, Houston, USA*, 2944-2948.

Liu, L., K. He, X. Xie, and J. Du (2007), Image enhancement with wave-equation redatuming: application to GPR data collected at public transportation sites, *J. Geophys. Eng.* **4**, 2, 139-147, DOI: 10.1088/1742-2132/4/2/003.

Margrave, G.F., and R.J. Ferguson (1999), Wavefield extrapolation by nonstationary phase shift, *Geophysics* **64**, 4, 1067-1078, DOI: 10.1190/1.1444614.

Oliveira, F., M. Pila, A. Novais, J. Costa, and J. Schleicher (2009), 2.5D True-amplitude diffraction-stack redatuming: numerical tests. **In**: *Proc. 11th Int. Congr. of the Brazilian Geophysical Society "Seismic Processing: Migration"*, 1-6.

Pila, M.F., J. Schleicher, A. Novais, and T.A. Coimbra (2014), True-amplitude single-stack redatuming, *J. Appl. Geophys.* **105**, 95-111, DOI: 10.1016/j.jappgeo.2014.03.010.

Schmalz, B., and B. Lennartz (2002), Analyses of soil water content variations and GPR attribute distributions, *J. Hydrol.* **267**, 3-4, 217-226, DOI: 10.1016/S0022-1694(02)00152-X.

Schneider, W.A. (1978), Integral formulation for migration in two and three dimensions, *Geophysics* **43**, 1, 49-76, DOI: 10.1190/1.1440828.

Tygel, M., J. Schleicher, P. Hubral, and L.T. Santos (1998), 2.5-D true amplitude Kirchhoff migration to zero offset in laterally inhomogeneous media, *Geophysics* **63**, 2, 557-573, DOI: 10.1190/1.1444356.

Wiggins, J.W. (1988), Attenuation of complex water-bottom multiples by wave-equation-based prediction and subtraction, *Geophysics* **53**, 12, 1527-1539, DOI: 10.1190/1.1442434.

Zhao, W., E. Forte, M. Pipan, and G. Tian (2013), Ground penetrating radar (GPR) attribute analysis for archaeological prospection, *J. Appl. Geophys.* **97**, 107-117, DOI: 10.1016/j.jappgeo.2013.04.010.

Remote Sensing Monitoring of Volcanic Ash Clouds Based on PCA Method

Chengfan LI, Yangyang DAI, Junjuan ZHAO, Shiqiang ZHOU, Jingyuan YIN, and Dan XUE

School of Computer Engineering and Science, Shanghai University, Shanghai, China; e-mail: david-0904@163.com (corresponding author)

A b s t r a c t

Volcanic ash clouds threaten the aviation safety and cause global environmental effects. It is possible to effectively monitor the volcanic ash cloud with the aid of thermal infrared remote sensing technology. Principal component analysis (PCA) is able to remove the inter-band correlation and eliminate the data redundancy of remote sensing data. Taking the Eyjafjallajokull volcanic ash clouds formed on 15 and 19 April 2010 as an example, in this paper, the PCA method is used to monitor the volcanic ash cloud based on MODIS bands selection; the USGS standard spectral database and the volcanic absorbing aerosol index (AAI) are applied as contrasts to the monitoring result. The results indicate that: the PCA method is much simpler; its spectral matching rates reach 74.65 and 76.35%, respectively; and the PCA method has higher consistency with volcanic AAI distribution.

Key words: remote sensing, principal component analysis (PCA), volcanic ash cloud, monitoring, aviation safety.

1. INTRODUCTION

The volcanic ash cloud, formed by the volcanic eruption, is usually consisted of pyroclastic particles, water vapor, SO_2 and HCL, *etc.* These components can be combined to form acidic aerosols when they erupt into the atmosphere troposphere, and cause global environmental problems (Wen and Rose 1994, Flynn *et al.* 2001, Ellrod 2004, Thomas *et al.* 2005, Andronico *et al.* 2009, Mastin *et al.* 2009, Webley and Mastin 2009). Currently, the civil air-

crafts' flight height is mostly 6-15 km and basically belongs to the troposphere. Once a large number of volcanic ash clouds appear in the stratosphere layer, it will easily cause aviation security risks. Remote sensing can obtain the real-time information of volcanic ash cloud. Thereinto, the infrared remote sensing has been rapidly developed since the infrared light was discovered in the year 1800. At present, there are several thermal infrared satellite remote sensing sensors which have been widely used in the field of air pollution, sea surface brightness temperature and volcanic activity monitoring (Qin *et al.* 2001, Jiménez-Muñoz and Sobrino 2003, Liu *et al.* 2003, Qu *et al.* 2006, Yao *et al.* 2007, Krueger 1983, McCarthy *et al.* 2008, Filizzola *et al.* 2007, Rose *et al.* 2008). In view of the significant impact of massive volcanic ash cloud on the global environment and aviation safety, how to detect volcanic ash cloud from remote sensing data has become one of the most important tasks in the volcanoes and security realms.

Principal component analysis (PCA) is an effective mathematical method whose aim is to transform multiple indicators into a few composite indicators using dimensionality reduction. It has been widely adopted in data compression, feature extraction and pattern identification fields (Hillger and Clark 2002a, b, Hillger and Ellrod 2003). The monitoring of volcanic ash cloud produced by the large-scale volcanic eruptions has been carried out since the 1980s. Zhu *et al.* (2011) have used FY-3A remote sensing data to present shortwave infrared-thermal infrared volcanic ash (STVA) method, and have achieved a higher detecting precision. Hillger and Clark (2002a, b) have processed the moderate resolution imaging spectroradiometer (MODIS) data of Popocatepetl and Cleveland volcanic ash clouds with the PCA method, and have accurately identified the volcanic ash cloud distribution in the lower atmosphere. The essence of volcanic ash cloud monitoring is also a pattern recognition problem, and along with a certain relevance and data redundancy coming from each band of remote sensing data, although the PCA method has great application potential, there are relatively fewer studies on the volcanic ash cloud monitoring using PCA method, and the research achievements are few.

According to the physical-chemical properties of volcanic ash cloud, taking the MODIS data of Eyjafjallajokull volcanic ash clouds on 15 and 19 April 2010 as the data source, this paper uses the PCA method to detect volcanic ash cloud with the aim to provide a reference for other volcanic ash cloud monitoring researches. Section 1 introduces the general situation and research status of volcanic ash cloud, Section 2 describes PCA basic principles and its realization process in remote sensing image, Section 3 is the detailed detection process of volcanic ash cloud from MODIS images using PCA method, Section 4 gives detection results, and Section 5 presents the conclusions and discussions.

2. THEORETICAL BASIS

PCA method is a kind of dimensionality reduction method whose assumption is that a given set of relevant variables is converted into another group of irrelevant variables by linear transformation, and these new variables are ordered sequentially in accordance with the variance in descending order (Lin *et al.* 2012, Zhao *et al.* 2012).

2.1 PCA principles

Suppose x is m-dimensional random vector, and the $E[x] = 0$, the realization process of PCA is shown below.

Step 1. Calculate: $Q = [x_1 - x, x_2 - x, \ldots, x_m - x]_{n \times m}$,   in which $x = \dfrac{1}{m} \sum\limits_{j=1}^{m} x_j$,

and then calculate matrix $R = Q^T Q$ according to Q.

Step 2. Calculate s maximum eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_s$ and corresponding orthogonal eignvectors $v_1, v_2, \ldots, v_s$ of R.

Step 3. Calculate the eigenvectors $w_1, w_2, \ldots, w_s$ of covariance matrix $C = QQ^T/m$. Hereinto, $w_d = Qv_d / \sqrt{\lambda_d}$, $d = 1, 2, \ldots, s$, and $W = [w_1, w_2, \ldots, w_s]$.

Step 4. Calculate the feature $Y = [y_1, y_2, \ldots, y_m]$ of x, where $y_p = W^T x_p$, $p = 1, 2, \ldots, m$.

2.2 PCA realization process in remote sensing image

According to PCA principles, the first principal component (PC 1) is required to have the largest possible variance, the second principal component (PC 2) is computed under the constraint that it is orthogonal to PC 1, and so forth (Lin *et al.* 2012). Next, this study takes the multispectral remote sensing images with N bands as an example, and then describes the realization process of PCA method in multispectral remote sensing images. Hereinto, the PCs obtained from remote sensing images using PCA method are called principal component images (PCIs) (Hillger and Clark 2002a, b, Li *et al.* 2011).

Let E be the transformation matrix, and B the multispectral remote sensing image with N bands, so the relationship between PCIs and the multispectral remote sensing images can be explained as follows:

$$\text{PCI} = E \cdot B = E \cdot \begin{bmatrix} b_{11} & b_{12} & \ldots & b_{1m} \\ b_{21} & b_{22} & \ldots & b_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ b_{n1} & b_{n2} & \ldots & b_{nm} \end{bmatrix} = E \cdot \begin{bmatrix} B_1 \\ B_2 \\ \ldots \\ B_n \end{bmatrix} \tag{1}$$

and, furthermore, we can rewrite the PCI as follows:

$$\begin{cases} \text{PCI}_1 = e_1 B = e_{11} B_1 + e_{12} B_2 + \ldots + e_{1n} B_n \\ \text{PCI}_2 = e_2 B = e_{21} B_1 + e_{22} B_2 + \ldots + e_{2n} B_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \text{PCI}_n = e_n B = e_{n1} B_1 + e_{n2} B_2 + \ldots + e_{nn} B_n \end{cases} \tag{2}$$

Similarly, the first principal component image (PCI 1) has the maximum variance and contains the maximum commonalities of each multispectral remote sensing image. Other PCIs often lay emphasis on the difference among the original remote sensing images, and this difference in a single original image is invisible. Besides, the noise (also called blemishes) appears primarily in the n-th principal component image (PCIn).

3.   REMOTE SENSING MONITORING OF VOLCANIC ASH CLOUD

3.1 General situation of study area

Iceland is located east of Greenland and immediately south of the Arctic Circle. Although its area is only $103\,000$ km^2, it is known as the Fire Island of Polar Circle. The Eyjafjallajokull volcano is located to the southeast of the Reykjavik, and began to erupt in the evening on 20 March 2010. Subsequently, the Eyjafjallajokull volcano erupted violently again on 14 April 2010. The volcanic ash cloud spread to Continental Europe and North Atlantic by wind, causing most European airports to be closed; more than 1000 flights were canceled, and the direct financial loss amounted to about hundreds of billions of euros. According to the relevant documentation and meteorological conditions, the MODIS images on 15 and 19 April 2010 are eventually used in this study in the monitoring of the volcanic ash cloud (Fig. 1).

(a)   (b)   (c)

Fig. 1: (a) MODIS image of Eyjafjallajokull volcanic ash cloud on 15 April 2010, (b) MODIS image of Eyjafjallajokull volcanic ash cloud on 19 April 2010, and (c) enlarged Eyjafjallajokull volcanic ash cloud, in which the red point is the Eyjafjallajokull caldera.

Figures 1a and b show the false color composite image of MODIS data band 2, 1, 3 acquired on 15 and 19 April 2010, respectively; the corresponding center wavelengths of three corresponding bands are 0.859, 0.645, and 0.469 µm. From Figs. 1a and b it is seen that the color of volcanic ash cloud is close to the one of the sea and land surface, so it is difficult to distinguish the volcanic ash cloud; it depends entirely on visual inspection. Figure 1c shows an enlarged image of the volcanic ash cloud and the volcanic caldera acquired from MODIS thermal infrared remote sensing images. As can be seen from Figs. 1a and c, when the volcanic ash cloud rushes upward and spreads southward and eastward by wind, the volcanic ash cloud components become thinner and spread out of the caldera's periphery. However, the spreading volcanic ash cloud does not cover the volcanic caldera and the caldera is clear in Fig. 1c.

3.2 PCA on the MODIS thermal infrared bands

Due to the method of PCA processing and effective spectral bands selection with MODIS, the thermal infrared bands on 15 and 19 April 2010 are the same; this paper just takes the case of the volcanic ash cloud on 19 April 2010. The MODIS sensor is an important remote sensor mounted in the Terra and Aqua satellites whose data can be freely received and used by worldwide users. MODIS sensor has a total of 36 discrete spectral bands, the spectral range covers the bands from visible to thermal infrared; its spatial resolutions are 250, 500, and 1000 m, respectively, and the maximum scan width is 2330 km. Spectral sensitivity is the relative efficiency of the signal as a function of the frequency of the signal; signal to noise ratio (SNR) is a measure that compares the level of a desired signal to the level of background noise. Both are an indication of the relative magnitude of the signal in these bands of the MODIS sensor, and are often used to evaluate its performance (Table 1).

In recent years, many researchers have been exploring the monitoring methods of volcanic ash cloud using different thermal infrared remote sensors, especially the advanced, very high resolution radiometer (AVHRR) remote sensor. Because the MODIS sensor is an updated product of AVHRR, how to reasonably use MODIS data to obtain volcanic ash cloud should be considered in the future study. In order to verify the thermal infrared band's sensitivity for the volcanic ash cloud in MODIS data (bands 20-36, except of band 26), this study adopts PCA method to process the MODIS thermal infrared bands. According to PCA principle, when the 16 thermal infrared bands are inputted, the corresponding 16 separated PCIs are obtained. The PCIs on 19 April 2010 are shown in Fig. 2.

Generally speaking, the principal component information in PCIs is also called explained variance. The PCI 1 has the biggest principal component in-

Table 1

The MODIS thermal infrared bands distribution and performance

Bands	Resolution [m]	Band width [μm]	Spectral sensitivity [W/m^2·μm·sr]	SNR [k]
20	1000	3.660-3.840	0.45	0.05
21	1000	3.929-3.989	2.38	2.00
22	1000	3.929-3.989	0.67	0.07
23	1000	4.020-4.080	0.79	0.07
24	1000	4.433-4.498	0.17	0.25
25	1000	4.482-4.549	0.59	0.25
27	1000	6.535-6.895	1.16	0.25
28	1000	7.175-7.475	2.18	0.25
29	1000	8.400-8.700	9.58	0.05
30	1000	9.580-9.880	3.69	0.25
31	1000	10.780-11.280	9.55	0.05
32	1000	11.770-12.270	8.94	0.05
33	1000	13.185-13.485	4.52	0.25
34	1000	13.485-13.785	3.76	0.25
35	1000	13.785-14.085	3.11	0.25
36	1000	13.085-14.385	2.08	0.35

formation of the original remote sensing images. As can been seen from Fig. 2, PCI 1 contains the vast majority of components' information. In this study, the image background mainly includes meteorological clouds and land cover information:

❑ For a meteorological cloud, the higher the meteorological cloud, the brighter the color. This is due to the different absorbed heat caused by varying height of the meteorological cloud. When the cloud is higher, its brightness temperature in the range of thermal infrared band will be lower and the color will be brighter due to the absorption of smaller amount of the earth radiation heat. When the cloud is lower, on the contrary, its brightness temperature is higher and the color is darker because of the absorption of much of the earth radiation heat;

❑ For land cover information, the gray color of light means low brightness temperature and black means high brightness temperature; the gray color represents gradual changes of the brightness temperature between light gray and black. In PCIs, its color gradually changes from light gray to black; this indicates that the surface brightness temperature of land cover information changes from low to high.
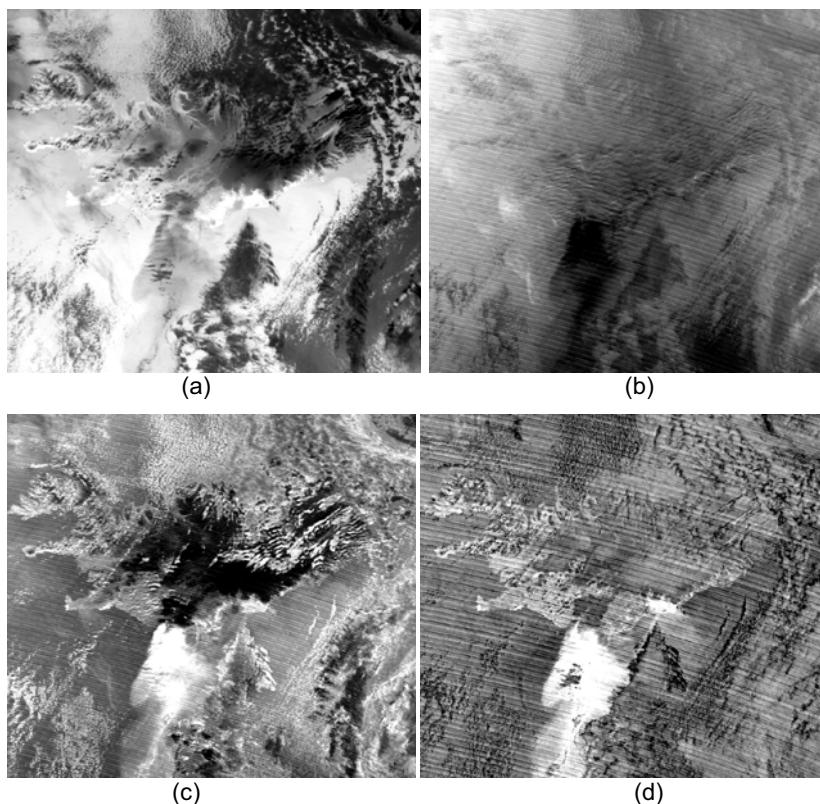
Fig. 2. PCIs of MODIS thermal infrared band acquired on 19 April 2010; (a), (b), (c), and (d) are PCI 1, 2, 4, and 8 images, respectively, other PCIs are basically noise information and are omitted.

For the obtained PCIs, the volcanic ash cloud information is detected clearly only in PCI 1, 2, 4, and 8; as other PCIs are basically full of noise, they are all ignored in the study (Fig. 2). In addition, from the viewpoint of a contrast between volcanic ash cloud and background feature, PCI 4 is clear with significant contrast in all PCIs, it is followed by PCI 1 and 8; and PCI 2 has less contrast. Even so, the distributions of volcanic ash cloud information are basically detected in the four PCIs.

3.3 Selection of effective spectral bands

In order to give a quantitative analysis of the detected volcanic ash cloud using PCI 1, 2, 4, and 8, the explained variance and SNR of different PCIs are calculated. According to the principle of diminishing amount, the PCIs contrasts of volcanic ash cloud are shown in Table 2.

Table 2

PCIs contrasts of volcanic ash cloud

PCI	Contrast effect	Explained variance [%]	SNR
4	Good ash signal	0.13	1.66
1	Good ash signal, but cloud contamination	0.20	1.33
8	Good ash signal, noisy	0.008	4.55
2	Densest ash only, low contrast	0.011	1.58

As can be seen from Table 2, the volcanic ash cloud information in the PCI 4 is that the cloud is obviously bright and has the biggest contrast to the image background and the best detection effect. The following one is PCI 8; although the contrast decreases and is between the volcanic ash cloud and the image background, the volcanic ash cloud can still be clearly detected; the volcanic ash cloud showing dark tones in PCI 1, the cloud shows bright tones, the land shows darker tones; the volcanic ash cloud shows dark tones in PCI 2, and the other features are basically in light tones. In addition, from the view of explained variance and SNR, the explained variance and SNR of PCI 4 have the best contrast effect, and reach 0.13 and 1.66, respectively, the explained variance and SNR of PCI 2 have the least obvious contrast effect, and reach 0.011 and 1.58, respectively. Thus, it has not been confirmed that the greater the explained variance and SNR value, the better the volcanic ash cloud contrast effect.

Table 3 shows the contributions from MODIS bands to PCI with a better volcanic ash cloud detection effect. For PCI 4, the contribution rate of band 30 is the largest, it is followed by band 36; for PCI 1, the contribution rate of band 36 is the largest, followed by band 25; for PCI 8, the contribution rate of bands 31 and 32 are the largest, respectively, the contribution rate of other bands is relatively smaller; for PCI 2, the contribution rate of bands 29 and 32 are the largest. Taken together, for

Table 3

Contributions from MODIS bands to PCIs [%]

PCI	20	25	29	30	31	32	33	34	36
4	2	9	0	52	0	0	1	1	17
1	3	13	3	10	2	0	2	5	55
8	1	0	6	0	64	11	1	0	0
2	0	0	32	3	1	26	0	0	0
Total	6	22	41	65	67	37	4	6	72

PCI 9, 8, 16, and 14, band 36 has the largest contribution to the volcanic ash cloud with the total contribution rate of 72%; followed by bands 31 and 30, respectively, whose total contribution rates reach 67 and 65%.

In addition, different combinations of PCIs are also able to reveal the object feature under different circumstances. For example, PCI 9 is the essential difference between ozone band and long-wave CO_2 band (corresponding to 9.7 and 14.2 μm); PCI 8 is mainly the difference between short-wave and long-wave CO_2 band (corresponding to 4.52 and 14.2 μm); PCI 16 is the difference between long wave window band and nearby dirty window band (corresponding to 11.0 and 12.0 μm); PCI 14 is the difference between water vapor band and dirty window band (corresponding to 8.6 and 12.0 μm). For different application fields, we are required to select the suitable PCIs in terms of the actual situation.

3.4 PCA on selected effective bands

Through the above analysis, the thermal infrared bands 30, 31, and 36 in MODIS have the biggest contribution to the volcanic ash cloud PCIs. Next, taking the original bands 30, 31, and 36 of MODIS sensors as the data source, this study will detect the Eyjafjallajokull volcanic ash cloud using 3-band PCA method on 15 and 19 April 2010, respectively. Figure 3 shows the selected MODIS images on 15 April 2010 and the corresponding PCIs. Figure 4 shows the selected MODIS images on 19 April 2010 and the corresponding PCIs.

From Figs. 3a-c and 4a-c it is seen that some fractus clouds have lower temperature in the original band 36 and show light tones, while some other, stratiform clouds have higher equivalent black body temperature (TBB) and show gray color. This is so because the volcanic ash cloud in the early eruption period and diffusion process contains a lot of heat and its brightness temperature is higher than that of TBB, which is more conventional. The brightness temperature characteristics of volcanic ash cloud in original bands 30 and 31 are similar to that of band 36. During the day, the brightness temperature of volcanic ash cloud is higher and the concentration is greater than that of meteorological clouds; all these factors make the image appear to be in dark tones.

From Figs. 3d-f and 4d-f it follows that the distinction between the volcanic ash cloud and the image background in PCI 2 is most obvious. The volcanic ash cloud is gray and has a clearly visible edge, relatively uniform texture; it also highlights the texture and spectral characteristics of volcanic ash cloud. In stark contrast, the surrounding white scattered clouds have fragmentary and chaotic texture. Although PCI 3 contains a certain amount of noise, the volcanic ash cloud information is highlighted. Due to the carried heat and radiation heat from itself, the volcanic ash cloud's brightness
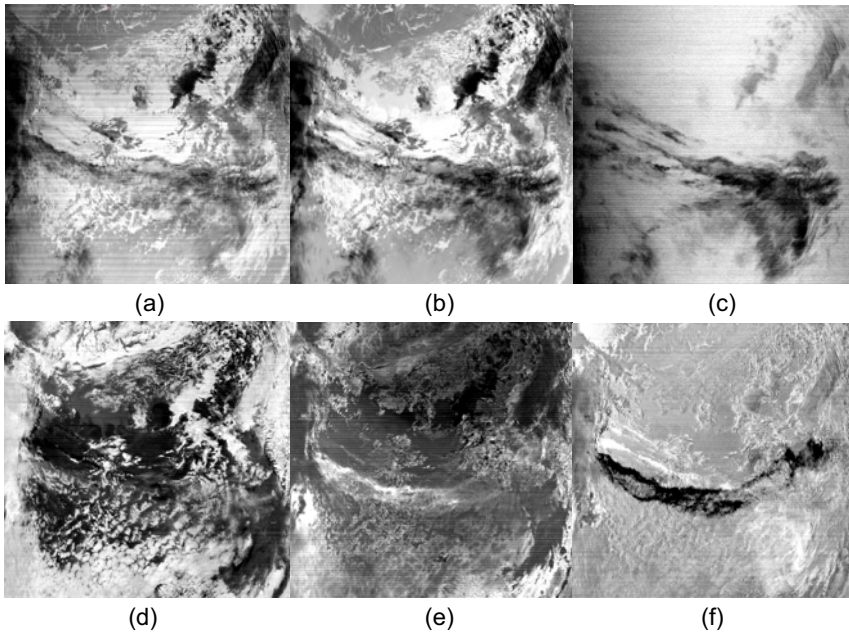
Fig. 3. The selected MODIS images and PCIs on 15 April 2010: (a)-(c) are the original bands 30, 31, and 36, respectively; (d)-(f) are PCI 1, 2, and 3, respectively.
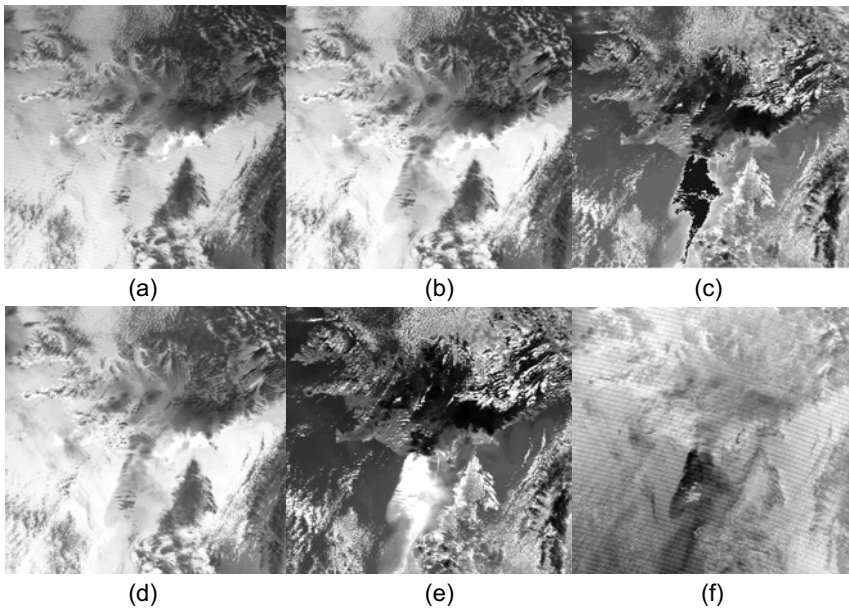
Fig. 4. The selected MODIS images and PCIs on 19 April 2010: (a)-(c) are the original bands 30, 31, and 36, respectively; (d)-(f) are PCI 1, 2, and 3, respectively.
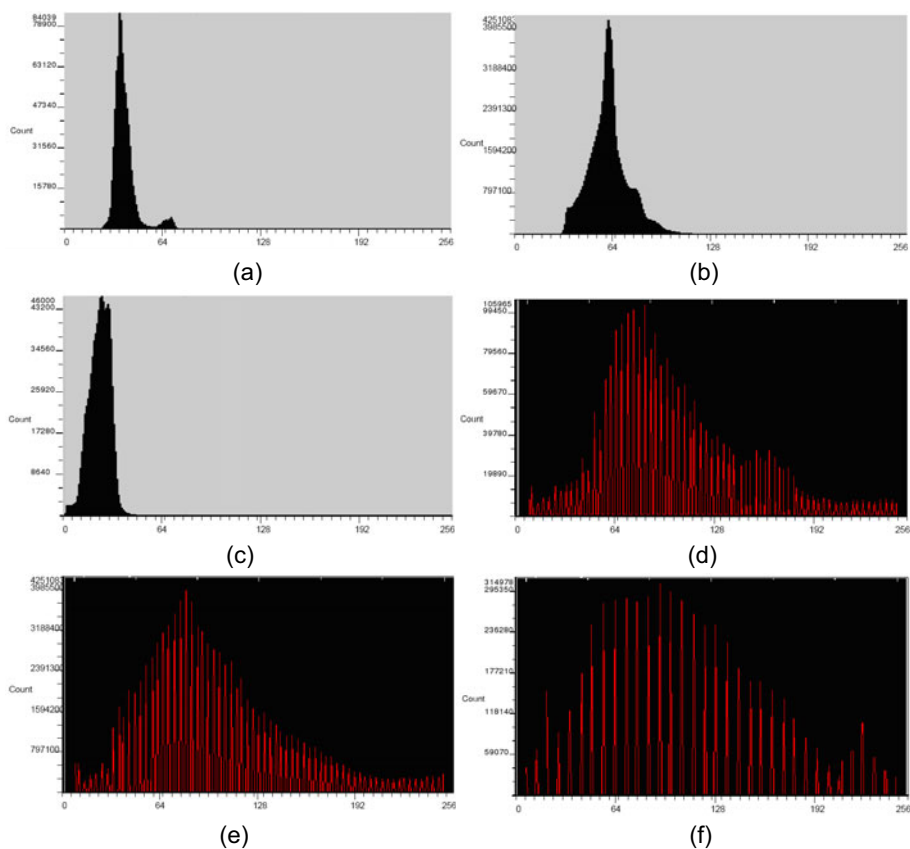
Fig. 5. The color histogram of the original bands and PCIs on 15 April 2010: (a)-(c) are the original bands 30, 31, and 36, respectively; (d)-(f) are corresponding PCI 1, 2, and 3, respectively.

temperature is higher than that of the surrounding land and cloud, and appears to be in dark tones.

In order to more accurately evaluate the difference between the original multispectral remote sensing image and PCIs, in this study, the color histogram of the original bands 30, 31, and 36 in MODIS and the corresponding PCI 1, 2, and 3 on 1**5** April 2010 (Fig. 5) and 19 April 2010 (Fig. 6) are calculated. Figures 5a-c and 6a-c show that the color histogram of the bands 30, 31, and 36 is very similar, and the three bands have high inter-band correlation. On the contrary, the color histogram of PCIs in the Figs. 5d-f and 6d-f is completely different and has different spatial distribution. This indicates that the PCIs by PCA processing can effectively remove the inter-band correlation among the original remote sensing images and greatly enhance the subsequent detection of volcanic ash cloud.
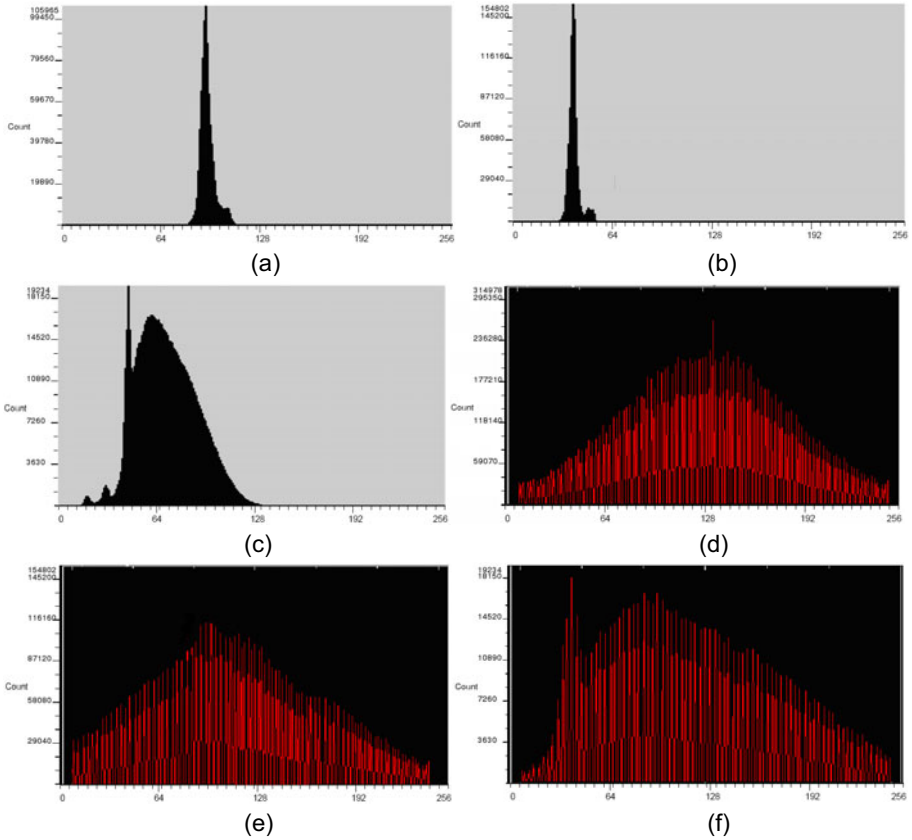
Fig. 6. The color histogram of the original bands and PCIs on 19 April 2010: (a)-(c) are the original bands 30, 31, and 36, respectively; (d)-(f) are corresponding PCI 1, 2, and 3, respectively.

3.5 PCIs noise elimination and segmentation

To some extent, the PCIs can basically reflect the distribution of the volcanic ash cloud. In order to more clearly highlight the volcanic ash cloud information in the practical applications, a further detection of the volcanic ash cloud from PCIs is necessary. The specific process is as follows:

❏ Noise elimination of volcanic ash cloud. In some cases, there are some minor noises in PCIs. In order to overcome the problem, the 3×3 filter processing is used to process the three obtained PCIs, respectively. Studies show that when the covering area of the multispectral remote sensing data is large enough, the statistics law of each band and its linear processing results in multispectral data basically present the normal distribution (Moghtaderi *et al.* 2007).

❑ False color synthesis. In order to display the volcanic ash cloud information more clearly, in this study, the obtained PCIs which are processed by noise elimination were given red (R), green (G), and blue (B) colors, respectively, and synthesized.

❑ Color density division. According to the multivariate statistical analysis method, the density division can get the best effect when the divided rule is that the mean value of false color synthesis PCIs plus n times of the standard deviation. Combined with previous research experience and many experiments, it is the ideal way to divide the density into three classes in this study, and the three class rules are 2.0, 2.5, and 3.0 times of the standard deviation. The detected volcanic ash cloud information on 15 and 19 April 2010 is shown in Fig. 7.

Figure 7 shows that the volcanic ash cloud on 19 April 2010 mainly spread towards the south, and then it spread towards the east under the action of gravity and wind. For Class 1, 2, and 3, the concentration of volcanic ash cloud composition within Class 1 region is the greatest; followed by Class 2 region, where the concentration is mediate; the concentration of volcanic ash cloud composition in Class 3 region is the thinnest. In addition, from the distribution point of volcanic ash cloud information, Class 3 is the most widely distributed one in the three classes of volcanic ash cloud region, it is followed by Class 2; and Class 1, which has the smallest distribution area. Furthermore, Class 1 distributes around the center of volcanic ash cloud; with the decrease of volcanic ash cloud concentration, Class 2 mainly locates around the Class 1, and the Class 3 distributes the peripheral region of volcanic ash cloud. In view of this, the usage of the PCA method yields a better
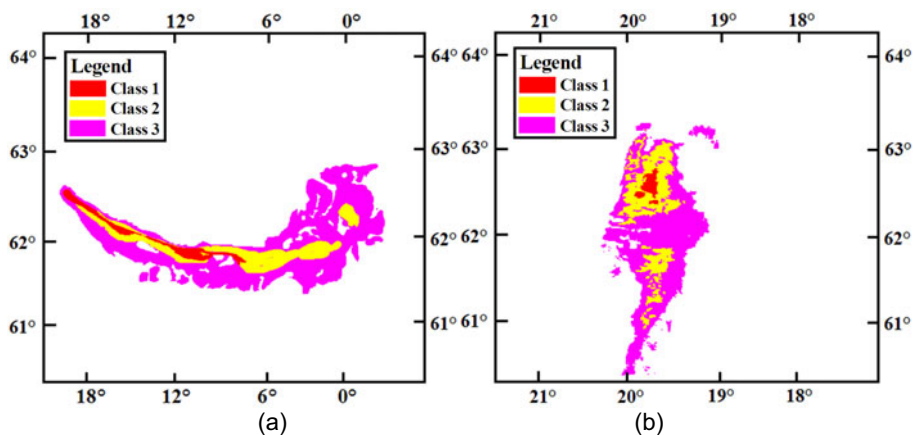
Fig. 7. Detected volcanic ash cloud of Eyjafjallajokull volcano: (a) 15 April 2010, and (b) 19 April 2010.

detecting effect. This method is able to detect not only the shape characteristics of the volcanic ash cloud information but also the different concentration distribution of volcanic ash cloud components. And what is more, the contrast between the obtained volcanic ash cloud and PCIs background characteristics in PCIs is obvious.

4. MONITORING RESULTS

In order to verify the detecting effects using PCA method, a comparative analysis was performed in this study on the spectral matching of volcanic ash cloud and volcanic absorbing aerosol index (AAI).

4.1 Spectral matching of volcanic ash cloud

In this study, the mineral spectral information of the United States Geological Survey (USGS) standard spectral database is adopted. According to the wavelength interval of MODIS data, the reflectance data in the USGS standard spectral database has been resampled to correspond to MODIS bands, and its spectral curve is established corresponding to MODIS data. Similarity matching between the spectral curve after the atmospheric correction and the standard spectral library of mineral ash after resampling is shown in Fig. 8.
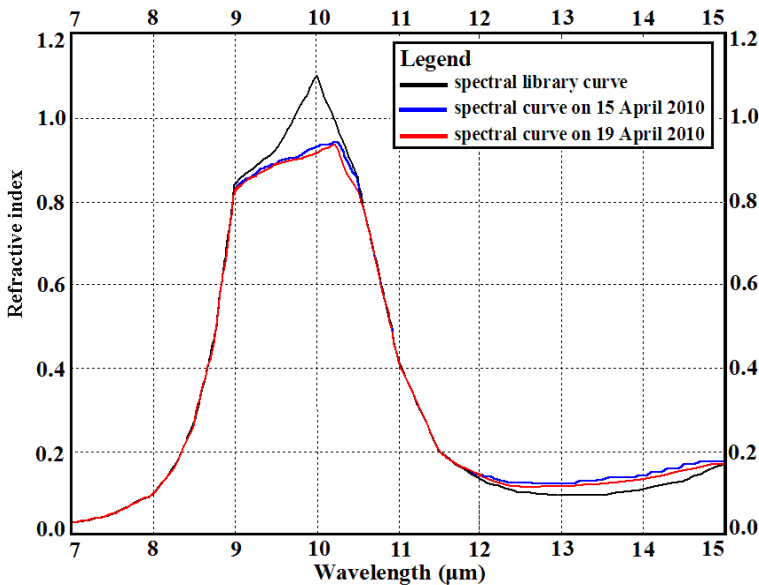
Fig. 8. The spectral curve matching rate of Eyjafjallajokull volcanic ash cloud on 15 and 19 April 2010.

As can be seen from Fig. 8, the matching rate between the spectral curve and spectral library of volcanic ash cloud reached 74.65% on 19 April 2010 and 76.35% on 15 April 2010, respectively. It shows that the volcanic ash cloud detection using PCA method has achieved good results. In addition, there are some significant differences within a certain wavelength area, such as 9-11 μm. Based on the analysis, it follows that the main cause is the absorption spectrum differences of thermal infrared bands in these wavelength ranges. And what is more, this is also consistent with the selecting basis of bands 30, 31, and 36 in this study.

Although the actual situation nearby the Eyjafjallajokull volcano, for example the snow-covered surface, ocean surrounding, atmospheric attenuation and terrain differences, reduces the detection precision of the volcanic ash cloud to some extent, the PCA method can still separate out the different object feature information from the mixed information in original remote sensing images based on decorrelation and elimination redundancy. Therefore, the PCA method contributes to detect volcanic ash cloud from the remote sensing image, and makes the detected volcanic ash cloud information closer to the actual distribution.

4.2 Volcanic AAI

The volcanic AAI is applied in this study, and it is calculated by the absorbing characteristics of volcanic ash cloud in ultraviolet bands. The AAI on 15 and 19 April 2010 is acquired by the 0.34 and 0.38 μm data of the global ozone monitoring experiment (GOME-2) sensor carried on the Metop-A satellite, and shown in Fig. 9.
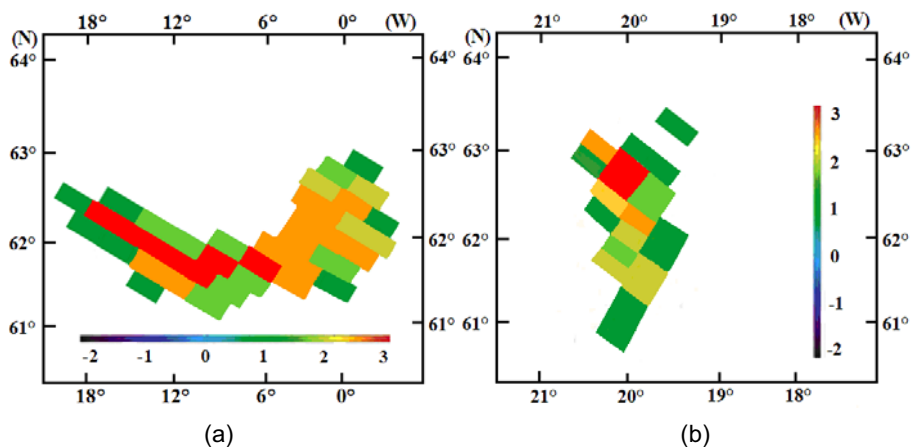
Fig. 9. The AAI of Eyjafjallajokull volcanic ash cloud: (a) 15 April 2010, and (b) 19 April 2010.

Figure 9 shows that the maximum concentration position of volcanic AAI basically appears near the volcano and is close to the centric line of volcanic ash cloud. In this paper, the distribution of volcanic ash cloud image on 15 and 19 April 2010 obtained by PCA method is similar to the AAI distribution; it accurately reflects not only the main part of the volcanic ash cloud distribution, but also a small amount of volcanic ash cloud distribution. Due to the lack of the ground measured data of the corresponding region, the volcanic ash cloud acquired by MODIS sensor is used to compare with that of volcanic AAI distributions. The results show that the proposed PCA method can get good monitoring effect of the volcanic ash cloud, and the monitored results have a high consistency and fitness with the volcanic AAI distribution.

5. CONCLUSIONS AND DISCUSSIONS

The PCA method not only removes the inter-band correlation and eliminates the data redundancy among the different remote sensing images, but also highlights the characteristics of different objects and improves the detection accuracy of thematic information. In this paper, the PCA method is used to detect the Eyjafjallajokull volcanic ash cloud from MODIS images on 15 and 19 April 2010. Thereinto, the volcanic ash cloud has been detected with the aid of selection of effective spectral bands, PCA processing, noise elimination, and segmentation. And then, the spectral matching rate and volcanic AAI have been adopted to assess the detected volcanic ash cloud. The results show that PCA method has good effect in the detection of volcanic ash cloud, whose spectral matching rate of volcanic ash spectral reaches 74.65 and 76.35%, respectively, and has high consistency with AAI distribution. The PCA method, which is used to detect the volcanic ash cloud from MODIS remote sensing images, has a certain theoretical significance and better practical values for volcanic activity monitoring and aviation safety. In comparison with other research methods (Hillger and Clark 2002a, b, Zhu *et al.* 2011), the PCA method is much simpler and has a certain detection precision.

Because different types of volcanic eruptions have different volcanic ash cloud components, the formed absorption spectrum characteristics are not exactly the same either. Although PCA method has a huge advantage in the field of volcanic ash cloud detection, it is necessary to select the suitable remote sensing data and wavebands in terms of the actual situation of study area and volcanic eruptions. In addition, compared to PCA method, independent component analysis (ICA) method has some more obvious advantages in the field of remote sensing data processing and volcanic activity monitoring (Qu *et al.* 2006, Li *et al.* 2013). It can not only remove the inter-band correlation among the different bands of remote sensing data, but also get mutually independent components, which has significant applications in vol-

canic ash cloud detection. However, the related researches are rarely involved, and this needs to be further discussed.

References

Andronico, D., C. Spinetti, A. Cristaldi, and M.F. Buongiorno (2009), Observations of Mt. Etna volcanic ash plumes in 2006: An integrated approach from ground-based and polar satellite NOAA-AVHRR monitoring system, *J. Volcanol. Geoth. Res*. **180**, 2-4, 135-147, DOI: 10.1016/j.jvolgeores. 2008.11.013.

Ellrod, G.P. (2004), Impact on volcanic ash detection caused by the loss of the 12.0 µm "Split Window" band on GOES imagers, *J. Volcanol. Geoth. Res*. **135**, 1-2, 91-103, DOI: 10.1016/j.jvolgeores.2003.12.009.

Filizzola, C., T. Lacava, F. Marchese, N. Pergola, I. Scaffidi, and V. Tramutoli (2007), Assessing RAT (Robust AVHRR Techniques) performances for volcanic ash cloud detection and monitoring in near real-time: the 2002 eruption of Mt. Etna (Italy), *Remote Sens. Environ*. **107**, 3, 440-454, DOI: 10.1016/j.rse.2006.09.020.

Flynn, L.P., A.J.L. Harris, and R. Wright (2001), Improved identification of volcanic features using Landsat 7 ETM+, *Remote Sens. Environ*. **78**, 1-2, 180-193, DOI: 10.1016/S0034-4257(01)00258-9.

Hillger, D.W., and J.D. Clark (2002a), Principal component image analysis of MODIS for volcanic ash. Part 1: Most important bands and implications for future GOES imagers, *J. Appl. Meteor*. **41**, 10, 985-1001, DOI: 10.1175/ 1520-0450(2002)041<0985:PCIAOM>2.0.CO;2.

Hillger, D.W., and J.D. Clark (2002b), Principal component image analysis of MODIS for volcanic ash. Part II: Simulation of current GOES and GOES-M imagers, *J. Appl. Meteor*. **41**, 10, 1003-1010, DOI: 10.1175/1520-0450 (2002)041<1003:PCIAOM>2.0.CO;2.

Hillger, D.W., and G.P. Ellrod (2003), Detection of important atmospheric and surface features by employing principal component image transformation of

GOES imagery, *J. Appl. Meteor*. **42**, 5, 611-629, DOI: 10.1175/1520-0450 (2003)042<0611:DOIAAS>2.0.CO;2.

Jiménez-Muñoz, J.C., and J.A. Sobrino (2003), A generalized single-channel method for retrieving land surface temperature from remote sensing data, *J. Geophys. Res*. **108**, D22, 4688-4695, DOI: 10.1029/2003JD003480.

Krueger, A.J. (1983), Sighting of El Chichón sulfur dioxide clouds with the Nimbus 7 total ozone mapping spectrometer, *Science* **220**, 4604, 1377-1379, DOI: 10.1126/science.220.4604.1377.

Li, C.F., J.Y. Yin, J.S. Dong, and D. Shen (2013), Monitoring of volcanic ash cloud based on thermal infrared satellite remote sensing, *Infrared Technol*. **35**, 8, 487-491.

Li, J., Z.G. Han, H.B. Chen, Z.L. Zhao, and H.Y. Wu (2011), Detection of heavy fog events over North China Plain by using the geostationary satellite data, *Remote Sens. Technol. Appl*. **26**, 2, 186-195 (in Chinese).

Lin, C., F. Peng, B.H. Wang, W.F. Sun, and X.J. Kong (2012), Research on PCA and KPCA self-fusion based MSTAR SAR automatic target recognition algorithm, *J. Electron. Sci. Technol*. **10**, 4, 352-357.

Liu, Z.W., A,R. Dang, Z.D. Lei, and Y.G. Huang (2003), A retrieval model of land surface temperature with ASTER data and its application study, *Progr. Geogr*. **22**, 5, 507-514 (in Chinese).

Mastin, L.G., M. Guffanti, R. Servranckx, P. Webley, S. Barsotti, K. Dean, A. Durant, J.W. Ewert, A. Neri, W.I. Rose, D. Schneider, L. Siebert, B. Stunder, G. Swanson, A. Tupper, A. Volentik, and C.F. Waythomas (2009), A multidisciplinary effort to assign realistic source parameters to models of volcanic ash-cloud transport and dispersion during eruptions, *J. Volcanol. Geoth. Res*. **186**, 1-2, 10-21, DOI: 10.1016/j.jvolgeores.2009.01.008.

McCarthy, E.B., G.J.S. Bluth, I.M. Watson, and A. Tupper (2008), Detection and analysis of the volcanic clouds associated with the 18 and 28 August 2000 eruptions of Miyakejima volcano, Japan, *Int. J. Remote Sens*. **29**, 22, 6597-6620, DOI: 10.1080/01431160802168400.

Moghtaderi, A., F. Moore, and A. Mohammadzadeh (2007), The application of advanced space-borne thermal emission and reflection (ASTER) radiometer data in the detection of alteration in the Chadormalu paleocrater, Bafq region, Central Iran, *J. Asian Earth Sci*. **30**, 2, 238-252, DOI: 10.1016/j.jseaes.2006.09.004.

Qin, Z.H., A. Karnieli, and P. Berliner (2001), A mono-window algorithm for retrieving land surface temperature from Landsat TM data and its application to the Israel–Egypt border region, *Int. J. Remote Sens*. **22**, 18, 3719-3746, DOI: 10.1080/01431160010006971.

Qu, C.Y., X.J. Shan, and J. Ma (2006), Application of satellite thermal infrared re-
mote sensing in detection of volcano activity, *Seismol. Geol.* **28**, 1, 99-110.

Rose, W.I., S. Self, P.J. Murrow, C. Bonadonna, A.J. Durant, and G.G.J. Ernst
(2008), Nature and significance of small volume fall deposits at composite
volcanoes: Insights from the October 14, 1974 Fuego eruption, Guatemala,
Bull. Volcanol. **70**, 9, 1043-1067, DOI: 10.1007/s00445-007-0187-5.

Thomas, W., T. Erbertseder, T. Ruppert, M. van Roozendael, J. Verdebout, D. Balis,
C. Meleti, and C. Zerefos (2005), On the retrieval of volcanic sulfur dioxide
emissions from GOME backscatter measurements, *J. Atmos. Chem.* **50**, 3,
295-320, DOI: 10.1007/s10874-005-5544-1.

Webley, P., and L. Mastin (2009), Improved prediction and tracking of volcanic ash
clouds, *J. Volcanol. Geoth. Res.* **186**, 1-2, 1-9, DOI: 10.1016/j.jvolgeores.
2008.10.022.

Wen, S.M., and W.I. Rose (1994), Retrieval of sizes and total masses of particles in
volcanic clouds using AVHRR bands 4 and 5, *J. Geophys. Res.* **99**, D3,
5421-5431, DOI: 10.1029/93JD03340.

Yao, Y.J., P. Nan, Z.L. Zhang, and B.S. Li (2007), Application of split window al-
gorithm in land surface temperature retrieval from thermal infrared remote
sensing data, *J. Lanzhou Univ. Technol.* **33**, 6, 89-92 (in Chinese).

Zhao, Q., Z.L. Xie, H. Li, and X.L. Li (2012), Color-feature extraction of remote
sensing image based on principal components analysis and K-means, *Mi-
croelectron. Comput.* **29**, 10, 61-68 (in Chinese).

Zhu, L., J. Liu, C. Liu, and M. Wang (2011), Satellite remote sensing of volcanic
ash cloud in complicated meteorological conditions, *Sci. China Earth Sci.*
54, 11, 1789-1795, DOI: 10.1007/s11430-011-4265-3.

Influences of Al, Mg, and K on Calculating Concentrations of Other Formation Elements and Their Corrections in Geochemical Logging

Wensheng WU[1], Maosong TONG[2], Wei NIU[1], and Li LUO[3]

[1]State Key Laboratory of Petroleum Resource and Prospecting,
China University of Petroleum, Beijing, China; e-mail: wwsheng@cup.edu.cn

[2]Daqing Drilling Corporation, PetroChina, Daqing, China

[3]CCDC Well Logging Company, Chongqing, China

A b s t r a c t

In geochemical logging, the standard capture gamma-ray spectra of Al, Mg, and K have no distinct characteristic energy peaks. This feature easily influences the spectral bands of characteristic peaks of other elements and thus affects the accuracy of calculating their concentrations. To study this influence, we constructed a model formation containing Al, Mg, and K, and employed the Monte Carlo N Particle Transport Code (MCNP) program to simulate their capture spectra. The results indicate that the calculated dry weights of Si are almost free from the influences of Mg, K, and Al. The ones of Ca and Fe are influenced, but only to a minor extent. The dry weight of S is concurrently influenced. Specifically, the K concentration causes S concentration to deviate significantly from its real concentration. By correcting for such influences, we can obtain relatively accurate values. Data processing for a real well validates the finding that Mg, Al, and K in the formation influence the precision of calculation of other elements and also testifies to the effectiveness of the correction method.

Key words: geochemical elemental logging, spectral stripping, elemental concentration.

1. INTRODUCTION

In geochemical logging, fast neutrons emitted from neutron source enter the formation to collide with nuclides in the formation, leading to three kinds of prompt nuclear reactions, *i.e.*, inelastic scattering, capture, and particle reactions. In all three cases, the resultant nucleus in an excited state decays and yields gamma rays with energies which are characteristic of the particular nucleus in the formation. Therefore, we can discriminate types of formation elements by detecting gamma-ray energy and determine the elemental content by measuring the intensity of specific-energy gamma rays. Because the measurement spectrum of a formation can be assumed as a linear superposition of the standard spectra of all elements in the formation in certain portions which add up to 1, with the standard elemental gamma-ray spectra, we can directly quantitatively analyze the mixed gamma-ray spectra of the measured formation to obtain the element's relative yields which are not the actual elemental weight percent or dry weight in the formation and cannot be directly used to evaluate the formation's mineralogy. The conversion factor between elemental relative yield and its dry weight is just the detection sensitivity factor (or sensitivity factor in short). Then, according to the sensitivity factor of each element in the formation and the formation oxide closure model, the weight percent or dry weight of every major element in the rock matrix is obtained, so as to further identify mineral composition of the rock matrix. In actual logging, the internal media in a well will influence the measured formation spectrum. When the internal fluid in a well has a high mineral salt concentration, the available formation signal in the measured spectrum degrades, since the elemental Cl in the fluid is a major neutron absorber, whereas the statistical uncertainty increases. Barite in the internal fluid in a well is a gamma-ray absorber and will deform the measured spectrum. So, the calculated relative yields need to be corrected for these influences before they are used to calculate elemental dry weights (Ellis and Singer 2008, Helmer *et al.* 1967, Hertzog *et al.* 1989, Herron 1986, Chapman *et al.* 1987, Grau and Schweitzer 1989, Galford *et al.* 2009).

When the source is Am-Be-source, geochemical logging usually utilizes the capture gamma-ray spectra to gain the elemental relative yields and dry weights. The main elements measured with this source include silicon, calcium, iron, sulfur, titanium, gadolinium, aluminum, magnesium, and potassium, among which silicon, calcium, iron, sulfur, titanium, and gadolinium are the six important elements indicative of formation minerals, having large capture cross-sections and distinct characteristic peaks in standard spectra, so their concentrations can be directly obtained from analysis of the capture gamma-ray spectra. Hydrogen and chlorine neutron capture signals are also detected, but they are not used for mineralogical evaluations.

One important reason for calculating concentrations of different elements through quantitative analysis of energy spectra is the differences in standard spectra of different elements. The more unique a feature an element's standard capture gamma-ray spectrum has, the better the measurability of the element. For elements with similar standard gamma-ray spectra, it is very difficult to calculate their concentrations through energy spectral analysis precisely. The three skeleton elements, *i.e.*, aluminum, magnesium, and potassium, have no distinct characteristic peaks in their standard spectra, making it difficult to obtain their concentrations from analysis of mixed formation spectra against their standard spectra; the standard capture gamma-ray spectra of these three elements are shown in Fig. 1. As can be seen from Fig. 1, neither Mg nor K has a very distinct characteristic peak. Al has no distinct characteristic peak other than a characteristic peak at 7.70 MeV, which is very close to 7.64 and 7.28 MeV, the strong energy peaks of Fe. Therefore, during Am-Be-source chemical elemental logging, the concentrations of elements Mg, Al, and K in the formation cannot be obtained through an analysis of capture spectra; instead, they have to be obtained by using other gamma-ray energy spectra logging methods. The new tool solutions, which are committed to improve the signal-to-noise ratio of formation responses by reducing undesired borehole and tool contributions, increase the energy resolution of spectra important for better differentiation of elements, which solves the identification of Al, Mg, and K to some extent (Galford *et al.* 2009, Radtke *et al.* 2012).
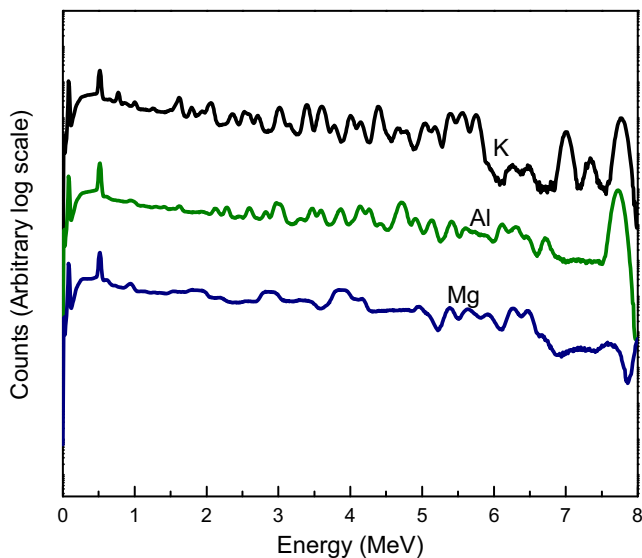
Fig. 1. Capture spectra of elements Mg, Al, and K.

MCNP5C, which can track photons according to the originating isotope, was used to carry out calculations for zero- and medium-porosity sandstone, limestone, and dolomite formations. On the basis of the detector response function, the MCNP5 simulation formation spectral responses for a geochemical tool can produce the best match to laboratory measurements made in quarried and fabricated rock formations. So, in this paper we employ the MCNP5 program (Sweezy *et al*. 2004) to study the influences of Al, Mg, and K on the results of analysis of the four important elements, Si, Ca, S, and Fe, in some model formation for GEM tool (Galford *et al.* 2009) that introduces additional hardware innovations to improve the signal-to-noise ratio of formation responses by reducing undesired borehole and tool contributions. The simulation results are used for developing corresponding correction methods. The processing of real well data will validate the effectiveness of the corrections.

2.   FORMATION NOT CONTAINING MG, AL, AND K

The simulation model is a cylinder with a height of 2000 mm and a diameter of 1600 mm. The borehole well diameter is 200 mm, and the in-hole fluid is fresh water. The GEM tool that consists of A standard chemical americium-beryllium neutron source and a large bismuth germanate (BGO) detector located above the source is put against the well wall. Since the tally mode F8 in MCNP5 is the tally of electronic pulse energy amplitude that records gamma ray energy deposition in formation tally cell, the count by F8 is more suitable for spectral logging, and close to the count of the actual logging instrument compared to other tally modes. Here F8 is used to record the induced capture gamma-ray spectra of BGO detector.

The formation swas deemed to be one entity filled with different volumes of SiO_2, $CaCO_3$, $FeCO_3$, $CaSO_4$, and H_2O; the major skeleton elements in the formation are Si, Ca, Fe, and S, as shown in Table 1.

Table 1

Volume contents of minerals in strata containing Si, Ca, Fe, and S

Model	Density [g/cm^3]	Mineral volume content [%]				
		SiO_2	$CaCO_3$	$CaSO_4$	$FeCO_3$	H_2O
1	2.589	50	25	10	5	10
2	2.690	45	30	15	5	5
3	2.639	60	15	10	7	8
4	2.597	65	20	5	3	7
5	2.577	30	50	3	6	11
6	2.548	20	45	20	2	13
7	2.545	27	26	12	15	20
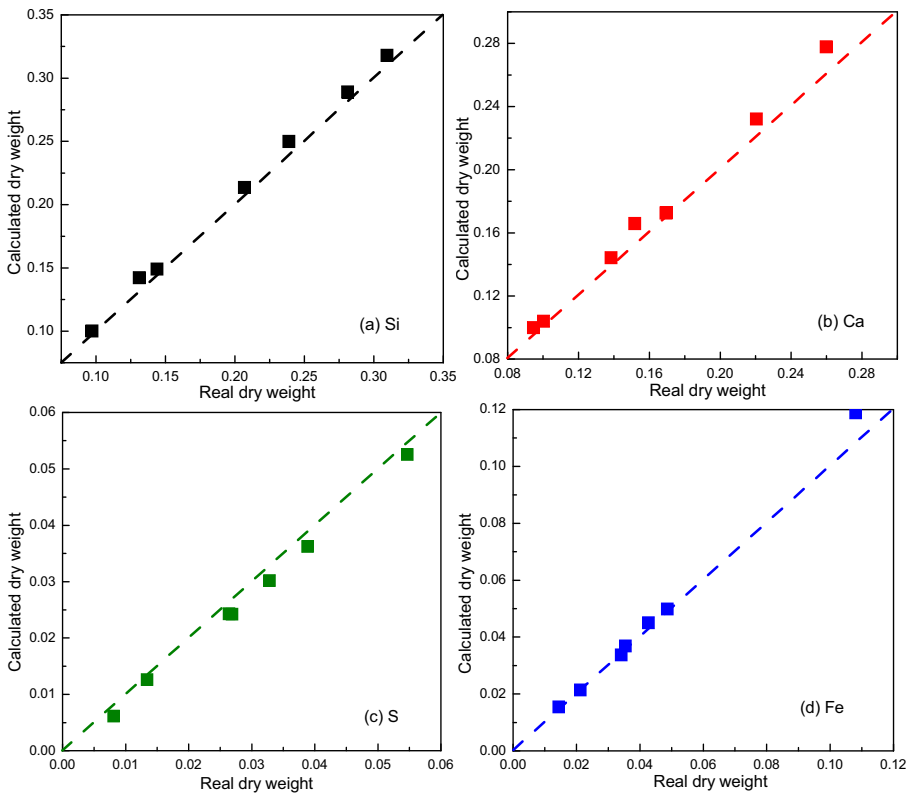
Fig. 2. Comparison between real dry weight and calculated dry weight of formation elements in Table 1.

A constrained and weighted least-squares method is used to analyze simulated capture gamma-ray spectra of seven formations listed in Table 1 to get elemental relative yields because the method can ensure that physically feasible solutions are obtained for the individual contributions, or elemental yields. We obtained the relative yields of four elements, Si, Ca, Fe, and S; then, according to the detection sensitivity factors of these four elements and depth normalization factors of the seven formations (Ellis and Singer 2008, Pang 1991, Pemper *et al.* 2006, Herron *et al.* 2002), we calculated the dry weights of Si, Ca, Fe, and S. Dry weights of four elements obtained from spectral analysis were compared with real dry weights of various elements in the formation, as shown in Fig. 2, in which the diagonal dashed lines are the equivalent lines of calculated values and real values of the formation. As can be seen in Fig. 2, the dry weights of Si, Ca, Fe, and S obtained from the spectral analysis have generally high precision, and their correlation coefficients relative to real values reach more than 0.998. Such high-precision
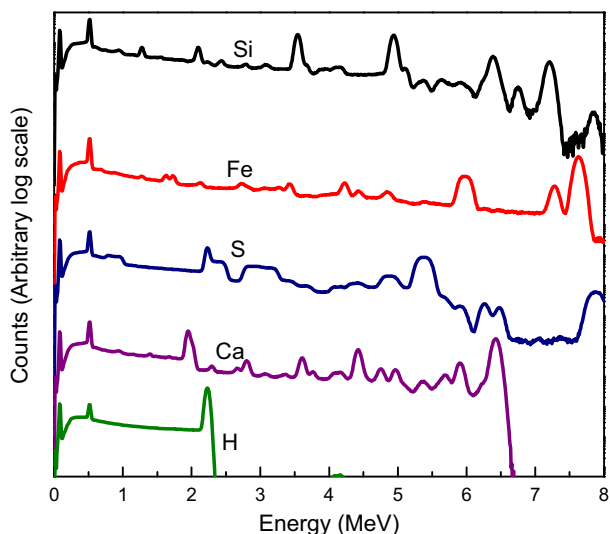
Fig. 3. Capture spectra of elements H, Si, Ca, Fe, and S.

spectral analysis results are attributed to the quantitative analysis method and to the important fact that, in standard spectra of the major elements H, Si, Ca, Fe, and S, there is little interference among characteristic peaks. For example, Ca's full-energy peaks at 1.94 and 4.42 MeV, Si's full-energy peak at 3.54 MeV, S's full-energy peak at 5.40 MeV, and Fe's full-energy peaks at 5.92 and 7.64 MeV are very distinct, with independent peak positions and high intensities, as shown in Fig. 3.

3. SPECTRAL ANALYSIS RESULTS OF THE FORMATION CONTAINING MG, AL, AND K

Mg-bearing formation

The formation volume was deemed to be 1, with different volumes of SiO_2, $CaCO_3$, $FeCO_3$, $CaSO_4$, and H_2O. A certain volume of dolomite $(CaMg(CO_3)_2)$ was added to the formation. The mineral volume contents of the formation models are shown in Table 2. Skeleton elements in these formation models include H, Si, Ca, S, Fe, and Mg.

According to the standard spectrum and detection sensitivity factor of each element and depth normalization factor of each formation, the mixed neutron-capture gamma-ray spectra of these formations were analyzed to get the dry weight of each element. Because Mg has no distinct characteristic peak, its concentration cannot be obtained from analysis of neutron-capture gamma-ray energy spectra. The dry weight of Mg has to be determined from

Volume contents of minerals in complicated Mg-bearing strata

Model	Density [g/cm³]	Mineral volume content [%]					
		SiO_2	$CaCO_3$	$CaMg(CO_3)_2$	$CaSO_4$	$FeCO_3$	H_2O
1	2.571	70	6	6	6	3	9
2	2.625	65	8	14	5	2	6
3	2.654	60	10	9	8	6	7
4	2.663	55	12	17	9	2	5
5	2.555	50	15	12	7	4	12
6	2.590	45	18	16	8	3	10
7	2.652	40	20	25	5	3	7
8	2.716	35	28	23	6	4	4
9	2.732	30	39	4	11	10	6
10	2.677	25	18	35	4	8	10
11	2.596	20	40	10	3	12	15
12	2.784	15	30	30	15	6	4
13	2.668	10	11	60	7	2	10
14	2.629	5	60	13	7	5	10

the lithology index P_e for litho-density logging. Assuming that the dry weight of Mg in the formation has been obtained, we can calculate the dry weights of elements Si, Ca, Fe, and S in various formations listed in Table 2 by using an analytical algorithm. The results are shown in Fig. 4.

As can be seen from Fig. 4, the three elements, Si, Ca, and Fe, in the Mg-bearing formation still have very high spectral analysis precision, with correlation coefficients between the calculated and real dry weights reaching about 0.99, but S has a distinctly lower spectral analysis precision, with the calculated value points distinctly deviated from the diagonal line. The analysis of the standard spectra of elements Si, Ca, and Fe shows that each standard spectrum has two or more strong characteristic energy peaks that can be easily identified. Mg has little influence on the characteristic peak counts of these three elements; thus, Si, Ca, and Fe still have relatively high spectral analysis precision, being almost immune to the influence of Mg. Nevertheless, it can be known from the standard spectra of S that the strongest characteristic peak is only one peak at 5.4 MeV, whereas the other characteristic peaks are relatively gentle or overlap with peaks of other elements, and the strongest characteristic peak of S at 5.4 MeV is of very weak contrast to those of Si, Ca, and Fe. Thus, the influence of the complicated and variable energy spectrum of Mg on the counts of the S characteristic peak will rela-
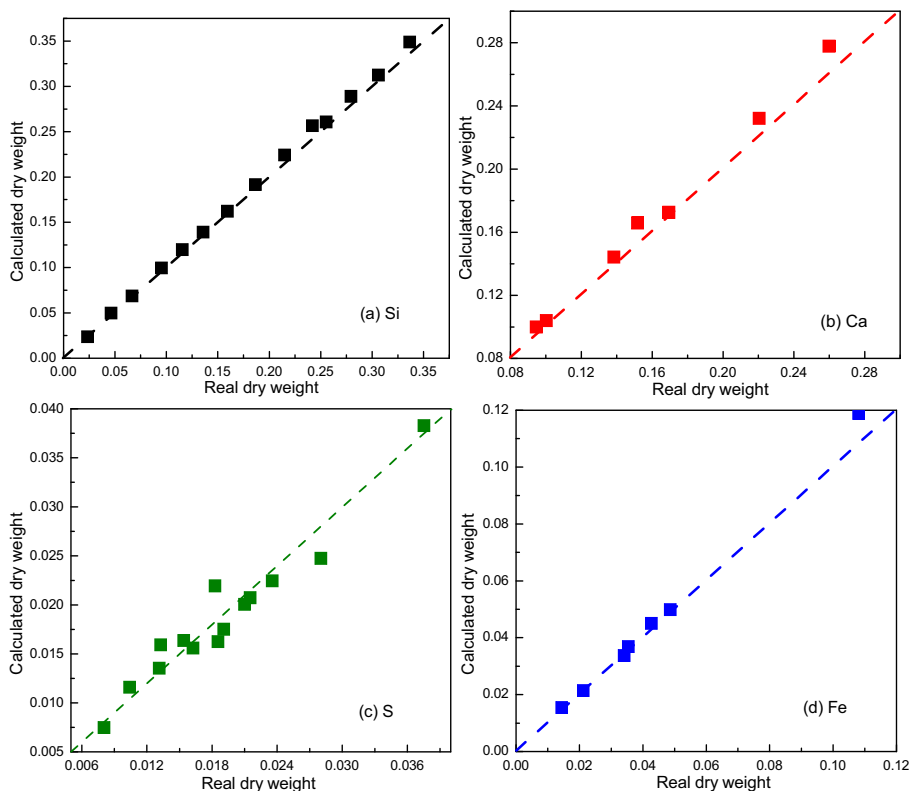
Fig. 4. Relationship of elemental concentration calculated from analysis of Mg-bearing formation *versus* real concentration.

tively increase, lowering the spectral analysis precision of S. However, since Mg appears relatively weak at 5.40 MeV, its influence on calculating S concentration is not great.

Formation containing Mg and Al

Deemed as one entity, the formation contains different volumes of SiO_2, $CaCO_3$, $FeCO_3$, $CaSO_4$, $CaMg(CO_3)_2$, and H_2O, plus a certain volume of anorthite ($CaAl_2Si_2O_8$). The mineral volume contents in the formation models are shown in Table 3. These formation models are based on those in Table 2, with Al added to their skeletons.

Assuming that concentrations of elements Mg and Al have been obtained, respectively, from the index P_e for litho-density logging and neutron aluminum activation logging (Herron and Herron 1996), we analyzed the mixed neutron-capture gamma-ray spectra of the above formation according

Table 3

Volume contents of minerals in Al-bearing strata

Model	Density [g/cm^3]	Mineral volume content [%]						
		SiO$_2$	CaCO$_3$	CaMg(CO$_3$)$_2$	CaAl$_2$Si$_2$O$_8$	CaSO$_4$	FeCO$_3$	H$_2$O
1	2.574	64	6	6	6	6	3	9
2	2.628	57	8	14	8	5	2	6
3	2.659	48	10	9	12	8	6	7
4	2.669	39	12	17	16	9	2	5
5	2.619	30	15	15	20	7	4	9
6	2.600	22	18	16	23	8	3	10
7	2.662	14	20	25	26	5	3	7
8	2.704	22	21	14	29	6	4	4
9	2.685	26	15	4	32	11	6	6
10	2.653	11	13	19	35	4	8	10
11	2.626	12	25	10	38	3	4	8
12	2.755	9	10	15	41	15	6	4
13	2.618	2	4	31	44	7	2	10
14	2.688	3	31	3	47	6	5	5

to the standard spectrum and detection sensitivity factor of each element and depth normalization factor of each formation, and we obtained the dry weight of each element, as shown in Fig. 5.

As can be seen from Fig. 5, Si concentration still has a relatively high spectral analysis precision. Compared with Fig. 4, at a high concentration of Ca, the calculated values deviate farther from the diagonal line and thus the spectral analysis precision decreases; for Fe, the calculated values deviate farther from the diagonal line, too, indicating that the spectral analysis precision of Fe concentration decreases. This can be explained from Fig. 6, which shows neutron-capture gamma-ray energy spectra of the SiO$_2$–CaCO$_3$ mixed formation and the SiO$_2$–CaAl$_2$Si$_2$O$_8$ mixed formation, where the Al-bearing stratum enables lower characteristic peaks of Ca, which causes the reduced spectral analysis precision of Ca concentration. The characteristic peak of Al at 7.70 MeV influences one characteristic peak of Fe, which is the cause of the reduced spectral analysis precision of Fe concentration. The lower the Fe concentration in the formation, the greater the influence of Al on the calculated Fe concentration. Compared with complicated Mg-bearing formation, the spectral analysis precision of S in the complicated formation containing Mg and Al does not vary much; as a whole, the calculated dry weight of S is kept consistent with its real dry weight. This indicates that, despite Mg and
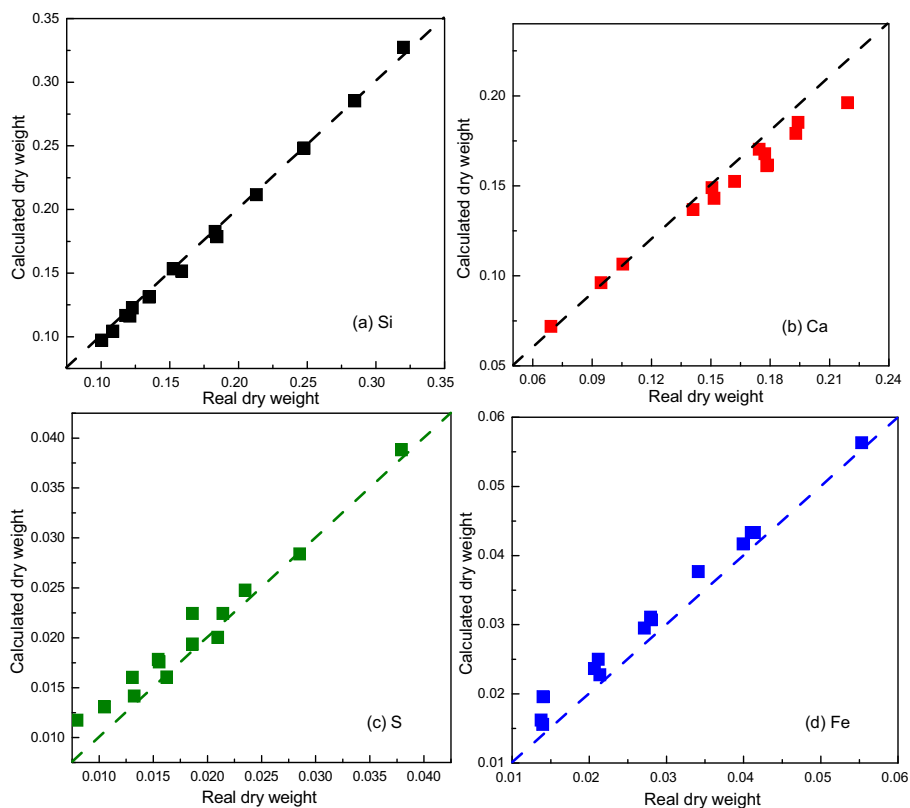
Fig. 5. Comparison between real dry weight and calculated dry weight of each element in Al-bearing formation.
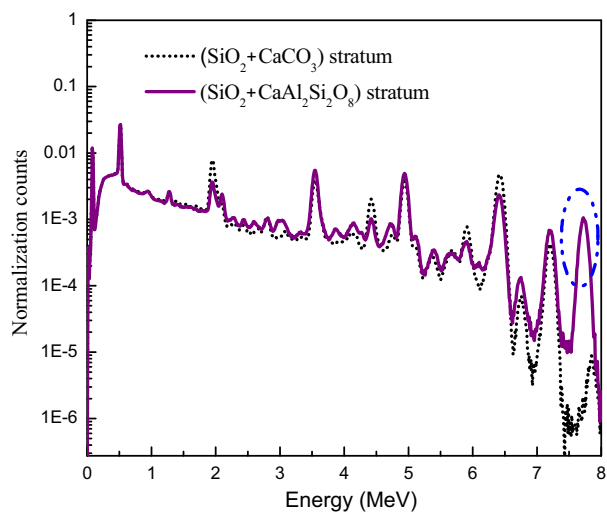
Fig. 6. Mixed capture energy spectra of the Al-bearing and non-Al-bearing formation.

Al inducing an interfering peak effect on S and lowering the spectral analysis precision of S such an influence is not especially great.

Formation containing Mg, Al, and K

Deemed as one entity, the formation contains different volumes of SiO_2, $CaCO_3$, $FeCO_3$, $CaSO_4$, $CaMg (CO_3)_2$, and H_2O, plus a certain volume of orthoclase ($KAlSi_3O_8$). The mineral volume contents of the various formation models are shown in Table 4. Besides H, Si, Ca, Fe, and S, the formation contains three more elements: Mg, Al, and K.

Table 4

Volume contents of minerals in complicated K-bearing strata

Model	Density [g/cm³]	Mineral volume content [%]						
		SiO_2	$CaCO_3$	$CaMg(CO_3)_2$	$KAlSi_3O_8$	$CaSO_4$	$FeCO_3$	H_2O
1	2.566	64	6	6	6	6	3	9
2	2.618	57	8	14	8	5	2	6
3	2.643	48	10	9	12	8	6	7
4	2.649	39	12	17	16	9	2	5
5	2.593	30	15	15	20	7	4	9
6	2.570	22	18	16	23	8	3	10
7	2.628	14	20	25	26	5	3	7
8	2.666	22	21	14	29	6	4	4
9	2.643	26	15	4	32	11	6	6
10	2.608	11	13	19	35	4	8	10
11	2.576	12	25	10	38	3	4	8
12	2.702	9	10	15	41	15	6	4
13	2.561	2	4	31	44	7	2	10
14	2.627	3	31	3	47	6	5	5

After the concentrations of K, Mg, and Al are gained, respectively, from natural gamma-ray spectral logging, the P_e logging and aluminum activation logging (Herron and Herron 1996), the neutron-capture gamma-ray energy spectra data of these formations were analyzed, and the obtained relative yields of elements were converted to dry weight of each respective element, as shown in Fig. 7.

Compared with Fig. 5, elements Si, Ca, and Fe in Fig. 7 have higher spectral analysis precision, with smaller differences between calculated dry weight and real one. This is because the energy spectra of Si, Ca, and Fe have many intense characteristic peaks, and any of these three elements in
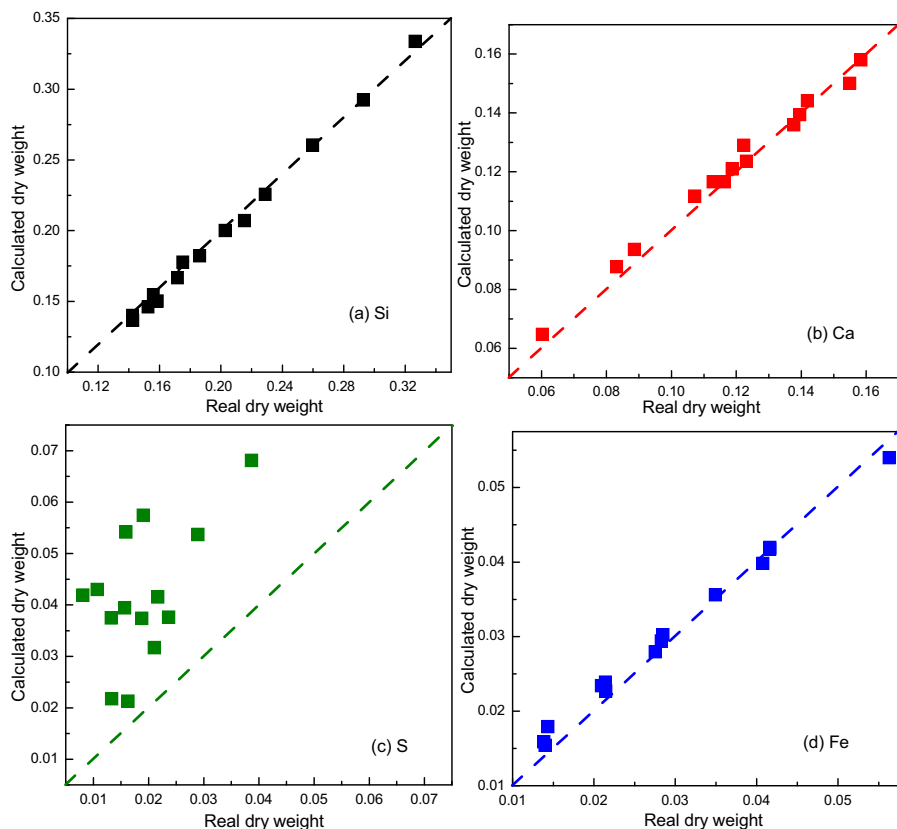
Fig. 7. Comparison between theoretical and real values of each element in compli-
cated K-bearing formation.

formation can generally keep at least one independent characteristic peak
immune to interference from other elements and less affected by K. Al has
one characteristic peak overlapping with one of Fe's characteristic peaks, but
it does not exert much influence on the spectral analysis result of Fe. Fig-
ure 7 clearly shows a greatly reduced spectral analysis precision of S, with
the correlation coefficient between calculated dry weight and real dry weight
falling to below 0.6, and calculated values of S are distinctly greater than the
real values, which are considered to be distorted. This can be explained from
Fig. 8, showing capture gamma-ray energy spectra of two formations: one
composed of SiO_2 and $CaCO_3$, and the other composed of SiO_2, $CaCO_3$, and
$KAlSi_3O_8$. The characteristic spectral band 5.0-6.0 MeV of K in the
K-bearing energy spectra induces intense interference upon S's characteristic
peak at 5.40 MeV and increases the characteristic peak count of S. In fact,
some of the counts of this enhanced characteristic peak of S originate from
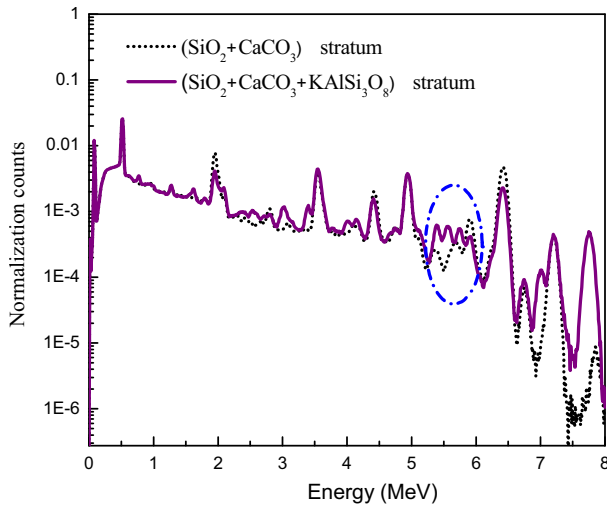
Fig. 8. Mixed capture energy spectra of K-bearing and non-K-bearing formations.

the characteristic peak count of K. For S, except for a distinct characteristic peak at 5.40 MeV, no other strong characteristic peak can control its spectra; thus, the spectral analysis result is greatly distorted.

4.   INFLUENCE CORRECTIONS

Taking Fig. 7c as an example where K distorts the spectral analysis accuracy of S, we subtracted the real dry weight of S in the formation from the calculated dry weight derived from spectral analysis to get the difference D, and with the difference D as the ordinate and the real dry weight of K in the formation as the abscissa, we obtained the result shown in Fig. 9. As can be seen from Fig. 9, with increasing K concentration, the difference between the calculated dry weight and the real dry weight of S exhibits a linear increase, indicating that this part of the difference is primarily caused by the influence of K concentration in the formation.

   A fitting of data points in Fig. 9 gives rise to a linear expression for the difference between calculated dry weight and real dry weight of S *versus* K concentration in the formation. With this relationship expression, we can correct the calculated dry weight of S according to the real dry weight of K in the formation. Correction of the calculated dry weight of S in Fig. 7c gives rise to the relationship between corrected S concentration and real dry weight, as shown in Fig. 10. Therefore, the corrected dry weight of S is relatively close to the real dry weight of S in the formation, indicating that this correction method is relatively effective. There is still a certain deviation be-
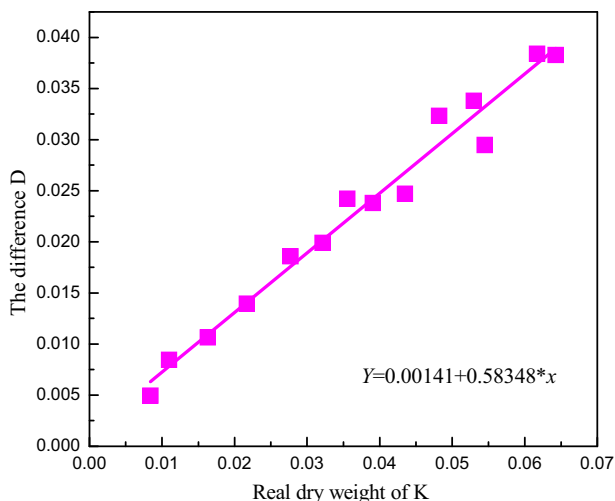
Fig. 9. Influence of K concentration on calculated value of S.

$Y=0.00141+0.58348*x$
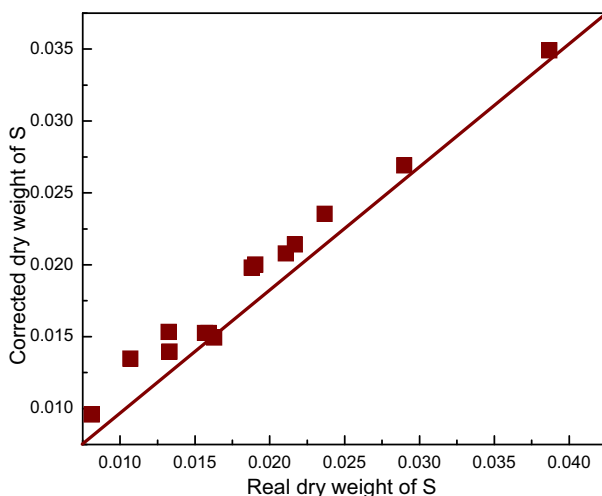
Fig. 10. Comparison between corrected dry weight and real dry weight of S.

tween the corrected dry weight and the real dry weight of S, because the influences of Al and Mg on S are still present. If influences of Al and Mg are further corrected for according to the above method, then the corrected dry weight will further approach its real dry weight. It should be noted that the concentration of Al, Mg, or K influences the concentration calculation of elements Ti and Gd besides those of Si, Ca, S, and Fe, which needs to be investigated.

It need be emphasized that the correction method is dependent not only on the level of disturbing elements, but also on the level of the main tested elements. As a consequence, the method can give the true correction in a limited range of tested elements.

5.   FIELD EXAMPLE

Chlorine in the borehole or formation fluid degrades the usable formation signals in the measured spectra and increases the statistical uncertainty for its major neutron absorber. Barite in the internal fluid in a well, a gamma-ray absorber, will deform the measured spectrum. In order to reduce the influences of barite and chlorine, we select well A with fresh mud and low formation water salinity. The mud in the borehole has no barite, and the formation containing elements Mg, Al, and K, is more complex than the simulation model used to determination of corrections. The constrained and weighted least-squares method is used to analyze capture gamma-ray energy spectra data of GEM tool to get relative yields of Si, Ca, Fe, and S. The correction of iron yields was done for the residual contribution from stainless steel in the tool. Then, these yields were converted into dry weights, as shown in Fig. 11. The solid lines represent elemental dry weights after correcting for influences of elements Mg, Al, and K based on elemental dry
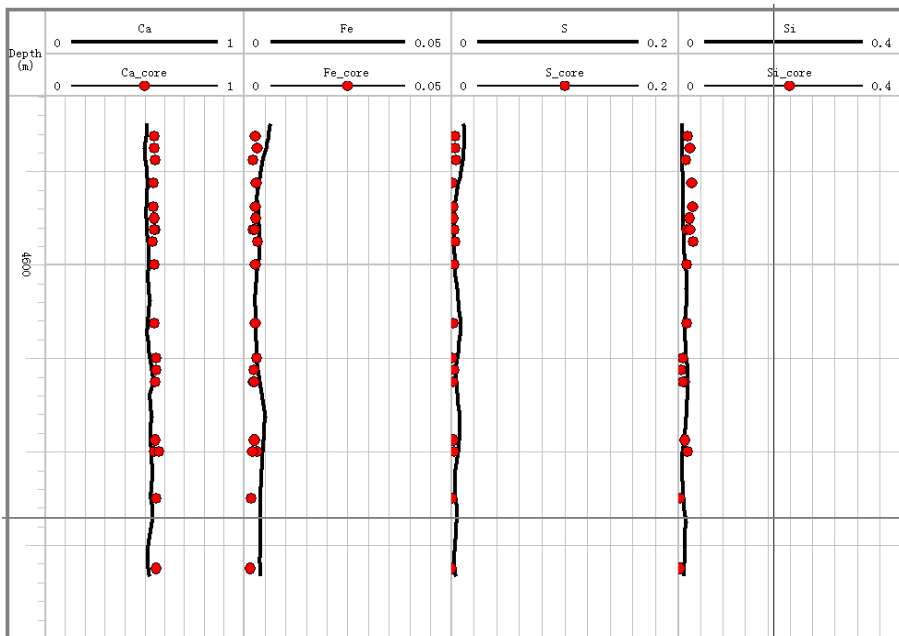
Fig. 11. Influences of Mg, Al, and K, and corrected real well data.

weights derived from spectral analysis using the constrained and weighted least-squares method; the discrete data points are real dry weights of elements derived from core experimental measurement. Before the corrections for the influences of Mg, Al, and K, we find that the calculated dry weight of Si is almost immune to, and the dry weights of Ca and Fe are less affected by Mg, Al, and K, whereas the dry weight of S is the most influenced and deviate from the real values greatly. After the corrections, the real sulfur dry weights derived from core experiments are closer to the corrected elemental dry weights curve, indicating that the above correction methods are effective.

6. CONCLUSIONS

Spectral analysis of both model formation and real formation demonstrates that, in Am-Be-source chemical elemental logging, elements Al, Mg, and K in the formation will influence the calculation of concentrations of other elements. Mg hardly influences the analysis precision of concentrations of Si, Ca, and Fe and influences S concentration to a certain extent. Al influences the calculation of concentrations of Ca, Fe, and S a little but hardly influences Si. K influences other elements not so distinctly but does influence the S concentration calculation significantly. Such influences can be eliminated by correction to obtain more accurate elemental dry weights.

References

Chapman, S., J.L. Colson, B. Everett, C. Flaum, M.M. Herron, R.C. Hertzog, G. Pirie, H. Scott, J.S. Schweitzer, J. La Vigne, J. Querein, and R. Wendlandt (1987), The emergence of geochemical well logging, *Tech. Rev*. **35**, 2, 27-35.

Ellis, D.V., and J.M. Singer (2008), *Well Logging for Earth Scientists*, 2nd ed., Springer, Dordrecht.

Galford, J.E., J.A. Quirein, S. Shannon, J.A. Truax, and J. Witkowsky (2009), Field test results of a new neutron-induced gamma-ray spectroscopy geochemical logging tool. **In:** *Proc. SPE Ann. Technical Conference and Exhibition, 4-7 October, 2009, New Orleans, USA*, Society of Petroleum Engineers, DOI: 10.2118/123992-MS.

Grau, J.A., and J.S. Schweitzer (1989), Elemental concentrations from thermal neutron capture gamma-ray spectra in geological formations, *Nucl. Geophys.* **3**, 1, 1-9.

Helmer, R.G., R.L. Heath, L.A. Schmittroth, G.A. Jayne, and L.M. Wagner (1967), Analysis of gamma-ray spectra from NaI(Tl) and Ge(Li) spectrometers. Computer programs, *Nucl. Instrum. Methods* **47**, 2, 305-319, DOI: 10.1016/0029-554X(67)90445-4.

Herron, M.M. (1986), Mineralogy from geochemical well logging, *Clays Clay Miner.* **34**, 2, 204-213, DOI: 10.1346/CCMN.1986.0340211.

Herron, M.M., S.L. Herron, J.A. Grau, N.V. Seleznev, J. Philips, A. El Sherif, S. Farag, J.P. Horkowitz, T.J. Neville, and H. Kai (2002), Real-time petrophysical analysis in siliciclastics from the integration of spectroscopy and triple-combo logging. **In:** *Proc. SPE Ann. Technical Conference and Exhibition, 29 September – 2 October 2002, San Antonio, USA*, Society of Petroleum Engineers, DOI: 10.2118/77631-MS.

Herron, S.L., and M.M. Herron (1996), Quantitative lithology: An application for open and cased hole spectroscopy. **In:** *SPWLA 37th Ann. Logging Symp., 16-19 June 1996, New Orleans, USA*, Society of Petrophysi- cists and Well-Log Analysts, SPWLA-1996-E, 14 pp.

Hertzog, R., L. Colson, O. Seeman, M. O'Brien, H. Scott, D. McKeon, P. Wraight, J. Grau, D. Ellis, J Schweitzer, and M. Herron (1989), Geochemical logging with spectrometry tools, *SPE Formation Eval.* **4**, 2, 153-162, DOI: 10.2118/16792-PA.

Pang, J.F. (1991), *γ-Ray Spectrum Data Analysis*, Shanxi Science & Technology Press, Xi'an (in Chinese).

Pemper, R.R., A. Sommer, P. Guo, D. Jacobi, J. Longo, S. Bliven, E. Rodriguez, F. Mendez, and X. Han (2006), A new pulsed neutron sonde for derivation of formation lithology and mineralogy. **In:** *Proc. SPE Ann. Technical Conference and Exhibition, 24-27 September 2006, San Antonio, USA*, Society of Petroleum Engineers, DOI: 10.2118/102770-MS.

Radtke, R.J., M. Lorente, B. Adolph, M. Berheide, S. Fricke, J. Grau, S. Herron, J. Horkowitz, B. Jorion, D. Madio, D. May, J. Miles, L. Perkins, O. Philip, B. Roscoe, D. Rose, and Ch. Stoller (2012), A new capture and inelastic spectroscopy tool takes geochemical logging to the next level. **In:** *SPWLA 53rd Ann. Logging Symp., 16-20 June 2012, Cartagena, Colombia,* Society of Petrophysicists and Well-Log Analysts, SPWLA-2012-103, 16 pp.

Sweezy J.E., T.E. Booth, and F.B. Brown (2004), MCNP-A general Monte Carlo N-particle transport code. Version 5, MCNP user manual, LA-UR-03-1987, Los Alamos National Laboratory, Los Alamos, USA.

Method for Identifying Micro-seismic *P*-Arrival by Time-frequency Analysis Using Intrinsic Time-Scale Decomposition

Ruihong ZHANG and Lihua ZHANG

Department of Civil Engineering, North China Institute of Science and Technology, Sanhe-Hebei, China; e-mails: zhangruihongw@163.com, zhanglihua@ncist.edu.cn

A b s t r a c t

A method to identify the *P*-arrival of microseismic signals is proposed in this work, based on the algorithm of intrinsic timescale decomposition (ITD). Using the results of ITD decomposition of observed data, information of instantaneous amplitude and frequency can be determined. The improved ratio function of short-time average over long-time average and the information of instantaneous frequency are applied to the time-frequency-energy denoised signal for picking the *P*-arrival of the microseismic signal. We compared the proposed method with the wavelet transform method based on the denoised signal resulting from the best basis wavelet packet transform and the single-scale reconstruction of the wavelet transform. The comparison results showed that the new method is more effective and reliable for identifying *P*-arrivals of microseismic signals.

Key words: micro-earthquake, *P*-arrival, ITD, time-frequency analysis, denoising.

1. INTRODUCTION

Microseismic monitoring technology provides a good method for monitoring rock mass rupture during exploration and excavation. For early warning of rupture and to prevent damage caused by rock mass fracture, a rapid and re-

liable monitoring method is needed to detect and accurately identify the arrival of the first microseismic signal.

As microseismic signals from mining activities usually propagate over short distances, the STA/LTA (the ratio of the Short Time Average over the Long Time Average) method of amplitude change is often used, with the first arrival of the P-wave taken as the trigger point. Although improved detection functions were developed (Allen 1978, 1982, McEvilly and Majer 1982, Baer and Kradolfer 1987, Earle and Shearer 1994), detecting the first arrival can result in a rough estimate because the accuracy of the STA/LTA-based method depends on the length of the time window and the predefined threshold (Ma 2008). Ye *et al.* (2008) used the maximum energy ratio of two time windows to detect the first arrival of microseismic waves. With this method, the results are still affected by the length of the time window and the shape of the waveform. In other studies that identified the first arrival of seismic waveforms (Zhang *et al.* 2003, Bai and Kennett 2000, Liu *et al.* 2013), large deviations exist between the real arrivals and the P-wave picking results (Dai and MacBeth 1995).

To overcome the difficulties mentioned above, in this paper, we introduce the method of intrinsic timescale decomposition (ITD) (Frei and Osorio 2007) and propose a highly accurate and efficient algorithm for identifying the first P-arrival of microseismic signals. Based on the time-frequency-energy denoised instantaneous amplitude and frequency information that was obtained, an algorithm that accurately identifies the P-arrival is presented by using four parameters. Based on the instantaneous piece-wise constant amplitude and instantaneous frequency information, the thresholds for the identification scheme are easier to predefine and the adaptive calculation time window for the STA/LTA calculation can be derived. We apply this method to simulated and real data and show that the proposed method can accurately identify the first P-arrival of seismic signals, providing a feasible technological approach for detecting P-wave arrivals.

2. INSTANTANEOUS TIME-FREQUENCY-ENERGY INFORMATION EXTRACTION AND DENOISING

2.1 Inherent timescale decomposition algorithm

Given the signal X_t, the operator L is defined whereby the baseline signal L_t can be extracted from the signal X_t so that the residual signal H_t is a proper rotation. Hence, X_t can be decomposed as L_t and H_t. Let $\{\tau_k, k = 1, 2, \ldots\}$ be the local extrema of X_t, X_k, and L_k denoted $X(\tau_k)$ and $L(\tau_k)$, respectively, then the baseline signal L_t on the interval $(\tau_k, \tau_{k+1}]$ between successive extrema is as follows (Frei and Osorio 2007):

$$L_t = LX_t = L_k + \left(\frac{L_{k+1} - L_k}{X_{k+1} - X_k} \right)(X_t - X_k) \ , \quad t \in (\tau_k, \tau_{k+1}] \ , \tag{1}$$

were

$$L_{k+1} = \alpha \left[X_k + \left(\frac{\tau_{k+1} - \tau_k}{\tau_{k+2} - \tau_k} \right)(X_{k+2} - X_k) \right] + (1 - \alpha)X_{k+1} \tag{2}$$

and α is set as 0.5. The residual signal H_t is as follows:

$$H_t = HX_t = X_t - L_t \ . \tag{3}$$

To make the residual function monotonic between extrema, as required for proper rotations, the baseline is constructed as a linearly transformed contraction of the original signal. Once the original signal is decomposed, the decomposition can be applied to the baseline signal again. The decomposition procedure is iterated several times. Finally, the original signal can be decomposed into a series of proper rotation components of successively decreasing instantaneous frequencies at each subsequent level and a low-frequency baseline signal. The decomposition iterative process is as follows (Frei and Osorio 2007):

$$X_t = HX_t + LX_t = HX_t + (H + L)LX_t = \left(H \sum_{k=0}^{p-1} L^k + L^p \right) X_t \ , \tag{4}$$

where $HL^k X_t$ is the proper rotation of the $(k+1)$th level, and $L^p X_t$ is the extracted lowest frequency baseline.

Through the operator L, the inherent information of the original signal X_t (for example the extremum position) can be passed down to the baseline and the residual parts. The residual component H_t and baseline L_t can be monotonic on the interval $(\tau_k, \tau_{k+1}]$, the extrema $\{\tau_k, k \geq 1\}$ of H_t coincide with the extrema of X_t, and the extrema $\{\tau_k, k \geq 1\}$ of X_t are the extrema or inflection points for L_t. The initial decomposition of the signal on the interval $[0, \tau_1]$ can be extended to values of $t < 0$ by reflections. Incomplete waves at the beginning and/or end of H_t, i.e., data prior to the first zero up-crossing or after the last zero up-crossing may be treated according to these definitions over the appropriate subinterval of $[t_1, t_5]$. H_t is a proper rotation on the interval $[\tau_1, \tau_N]$ for any $N \geq 1$. The proof is as follows (ITD; Frei and Osorio 2007).

Assuming that X_t has a local maximum at τ_k and adjacent local minima at τ_{k-1} and τ_{k+1}, then H_t has a local maximum at τ_k and is monotonic on $[\tau_{k+1}, \tau_k]$ and on $[\tau_k, \tau_{k+1}]$ because:

$$H_t = X_t - L_t = \begin{cases} X_t - (\alpha_K X_t + \beta_K) & \text{for} \quad t \in [\tau_{k-1}, \tau_k] \\ X_t - (\alpha_{K+1} X_t + \beta_{K+1}) & \text{for} \quad t \in [\tau_k, \tau_{k+1}] \end{cases} , \tag{5}$$

where $\alpha_K = (L_k - L_{k-1})/(X_k - X_{k-1})$, $\forall$k.

If X_t is monotonic on $[\tau_{k-1}, \tau_k]$ and on $[\tau_k, \tau_{k+1}]$, then H_t is monotonic on $[\tau_{k-1}, \tau_k]$ and on $[\tau_k, \tau_{k+1}]$.

If $L_{k-1} > X_{k-1}$, $L_{k+1} > X_{k+1}$, and $L_k < X_k$ then $H_{k-1} = X_{k-1} - L_{k-1} < 0$, $H_k = X_k - L_k > 0$, and $H_{k+1} = X_{k+1} - L_{k+1} < 0$. Therefore, H_t has a local maximum at $t = \tau_k$.

2.2 The extraction of instantaneous information

When a microseismic signal is decomposed into a series of intrinsic rotational components and a low-frequency baseline signal using ITD, the inherent instantaneous time-frequency-energy information of the input signal can be obtained by analyzing the morphology and other features of the waves. Specifically, for the intrinsic rotational components, based on the single waves and their monotonic segment between adjacent extrema, we get the instantaneous amplitude information, which is piece-wise constant, and the constant value is determined by the extremum values of the proper rotations between zero crossings from the signal X_t. For one full wave (*i.e.*, the portion of the signal between successive zero up-crossings) at a time, the instantaneous amplitude information is as follows (Frei and Osorio 2007):

$$A = \begin{cases} A_1, & t \in [t_1, t_3) \\ -A_2, & t \in [t_3, t_5) \end{cases}, \tag{6}$$

where $A_1 > 0$ and $A_2 > 0$ are the absolute amplitude of the positive and negative half-waves between successive zero up-crossings t_1 and t_5, respectively, and t_2 is the time of the positive half-wave maxima (A_1), t_3 is the time of the zero down-crossing, and t_4 is the time of the negative half-wave minima ($-A_2$).

In accordance with the instantaneous amplitude, the instantaneous frequency is also piece-wise constant and can be computed between times of successive extrema of the ITD-generated proper rotation components, as follows (Frei and Osorio 2007):

$$f_t = \begin{cases} \dfrac{0.5}{t_3 - t_1}, & t \in [t_1, t_3) \\ \dfrac{0.5}{t_5 - t_3}, & t \in [t_3, t_5) \end{cases}. \tag{7}$$

For simplicity, an alternative way to write the instantaneous period is

$$T_t = \begin{cases} 2(t_3 - t_1), & t \in [t_1, t_3) \\ 2(t_5 - t_3), & t \in [t_3, t_5) \end{cases}. \tag{8}$$

At the beginning of a proper rotation component, the incomplete waves, *i.e.*, the data prior to the first zero up-crossing, or the incomplete waves at the end of a proper rotation component, *i.e.*, the data after the last zero up-crossing, may be processed according to these definitions over the appropriate subinterval of $[t_1, t_5]$.

2.3 ITD-based time-frequency-energy denoising

The accuracy of the detection of the first arrival is dependent on the SNR of the recorded signals; therefore denoising is an important stage of the arrival detection process. It is necessary to reduce the interference noise of the microseismic signals as much as possible. As any slight noise, distortion, or phase shift on the first arrival can result in an incorrect time picking (Galiana-Merino *et al.* 2003), a denoising method based on ITD (Zhang *et al.* 2012) is more suitable for detecting the seismic signal because the abrupt change of the amplitude at the beginning of the first pulse is preserved. We therefore apply the ITD method to the microseismic signals.

To carry out ITD-based time-frequency-energy denoising, two steps are needed: (i) the instantaneous frequencies and instantaneous amplitudes are extracted for each of the ITD-generated proper rotation components of the original signal; and (ii) the denoised components are obtained by using the amplitude thresholds and period (frequency) thresholds.

The decomposition levels are dependent on the ratio of the average amplitude energy of the denoised components to the amplitude energy of the total denoised signal. Generally, we used the first several components containing the main frequency bands of a microseismic event. Ultimately, the time-frequency-energy denoised instantaneous amplitude signal and frequency characteristics were obtained for the first pulse detection.

Figure 1 illustrates the process of ITD-based time-frequency-energy denoising. First, a synthetic signal was generated at frequencies of 6, 7, and 10 Hz (Fig. 1a). The synthetic signal begins at 3 s, using a sample frequency of 200 Hz. The pure synthetic signal was contaminated with Gaussian noise so that the SNR is 10 dB (Fig. 1b). Figure 1c shows the original decomposed piece-wise constant instantaneous amplitude of the first four components, and Fig. 1d shows the original piece-wise constant instantaneous period of the first four components. Figure 1e shows the denoised instantaneous amplitude of the first four components, and Fig. 1f shows the filtered instantaneous period of the first four components. Using the ITD method, the level-dependent amplitude thresholds and frequency thresholds were applied according to the noise samples, which were several seconds ahead of the signal (Zhang *et al.* 2012); meanwhile the denoised instantaneous amplitudes were filtered by the high-frequency thresholds and the instantaneous period in-
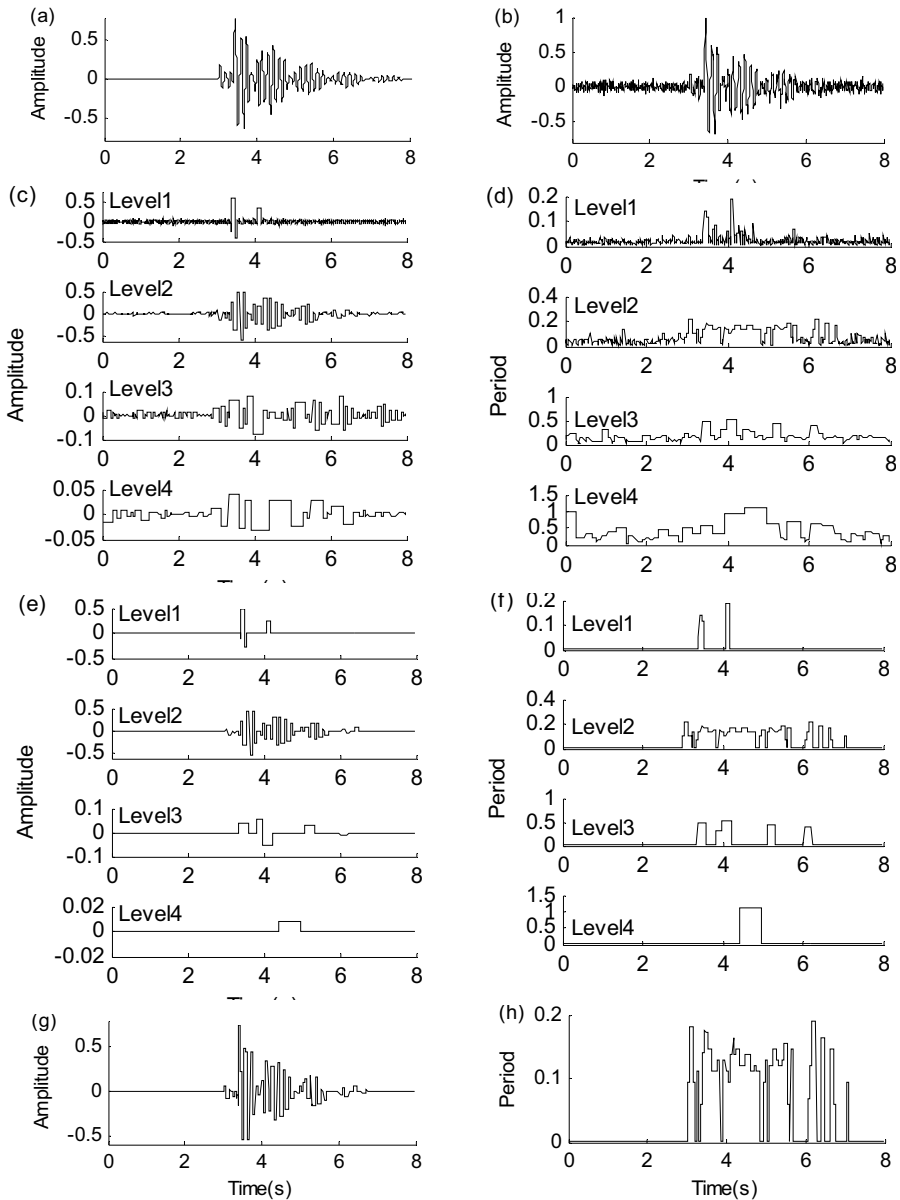
Fig. 1. Denoising process for a contaminated synthetic signal using the ITD method: (a) pure synthetic signal generated by several sine waves damped by an exponential, (b) synthetic signal contaminated by Gaussian noise, (c) the original piece-wise constant amplitude of the first four components, (d) the original piece-wise constant period of the first four components, (e) time-frequency-energy denoised instantaneous amplitude of the first four components, (f) filtered instantaneous period of the first four components, (g) denoised amplitude from the first three amplitude components, and (h) equivalent period calculated from the first two components.

formation was also dependent on the denoised amplitude information. The extrema of the denoised discomposed components coincide with those of the original signal, and the instantaneous period components are in good agreement with the instantaneous amplitude components, demonstrating the advantage of the proposed method for first arrival detection.

Table 1 presents the average amplitude and frequency of the denoised first four components. The two maximum energy components of levels 2 and 1 account for 80% of the total energy. The energy from level 4 and subsequent levels comprises nearly 9% of the total energy, the frequency of level 3 is lower than the generated frequency, and further decomposition is not needed. The sum of the first three amplitude components that contain the components of the three maximum energy levels as the instantaneous amplitude information is reasonable (Fig. 1g). We use the equivalent period as the instantaneous frequency information (Fig. 1h). The equivalent period is the energy-weighted sum of the frequency components by the average amplitude ratio of the two maximum energy components. The equivalent average frequency of levels 2 and 1 is 6.89 according to this conversion which is a good estimate of the main frequency of 7.

Table 1

Average amplitude and frequency of the denoised components

Component level	Average frequency [Hz]	Average amplitude	Amplitude ratio [%]
1	6.92	0.0098	11.2
2	6.88	0.0602	69.0
3	3.76	0.0095	10.9
4	1.39	0.0033	3.8

3. THE PROPOSED ALGORITHM

An algorithm for the first pulse identification was proposed based on the denoised piece-wise constant amplitude and the equivalent period (frequency) information. First, the thresholds of the instantaneous amplitude ratio (STA/BTA) and the instantaneous period (FSTA) were predefined. STA is the maximum amplitude of the half-wave that the detected point belongs to, BTA is the maximum amplitude of the wave that the previous point belongs to, and FSTA is the instantaneous period average of the half-wave that the detected point belongs to. When STA/BTA and FSTA exceed their predetermined thresholds, the average amplitude ratio (LTA/BTA) and the average period (FLTA) in the next period of time of the signal will be calculated based on the information of that period. LTA and FLTA are the long-term-

average amplitude and the average period of the next consecutive period time related to the detected point, respectively. When LTA/BTA and FLTA are larger than their predetermined thresholds, this indicates that a micro-seismic event is occurring. The first two parameters ensure that the instantaneous amplitude and period are consistent with a microseismic event. The last two parameters can prevent a false detection caused by individual disturbance fluctuations. The detailed steps are as follows:

❑ The original signals are decomposed by ITD and the components of the first three or four levels are generated; the four components generally contain the three maximum energy levels and include the main frequency bands of the original signals.

❑ Instantaneous amplitude and period (frequency) information is extracted from the decomposition components. The instantaneous amplitude components are denoised and the period components are filtered by level-dependent amplitude thresholds and level-dependent period thresholds, respectively. The denoised instantaneous amplitude is filtered by the high-frequency threshold and the instantaneous period is also dependent on the denoised amplitude information; thus, the amplitude and period information are interdependent.

❑ The first several levels (generally three levels) of the denoised amplitude components are summed as the instantaneous amplitude information; these components contain the two maximum energy levels. The energy-weighted period components of the two maximum energy levels are summed as the equivalent period information; the weighted coefficients may be determined by the average amplitude ratio of the two maximum energy levels (such as levels 1 and 2 in Table 1).

❑ STA/BTA and FSTA are detected for each point. When FSTA and STA/BTA are higher than the two predefined thresholds, the next step is followed.

❑ LTA and FLTA of the next consecutive period time connected with the detected point are calculated. The optimum time window for calculating LTA and FLTA depends on the period information. If LTA/BTA and FLTA are higher than the two thresholds then a phase arrival is affirmed and the first-arrival time is determined.

The first-arrival is defined as the first sample where all four parameters (STA/BTA, FSTA, LTA/BTA, and FLTA) exceed their thresholds. As the instantaneous piece-wise constant amplitude is determined by the extremum values between the zero crossings rather than by the transient amplitude, it is easy to predefine the amplitude threshold for the first-arrival detection because the abrupt extremum value is far larger than the low transient amplitude value at the beginning of a pulse. For a routine transient amplitude signal, because of the low transient amplitude at the beginning of the first

pulse, the threshold is difficult to predetermine and often gives false picking results. Based on the period (frequency) information of the signal, the location of the next wave information can be determined; therefore, the adaptive time window can be obtained for the proposed algorithm.

It is worth mentioning that as the proposed scheme is applied to the previously denoised signal, there may be a number of samples preceding the first arrival that were reduced to zero or to a very low value, which could produce a low BTA value. Therefore, stability factors should be introduced when calculating the ratio-related values.

4. MODELING RESULTS BY MATLAB

To evaluate the proposed method, first-arrival identification was performed for the contaminated synthetic signal of Fig. 1b, first by the proposed method and then using the wavelet transform (WT) method, and the results were compared.

The first-arrival detection of the proposed method is illustrated in Fig. 2a. The equivalent periods were calculated from the instantaneous peri-
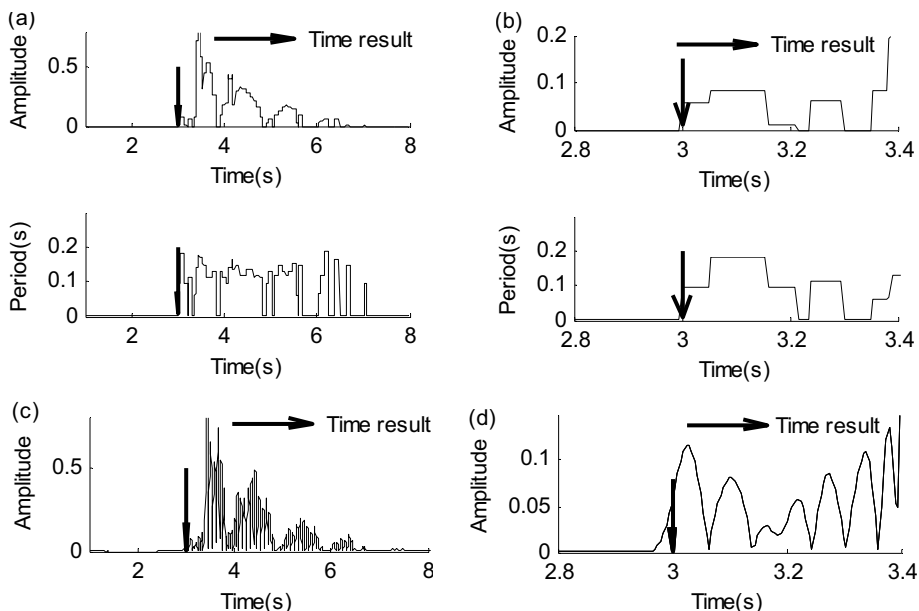
Fig. 2. First-arrival identification by two methods for the contaminated synthetic signal shown in Fig. 1b: (a) the absolute denoised amplitude and equivalent period used by the proposed method and the picking result, (b) the local magnified drawing near the picking results of Fig. 2a, (c) the normalized absolute reconstructed signal used by the WT method and the picking result, and (d) the local magnified drawing near the picking results of Fig. 2c.

od components of levels 1 and 2; the energy weighted coefficients are the average amplitude ratios of the two maximum energy components. The instantaneous amplitude signals used the sum of the first two denoised amplitude components including the two maximum energy levels. The thresholds of FSTA and FLTA were estimated based on the equivalent period threshold for period filtering and the average equivalent period, respectively. The thresholds used for STA/BTA, LTA/BTA, FSTA, and FLTA were 1.5, 1.5, 0.03, and 0.06, respectively. The automatic first-arrival picking result was 3.000 s, indicated by the vertical dashed line.

The WT method of first-arrival detection uses the normalized single-scale reconstructed signals by wavelet transformation of the denoised signals. The denoised signals were obtained through the optimum wavelet packet transform methods by using a node-dependent threshold of the noise (Zhang *et al.* 2011). The first-arrival times were detected by the parameter thresholds of STA/LTA and LTA. When the ratio of STA/LTA and LTA exceeded their predetermined thresholds, the first *P*-arrival time was identified (Galiana-Merino *et al.* 2007)

The Daubechies 8 orthogonal mother wavelet (Daubechies 1992) was employed, four levels of decomposition were used for the WT algorithm, and a normalized single-scale reconstructed signal of the 4th scale was obtained for the first-arrival detection. The normalized absolute single-scale reconstructed signal used by the WT method is shown in Fig. 2c. The calculation time windows used for the WT algorithm were: an STA window of 20 samples (the sampling rate of the signal was 5 ms, equivalent to one period of a 10-Hz pulse), an LTA window of 80 samples (four times the STA window length). The threshold of STA/LTA was 2.6, and that of LTA was 0.002. The automatic picking result indicated by the long vertical dashed line is 3.000 s.

Figure 2b and d shows a magnified view of Fig. 2a and c near the first-arrival picking results of the two methods, respectively. The picking point of each method is indicated by a vertical dashed line.

For the proposed method, the instantaneous amplitude and equivalent period of the original signal are obtained, and the denoised amplitude preserves the sudden change of amplitude at the beginning of the first pulse, as well as prevents the distortion and phase shift phenomenon (Fig. 2b). Moreover, the instantaneous equivalent period of the original signal that is consistent with its amplitude information provides an important evaluation criterion for the first pulse identification. This suggests that the proposed method provides better identification of the first pulse. The WT denoised waveform (Fig. 2d) is smooth; however, because of the inherent characteristic of the wavelet, distortion, and a phase-shift phenomenon occur in the denoised reconstruct-

ed signal at the beginning of the first arrival; this can easily lead to an identi-
fication of the first arrival earlier than it actually was.

5.   APPLICATION EXAMPLES

Two applications of the two methods are illustrated as follows.

Figure 3 shows the first-arrival picking results for manmade blasting
signals of high SNR using the two methods mentioned above. The six
original signals of the manmade blasting are shown in Fig. 3a. The equiva-
lent periods used by the proposed method are presented in Fig. 3b. In
Fig. 3c, the denoised amplitude information used by the proposed method
and the denoised normalized single scale reconstructed signals used by the
WT method are compared with the original signals.

For the first arrival detection by the proposed method, the first maximum
energy level is level 2 and the second maximum energy level is level 3 for
all signals. Therefore, the equivalent periods were all calculated from the in-
stantaneous period components of levels 2 and 3; the instantaneous ampli-
tude signals used the sum of the first three amplitude components including
the two maximum energy levels. The first component of level 1 can be ig-
nored because of its very low energy and high frequency. The identification
thresholds used here for STA/BTA, LTA/STA, FSTA, and FLTA are 1.5,
1.5, 0.002, and 0.004, respectively.

For the first arrival detection by the WT method, the threshold of
STA/LTA was 2.6, and that of LTA was 0.001. The main frequency band
derived from the FFT of the original signals was 10-60 Hz. The time win-
dows for the calculation were: an STA window of 50 samples (the sampling
rate of the signal is 0.2 ms, equivalent to half a period of a 50-Hz pulse), and
an LTA window of 200 samples (four times the STA window length).

Table 2 shows the automatic picking results of the proposed method and
the WT method and their errors compared with the manual picking results.
As the original manmade blasting signals have a high SNR, the manual re-
sults have small errors compared with the real results. The difference be-
tween the automatic picking results and the manual results may be taken as
the errors of the picking results.

The first-arrival picking results for low SNR microseismic signals by the
two automatic methods are shown in Fig. 4. Figure 4a shows the waveforms
of six microseismic signals recorded in a mine. The equivalent periods used
by the proposed method are presented in Fig. 4b. The amplitude information
used by the proposed method and the normalized reconstructed signals used
by the WT method compared with the original signals near the picking re-
sults are shown in Fig. 4c.

In the proposed method, the two maximum energy levels are levels 2
and 3 for signals S1-S6 except of S4; the maximum energy level for S4 is
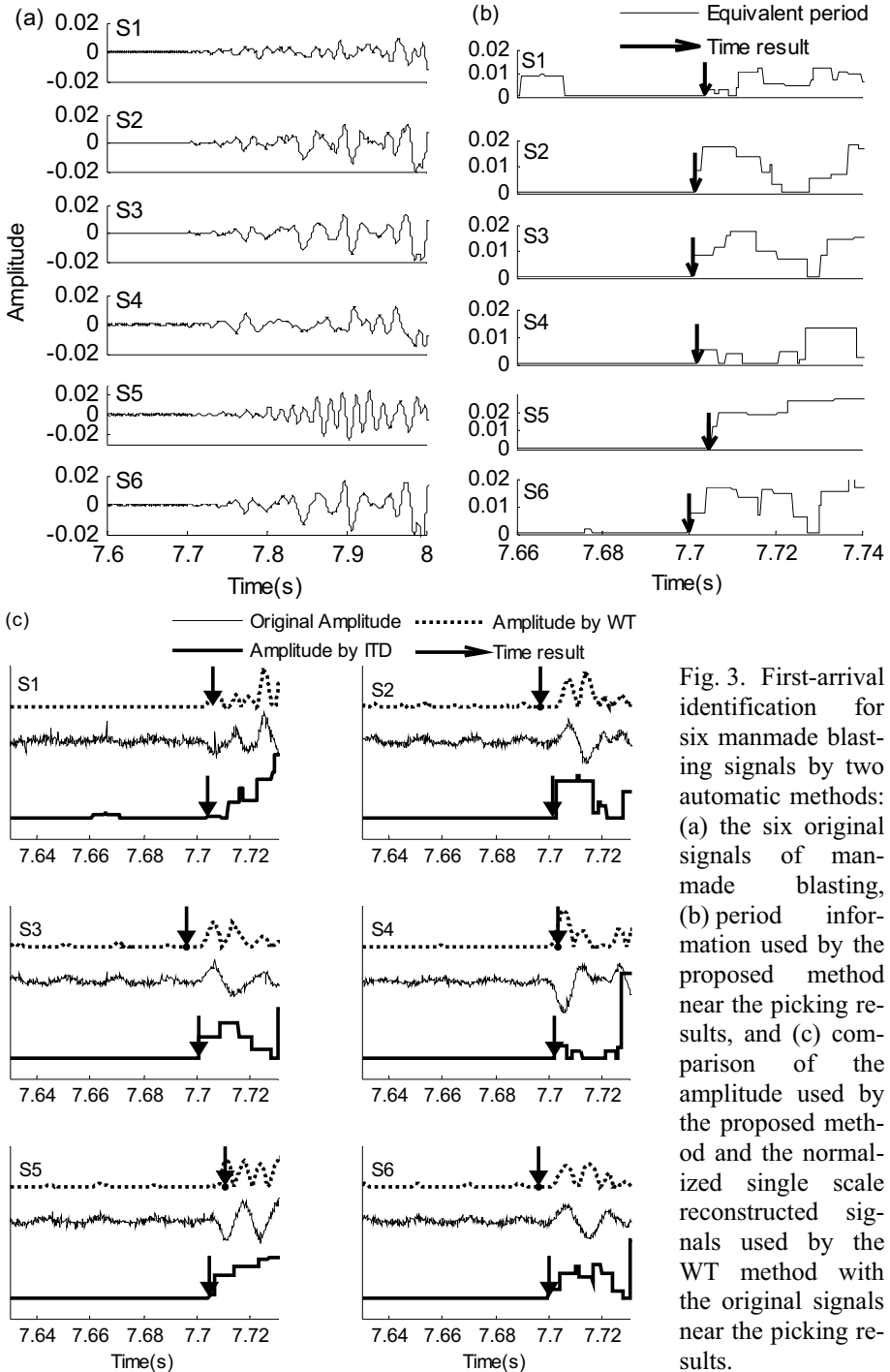
Fig. 3. First-arrival identification for six manmade blasting signals by two automatic methods: (a) the six original signals of manmade blasting, (b) period information used by the proposed method near the picking results, and (c) comparison of the amplitude used by the proposed method and the normalized single scale reconstructed signals used by the WT method with the original signals near the picking results.

Picking results comparison of manmade blasting waves

Signal number	Manual results	Proposed method		WT method	
		Results	Error*	Results	Error*
1	7.7030	7.7038	0.0008	7.7062	0.0032
2	7.7030	7.7014	−0.0016	7.6970	−0.0060
3	7.7020	7.7008	−0.0012	7.6962	−0.0058
4	7.7010	7.7020	0.0010	7.7032	0.0022
5	7.7070	7.7048	−0.0022	7.7106	0.0036
6	7.7010	7.6998	−0.0012	7.6960	−0.0050

*)The error means the difference between the manual picking result and the auto-mated methods result.

level 1. For consistency, the amplitude signals and the equivalent periods were calculated by using the first two maximum energy levels for all the signals. The thresholds for STA/BTA, LTA/STA, FSTA, and FLTA were 1.5, 1.5, 0.001, and 0.002.

For the WT method, the threshold of STA/LTA was 2.6 and the threshold of LTA was 0.003. The main frequency band of the original signals was 50-150 Hz. The time windows for the calculation were: an STA window of 50 samples (signal sampling rate of 0.2 ms, equivalent to a period of a 100-Hz pulse), and an LTA window of 200 samples (four times the STA window length).

Table 3 shows the automatic picking results of the proposed method and the WT method and their errors compared with the manual picking results. As the microseismic signals have low SNR, the manual results are less reliable than those of the manmade blasting signals. The errors of the automatic picking results were also estimated from the difference between the automatic picking results and the manual results.

Comparing the results from Tables 2 and 3, the errors of the WT method were bigger than the errors of the proposed method in most cases, and were less than those of the proposed method in a few cases. For high SNR signals, the differences between the proposed method and the manual results are less than 0.00 ± 0.002 s in most cases; for low SNR signals, the errors of the proposed method are within 0.00 ± 0.003 s in most cases; if the discomposed components of level 4 were included in the amplitude and the period information for low SNR microseismic signals, the errors between the proposed method and the manual results would be less, especially for signal S6 in Fig. 4.

In fact, the WT denoised signals were roughly the same as the ITD denoised signals in most cases. For the high-SNR signals, the WT denoised
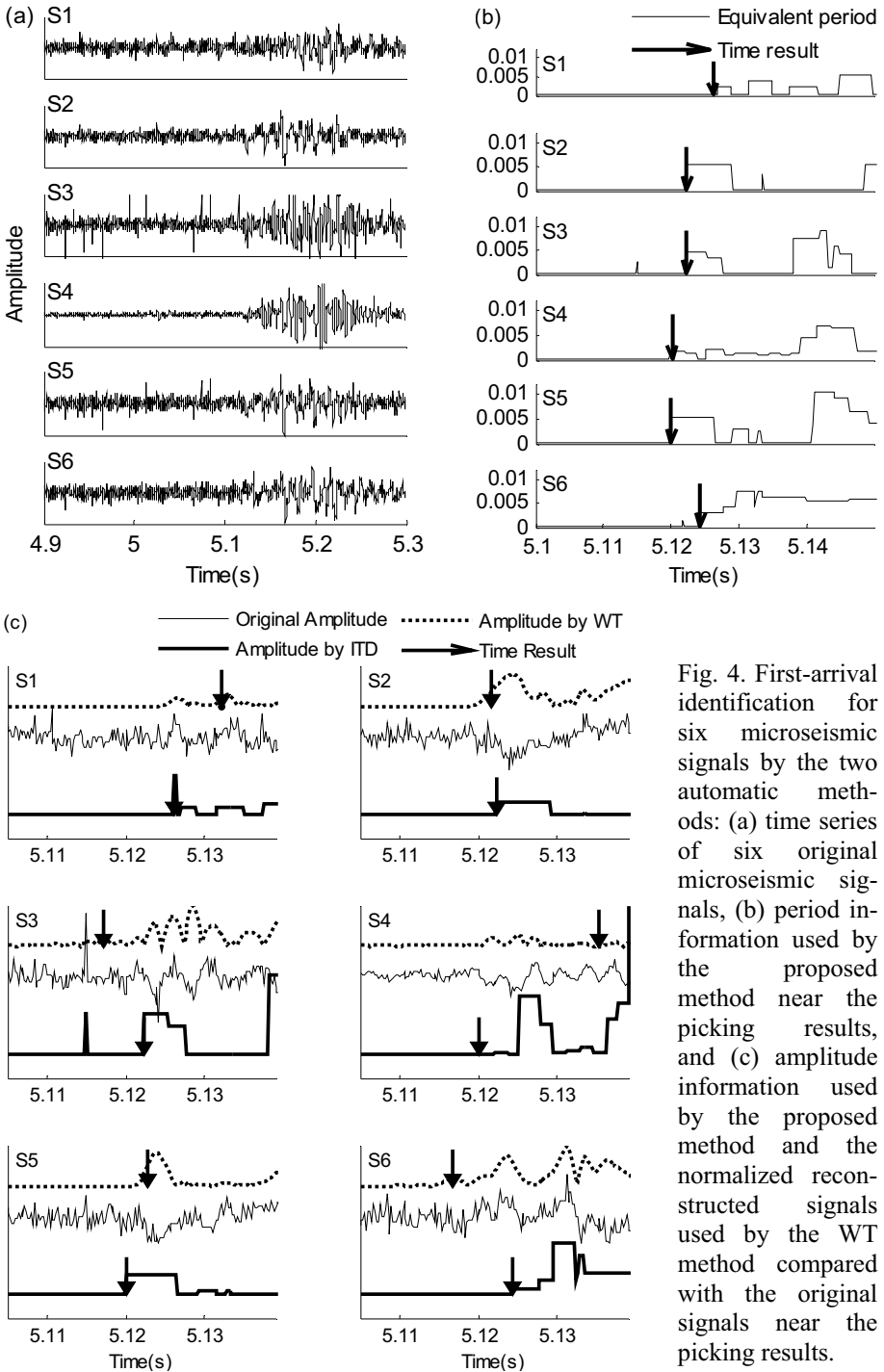
Fig. 4. First-arrival identification for six microseismic signals by the two automatic methods: (a) time series of six original microseismic signals, (b) period information used by the proposed method near the picking results, and (c) amplitude information used by the proposed method and the normalized reconstructed signals used by the WT method compared with the original signals near the picking results.

Table 3

Picking results of the microseismic signals

Signal number	Manual results	Proposed method		WT method	
		Results	Error*	Results	Error*
1	5.1250	5.1262	0.0008	5.1322	0.0072
2	5.1220	5.1224	0.0004	5.1216	−0.0004
3	5.1210	5.1224	0.0014	5.1172	−0.0038
4	5.1200	5.1202	0.0002	5.1352	0.0152
5	5.1230	5.1200	−0.0030	5.1228	−0.0002
6	5.1190	5.1244	0.0054	5.1168	−0.0022

*)The error means the difference between the manual picking results and those of the automated methods.

signals showed excellent accuracy and were smoother, especially for the first wave arrival. The main reason for the bigger picking errors of this method is the criterion for the first-pulse identification. For signals in the wide main frequency bands, only amplitude information can be used for the detection, and parameters such as the window length and thresholds which affect the accuracy of the picking results are difficult to determine, especially for low SNR signals. For low SNR signals, there are obvious noise wave in the denoised reconstructed signals, which may lead to incorrect picking results. The picking results of the WT method depend mostly on the LTA threshold; thus, when the time window or the thresholds change, the results will change accordingly.

For the ITD method, the denoised signals preserved the sudden change of the amplitude at the first-arrival, as well as extracting the period information consistent with the instantaneous amplitude. These two features make it easy to establish the evaluation criterion for the first-arrival identification. For every point, using the instantaneous piece-wise constant amplitude equal to the extremum value of the instantaneous half-wave that the point belongs to instead of the point's transient amplitude, makes the amplitude threshold easier to preset. The instantaneous equivalent period of the original signals was used not only for the evaluation criterion, but also for determining the location of the consecutive wave; therefore, the time windows can be adaptively determined. The thresholds of the proposed method were estimated based on the reference values from the period information and the picking time results were stable. These are the main advantages of the proposed ITD method, suggesting that in most cases the results of the proposed method are more reliable and stable than those of the WT method.

6.   CONCLUSIONS

Using the time-frequency-energy denoised amplitude signals and the filtered equivalent period (frequency), we present a method for detecting first-arrival signals and apply it to simulated signals and manmade blasting signals of high SNR and microseismic signals of low SNR for automatic first-arrival detecting.

The instantaneous amplitude and equivalent period information for the first-arrival identification were obtained by the proposed method using several decomposed component levels. In general, the first decomposed component of level 1 has some effect on the long-term-average amplitude and the average period for the first pulse identification in some cases; the second and third components are effective signals – the two maximum energy levels that contain the dominant frequency bands. The amplitude information uses the sum of the first several instantaneous amplitude components which should include the two maximum energy levels. The equivalent period information uses the energy weighted sum of the period components of the two maximum energy levels weighted by their average amplitude ratio.

The first pulse identification depends on four amplitude thresholds and period thresholds, including the instantaneous amplitude ratio and instantaneous period, and the average amplitude ratio and average period of the next waves connected with the detected point. The denoised piece-wise constant amplitude is used for each point to detect the first arrivals; the abrupt amplitude at the beginning of the first pulse makes the amplitude threshold easy to preset. The period thresholds were estimated based on the extracted period information. According to the equivalent period information, the adaptive time window was obtained for long-time average amplitude calculation.

When the sampling frequency is low, the abrupt amplitude information at the beginning of the first arrival pulse is retained in the proposed method; for the WT method, because of the inherent characteristics of the wavelet, there exists a distortion and phase shift phenomenon at the beginning of the first arrival in the denoised signals. Thus, the picking accuracy of the proposed method is less influenced by the sampling frequency of the original signals compared with that of the WT method.

Compared with the WT method, in most cases, the picking results of the proposed method were stable, more effective, and more reliable. The deviations between the results of the proposed method and the manual results are within 0.00 ± 0.003 s in most cases, less than those of the WT method. The proposed method showed better performance in identifying P-arrivals of microseismic signals.

References

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

Allen, R. (1982), Automatic phase pickers: Their present use and future prospects, *Bull. Seismol. Soc. Am.* **72**, 6B, 225-242.

Baer, M., and U. Kradolfer (1987), An automatic phase picker for local and teleseismic events, *Bull. Seismol. Soc. Am.* **77**, 4, 1437-1445.

Bai, C., and B.L.N. Kennett (2000), Automatic phase-detection and identification by full use of a single three-component broadband seismogram, *Bull. Seismol. Soc. Am.* **90**, 1, 187-198, DOI: 10.1785/0119990070.

Dai, H., and C. MacBeth (1995), Automatic picking of seismic arrivals in local earthquake data using artificial neural network, *Geophys. J. Int.* **120**, 3, 758-774, DOI: 10.1111/j.1365-246X.1995.tb01851.x.

Daubechies, I. (1992), *Ten Lectures on Wavelets*, SIAM, Philadelphia, 357 pp.

Earle, P.S., and P.M. Shearer (1994), Characterization of global seismograms using an automatic-picking algorithm, *Bull. Seismol. Soc. Am.* **84**, 2, 366-376.

Frei, M.G., and I. Osorio (2007), Intrinsic time-scale decomposition: time-frequency-energy analysis and real-time filtering of non-stationary signals, *Proc. R. Soc. A* **463**, 321-342, DOI: 10.1098/rspa.2006.1761.

Galiana-Merino, J.J., J. Rosa-Herranz, J. Giner, S. Molina, and F. Botella (2003), De-noising of short-period seismograms by wavelet packet transform, *Bull. Seismol. Soc. Am.* **93**, 6, 2554-2562, DOI: 10.1785/0120010133.

Galiana-Merino, J.J., J. Rosa-Herranz, P. Jáuregui, S. Molina, and J. Giner (2007), Wavelet transform methods for azimuth estimation in local three-component seismograms, *Bull. Seismol. Soc. Am.* **97**, 3, 793-803, DOI: 10.1785/0120050225.

Liu, J.S., Y. Wang, and Z.X. Yao (2013), On micro-seismic first arrival identification: A case study, *Chin. J. Geophys.* **56**, 5, 1660-1666, DOI: 10.6038/cjg20130523.

Ma, Q. (2008), Study and application on earthquake early warning, Ph.D. Thesis, Institute of Engineering Mechanics, China Earthquake Administration, Harbin, China (in Chinese).

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

Ye, G.X., F.X. Jiang, and S.H. Yang (2008), Possibility of automatically picking first arrival of microseismic wave by energy eigenvalue method, *Chin. J. Geophys*. **51**, 5, 1574-1581.

Zhang, H., C. Thurber, and C. Rowe (2003), Automatic P-wave arrival detection and picking with multiscale wavelet analysis for single-component recordings, *Bull. Seismol. Soc. Am.* **93**, 5, 1904-1912, DOI: 10.1785/0120020241.

Zhang, R.H., D.C. Lin, and L. Qiao (2011), Application of the optimum wavelet packet transform in de-noising seism signals, *J. Seismol. Res.* **34**, 3, 358-364.

Zhang, R.H., H.W. Qi, D.C. Lin, and L. Qiao (2012), Micro-seismic signal de-noising based on intrinsic time-scale decomposition algorithm, *Safety Coal Min.* **43**, 3, 164-168.

Hydrodynamic Equilibrium for Sediment Transport and Bed Response to Wave Motion

Leszek M. KACZMAREK[1], Szymon SAWCZYŃSKI[2], and Jarosław BIEGOWSKI[3]

[1]University of Technology, Department of Geotechnics, Koszalin, Poland
e-mail: leszek.kaczmarek@tu.koszalin.pl

[2]University of Warmia and Mazury, Department of Mechanics
and Civil Engineering Constructions, Olsztyn, Poland
e-mail: sz.sawczynski@uwm.edu.pl (corresponding author)

[3]Polish Academy of Sciences, Institute of Hydroengineering, Gdańsk, Poland
e-mail: jarbieg@ibwpan.gda.pl

A b s t r a c t

An experimental and theoretical identification of hydrodynamic equilibrium for sediment transport and bed response to wave motion are considered. The comparison between calculations and the results of laboratory experiments indicates the linear relation between sediment transport rate and the thickness z_m of bed layer in which sediments are in apparent rectilinear motion. This linear relationship allows to use the first order "upwind" numerical scheme of FDM ensuring an accurate solution of equation for changes in bed morphology. However, it is necessary to carry out a decomposition of the sediment transport into transport in onshore direction during wave crest and offshore direction during wave trough. Further, the shape of bed erosion in response to sediment transport coincides with the trapezoid envelope or with part of it, when some sediments still remain within it. Bed erosion area is equal to the one of a rectangle with thickness z_m.

Key words: hydrodynamic equilibrium, linear equation of morphological changes, thickness of eroded/accumulated sediment, sand trap.

1. INTRODUCTION

In the coastal zone there is a strong impact of waves and currents on the bed resulting in almost permanent movement of sediment closely related to the variability of bed shape. The importance of the sediment movement within the coastal areas emphasizes the need for mathematical description of sediment transport and bed morphology. Cyclic erosion and accumulation processes, due to the volume of transported sediment, induce dynamic equilibrium of the local seabed. Thus, to describe the bed morphology changes, it is possible to use conservation laws concerning transport of sediments and to use numeric models in order to determine the prediction of changes in the spatial variation of sediment transport within the coastal zone. Hence, morphological models of the coastal zone are essential to predict changes in bathymetry within the coastal zone. They also allow for the analysis of changes in bed elevations in areas in the vicinity of hydraulic structures at the stage of design, construction and use.

Changes in the elevation of the bottom level can be determined by the solutions of the commonly used equation of conservation of mass (Exner equation, *e.g.*, Yalin and da Silva 2001) for bed sediment transport. For the one-dimensional case this can be written as follows:

$$\frac{\partial z_b}{\partial t} + \frac{1}{1-n_p}\frac{\partial q_{(x)}}{\partial x} = 0 \ , \tag{1}$$

where z_b is the bed level elevation [m], x is a horizontal coordinate [m], t is time [s], n_p is the sediment porosity [-], and $q_{(x)}$ is the sediment transport rate [m^2/s] towards x.

To solve Eq. 1 it is required to use various numeric schemes (Callaghan *et al.* 2006, Chiang and Hsiao 2011, Johnson and Zyserman 2002, Long *et al.* 2008). As shown, among others by Johnson and Zyserman (2002), the numerical schemes used to solve Eq. 1 generate spatial numerical oscillations. These oscillations of morphological models appear as a result of non-linear relation between sediment transport rate $q_{(x)}$ and the elevation of the bed level z_b. The sediment transport rate is also determined on the basis of non-linear, complex hydrodynamic relations. These non-linear relations and numerical errors of the particular sub-models can generate instability and uncertainty of the results of calculation, especially because the nature of these relations is still poorly understood. These models are therefore unable to carry out the exact (physical) prediction of bathymetry in coastal areas, especially in long-term simulations (Johnson and Zyserman 2002).

Over the last few decades, a few techniques, that should improve the stability and accuracy of solutions incorporated into models of bathymetric changes, have been proposed. As shown by Long *et al.* (2008), many of the

latest calculation models contain numerical paradigms which smoothen os-cillations between results describing the evolution of the seabed morpholo-gy. In order to ensure adequate, long-term simulation of bathymetric changes caused by interactions between waves and currents near the seashore, a cal-culation model should enable to control spatial oscillations and guarantee the accuracy of results. Simultaneously, it should reflect the physical aspects of these changes.

In recent years, several morphological models have been verified with respect to the control of oscillations. Callaghan *et al.* (2006) reviewed a few numerical schemes and concluded that the Lax-Wendroff's schemes and all their subsequent modifications are unstable with respect to long-term simu-lated bathymetric changes. It is therefore necessary to expand these schemes, for example by adding such information as artificial viscosity, which will prevent numerical oscillations generated by these methods. Besides, Calla-ghan *et al.* (2006) emphasized the difficulties with describing the bed form propagation phase speed as a vital parameter responsible for the stability of the reviewed numerical schemes.

The stability of numerical schemes is typically referred to the require-ment concerning the Courant number condition: $C_r = C_a \Delta t / \Delta x \leq 1$, where C_a stands for the bed form propagation phase speed. Thus, it is possible to im-prove the stability of numerical models by shortening the duration of simula-tion. It will, however, demand a greater computational power of computers. If the diffusion term is correctly removed from the transport equation, the boundary value of the Courant number can be exceeded (Chiang and Hsiao 2011). This is attainable through the introduction of diffusion constants, which are chosen according to the actual environmental conditions (Cayocca 2001, Chiang *et al.* 2010, Kuroiwa *et al.* 2004, Struiksma *et al.* 1985, Watanabe 1988).

In the authors' approach, it is assumed that due to the shear stress impact on the bottom, the sediment is pulled off directly from the bottom, which is understood as an averaged response of the bed to the particular hydrodynam-ic conditions in time Δt and from the area with a length Δx. The entire sedi-ment moving in saltation motion above the bottom originates exclusively from the bottom (*e.g.*, Bialik 2013, Moreno and Bombardelli 2012). In hy-drodynamic equilibrium conditions, the stream of sediments picked up from the bed with a length Δx, averaged in time Δt, is equal (at each level) to the stream of sediments falling onto the bed. No restrictions are imposed on both Δx and Δt values. The only requirement is that the sediment stream q_x pass-ing through the selected profile (1-1 in Fig. 1a) has a constant value. The sediments have been picked up from the bed with a length Δx, over the time Δt before they pass through the selected profile. Usually, the Δx value is of
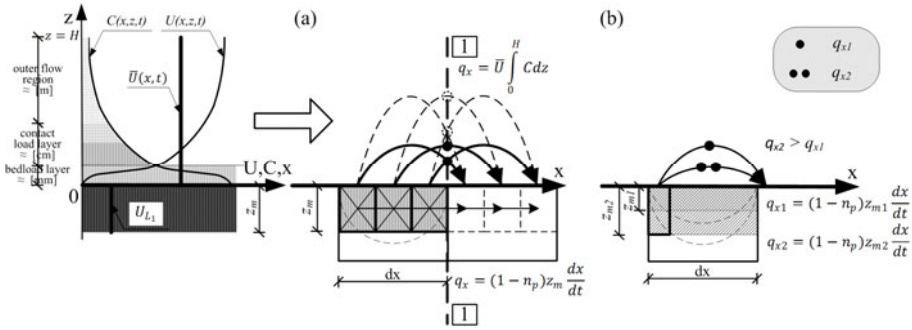
Fig. 1. Scheme of the sediment transport in hydrodynamic equilibrium conditions: (a) sediment stream passing through the 1-1 profile and equilibrium between sediment transport in the bed layer of the thickness z_m and the one above the bottom, and (b) the saltation impulse of sediment grains resulting in transport $q_{x2} > q_{x1}$ and causing bed erosion of the thickness $z_{m2} > z_{m1}$.

the order of 0.05-0.10 m in numerical modelling against laboratory experiments (*e.g.*, van Rijn *et al.* 2005, Walstra *et al.* 2005), while 5-10 m in numerical modelling *versus* field measurements (*e.g.*, Kaczmarek *et al.* 2004).

The time Δt is expected to be correlated with the duration of an erosion process from the bed with a length Δx caused by sediment transport q_x passing through the vertical profile (Fig. 1). It is postulated (see Kaczmarek *et al.* 2004) that erosion depth z_{max} should not exceed the value of 0.015-0.030 m in order to satisfy the requirement that the sediment transport has a constant value in time Δt. This means that the time interval Δt can attain relatively great value when the erosion process caused by sediment transport is not intensive.

2. FORMULA FOR THE PROBLEM

2.1 Theoretical background

The movement of the sediment in the layer above the bottom with the average speed of

$$\bar{U} = \frac{\int_0^H UCdz}{\int_0^H Cdz} \qquad (2)$$

implies a kind of sediment apparent rectilinear motion which takes place in the bed, in the layer with the thickness z_m (Fig. 1). Wherein H determines water depth, C is the volumetric concentration, averaged in time dt, and U means the sediment velocity, averaged in time dt. The thickness z_m means a thickness of $z_m \times dx \times 1$ cell, which is eroded in time dt as a result of the

transport q_x of sediments passing thorough the transverse profile. This thickness is described by the following relation:

$$z_m = \frac{1}{\left(1-n_p\right)} \frac{q_x dt}{dx} \; ,$$ (3)

where q_x means the sediment transport rate in the conditions of hydrodynamic equilibrium, while n_p is the porosity of sediments.

Cell erosion of the thickness z_m proceeds at the forward speed of U_{L_1} :

$$U_{L_1} = \frac{dx}{dt} \; ,$$ (4)

while

$$\frac{\bar{U}}{U_{L_1}} = \kappa \; .$$ (5)

Conditions of the hydrodynamic equilibrium assume that sediment transport rate q_x which takes place in the bed layer of the thickness z_m and which results from Eq. 3 in traditional (*e.g.*, van Rijn 1984) form:

$$q_x = \left(1-n_p\right) U_{L_1} z_m$$ (6)

is equal to the sediment transport above the bottom (Fig. 1). The sediment transport above the bed is described by the following relation:

$$q_x = \bar{U} \int_0^H C dz = U_{L_1} \kappa \int_0^H C dz$$ (7)

from which one can obtain:

$$\left(1-n_p\right) z_m = \kappa \int_0^H C dz \; .$$ (8)

The essential role in initiating the motion of sediment is performed by grain saltation (*e.g.*, Bialik and Czernuszenko 2013, Bialik *et al.* 2012, Wiberg and Smith 1985) (Fig. 1a). Figure 1a shows that a saltation impulse moves a total number of sediment grains in a jump over the length dx. Thus, the saltation impulse (Fig. 1b) of sediment grains resulting in transport $q_{x2} > q_{x1}$ causes in time dt the erosion of the seabed of the thickness $z_{m2} > z_{m1}$, in a cell of the length dx. This means that under the hydrodynamic equilibrium the thickness z_m is a function of the sediment transport rate whereas forward speed U_{L_1}, described by Eq. 4, does not depend on the said rate. In other words, under hydrodynamic equilibrium condition, the sediment transport above the bed is equal to the one inside the bed for any length

dx and time dt as long as the forward speed U_{L_1} does not depend on sediment transport rate q_x and coordinate x. Consequently, a grain jump above the bottom with any trajectory height, which corresponds to an apparent stroke of a layer inside the bed, is defined by Eq. 7 also for any length dx and time dt. The only requirement is that the sediment transport has a constant value in time dt and along the entire length dx.

These conclusions coincide with what Einstein proposed in 1950. Based on experimental observations, Einstein (1950) assumed that the averaged distance travelled by a sand particle between erosion and subsequent deposition, is simply proportional to the grain diameter and independent of the hydraulic conditions and the amount of sediment in motion. It means that every jump of any length is equally probable at every x point. As a result, any x point is the origin of different jumps, altogether resulting in sediment transport q_x. In that case, the stream of sediments picked up in time dt from every x point along the distance dx is equal to the stream of sediments following onto the bed. Figure 1 shows the simplified case, when all considered jumps are defined by one particular length. In addition, for any value of dx, the saltation motion of sediment grains with the transport rate of, respectively, q_{x1} and q_{x2} will trigger (after the sufficiently long times t_2 and t_1) the erosion of the bed with the same thickness z_{max}, so from Eq. 3 it results that:

$$q_{x2}t_2 = q_{x1}t_1 = \left(1 - n_p\right)z_{max}dx \ . \tag{9}$$

The thickness z_{max} is therefore also the value which is independent of the sediment transport rate q_x and it should be determined on the basis of experiments.

A good empirical illustration of a motion of sediment under the hydrodynamic equilibrium is the flow of water over a box filled with sediment (Fig. 1). If the bottom outside the box is made of concrete, then it is not difficult to image that all the sediment moving above the bottom originates from that bed structure. It can be assumed that under determined water flow conditions, the erosion of sediments from the box (Fig. 1a) is caused by the simultaneous (along the entire length dx) grain strokes with the length dx and with the total size z_m or by corresponding to them grain jumps with any trajectory height. In such a case, the bottom sediment erosion is constant at the entire dx length and it amounts to z_m (Fig. 1).

This picture is not exclusively approximated by situation with jumps defined by one particular length dx. It is also simplified because of the fact that the material is transported from the central part of the dx area and from only one extreme area (Fig. 1a) at any moment and for every height of jump trajectory. Indeed, the erosion of the bed is not constant over the entire dx length. The middle part is eroded twice strongly than extreme areas. It is

postulated, however, that the actual size of the bed erosion − in the entire area of dx − is equal to the value calculated assuming a constant erosion along dx.

2.2 Scope of the work

The main aim of this study is to demonstrate (on the basis of the results obtained from laboratory experiments of "IBW PAN 1996" (Kaczmarek 1999) and "IBW PAN 2010" (Sawczyński 2012)) the postulate of sediment transport in hydrodynamic equilibrium conditions. Measurements were carried out in the wave flume of the Institute of Hydroengineering, Polish Academy of Sciences (IBW PAN) in Gdańsk in 2010 (Sawczyński 2012). The aim of the experiment was to make measurements of the thickness of eroded bed sediment $(z_m)_e$ in the context of comparative analysis of the results of "IBW PAN 1996" experiment, during which the measurements of cumulative thickness $(z_m)_a$ were carried out (Kaczmarek 1999). Within the frame of the study, two tests, T4 and T5, were performed for the eroded thickness of the sediment. The laboratory experiments T4 and T5 indicate that the bathymetry of erosion area coincides with the trapezoid envelope or with a part of it, when a part of sediments still remains within it. In both cases the erosion area is equal to the rectangle one with thickness $(z_m)_e$ along the entire length dx.

The results of the measurements are to be used as a confirmation of the hypothesis of hydrodynamic equilibrium, *i.e.*, the possibility to adopt the linear relation 6 between the sediment transport rate q_x and thickness z_m of a bed layer where densely packed grains of the sediment participate in the apparent translational and rectilinear motion. This motion is caused by the actual sediment grain flow over the bed, in the form of steps. This linear relationship 6 gives the opportunity to use a numerical scheme of the first order "upwind" the finite difference method (FDM). This method provides an accurate solution of the equation describing the changes in bathymetry in time and space (Kaczmarek *et al.* 2011, Sawczyński 2012):

$$\frac{\partial z_m}{\partial t} + \frac{1}{1-n_p}\frac{\partial q_x}{\partial x} = 0 \ . \tag{10}$$

At the same time, the following relation is valid:

$$z_b\left(x, t+dt\right) = z_b\left(x, t\right) + \frac{\partial z_m}{\partial t}dt \ . \tag{11}$$

Thus, as long as the sediment transport remains in the hydrodynamic equilibrium conditions, the first order "upwind" numerical scheme of FDM can be used to obtain an accurate solution of Eq. 10. However, as it is shown

in Section 3, the decomposition of the sediment transport into the one in on-shore direction during wave crest and offshore direction during wave trough should be done.

In contrast to the proposal described above, numerical schemes based on non-linear relation between the sediment transport rate q_x and the elevation of the bed level z_b, used in Eq. 1 to describe changes in bathymetry, are unstable ones and generate oscillations in calculation results. It was shown inter alia by Sawczyński (2012) and Szymkiewicz (1999).

It should be emphasized that the postulate of the linear relationship 4 between sediment transport rate q_x and the thickness z_m of layer in apparent motion does not correspond to the postulate of a linear relationship between the sediment transport rate q_x and the elevation of the bed level z_b. The last relationship is still postulated as non-linear.

3. SEDIMENT TRANSPORT IN THE WAVE MOTION

The instantaneous velocities u and sediment concentrations c in water flow under surface waves can be written in the following form:

$$u = \langle U \rangle + \tilde{U} \ , \tag{12}$$

$$c = \langle C \rangle + \tilde{C} \ , \tag{13}$$

where $\langle ... \rangle = \dfrac{1}{T} \displaystyle\int_0^T (...)\, dt$ means the averaged value in time of the wave period (T), while $\tilde{U}$ and $\tilde{C}$ are the oscillating components of the velocity and concentration, respectively.

The product of velocity and concentration is defined as the volumetric stream related to the unitary area presented as follows:

$$\langle \phi \rangle = \langle U \rangle \langle C \rangle + \langle \tilde{U} \tilde{C} \rangle = U_L \langle C \rangle \ . \tag{14}$$

The velocity U_L, according to Plumb's (1979) study, is a kind of Lagrange's speed similarly to forward speed U_{L_1} of the bed layer with the thickness z_m, as determined by Eq. 3.

It is postulated to implement the decomposition of the sediment flow induced by waves into ones during wave crest (T_c) and wave trough (T_t), respectively. Consequently:

$$T = T_c + T_t \ . \tag{15}$$

The parameters related to the wave crest and wave trough are marked as $(...^+)$ and $(...^-)$, respectively. Thus, the volumetric stream may be presented in the following form:

$$\langle \phi \rangle = \langle \phi^+ + \phi^- \rangle \ , \tag{16}$$

where

$$\langle \phi^+ \rangle = U_L^+ \langle C^+ \rangle \ , \tag{17}$$

$$\langle \phi^- \rangle = U_L^- \langle C^- \rangle \ , \tag{18}$$

while $\langle ...^+ \rangle = \dfrac{1}{T} \displaystyle\int_0^{T_C} \left(...^+ \right) dt$ and $\langle ...^- \rangle = \dfrac{1}{T} \displaystyle\int_{T_C}^{T} \left(...^- \right) dt$ are averaged values within the wave period (T).

The complete sediment transport rate is a result of integration of the stream of $\langle \phi^{+/-} \rangle$ over the depth H, as follows:

$$\int_0^H \langle \phi^{+/-} \rangle dz = q_x^{+/-} \ , \tag{19}$$

which corresponds to the sediment transport rate $q_x^{+/-}$ in the layer of the thickness $z_m^{+/-}$ with the forward speed $U_{L_1}^{+/-}$ in hydrodynamic equilibrium conditions, while:

$$z_m = z_m^+ + z_m^- \tag{20}$$

and

$$z_m^{+/-} = \frac{1}{\left(1 - n_p\right)} \frac{q_x^{+/-} dt}{dx^{+/-}} \ . \tag{21}$$

The symbol $(...^{+/-})$ combines two cases, *i.e.*, the one described by $(...^+)$ for the wave crest and another one, by $(...^-)$, for the wave trough.

Using the Eqs. 2, 4, and 5 in comparison with 17, 18, and 19, relation 21 may be transformed to the following expression describing the sediment transport rate:

$$q_x^{+/-} = \left(1 - n_p\right) z_m^{+/-} U_{L_1}^{+/-} = U_{L_1}^{+/-} \kappa^{+/-} \int_0^H \langle C^{+/-} \rangle dz \ , \tag{22}$$

where

$$\kappa^{+/-} = \frac{U_L^{+/-}}{U_{L_1}^{+/-}} \ . \tag{23}$$

Next, using relation 22, Eq. 10 takes the following form:

$$\frac{\partial z_m}{\partial t} + \frac{1}{\left(1 - n_p\right)} \left(\frac{\partial q_x^+}{\partial x^+} + \frac{\partial q_x^-}{\partial x^-} \right) = 0 \ . \tag{24}$$

Herein, the following should be taken into consideration:

$$\frac{\partial q_x^+}{\partial x^+} \neq \frac{\partial q_x^-}{\partial x^-} , \tag{25}$$

when waves propagate over inclined bed. Indeed, the values of q_x^+ and q_x^- are interpreted as sediment transport rates under hydrodynamic equilibrium conditions during wave crest and trough duration, respectively. Both values are averaged within wave period T. These sediment transport rates (in accordance with Eq. 22) are in linear relations with thicknesses $z_m^{+/-}$. However, in view of Eq. 11, the relation between these sediment transport rates q_x^+ and q_x^- and bed level of z_b is still non-linear.

In order to solve Eq. 24, the first order "upwind" numerical scheme of FDM is proposed, in which for any assumed length dx it is possible to calculate the time interval dt according to Eq. 3:

$$dt = \frac{\left(1 - n_p\right) z_{\max} dx}{q_{\mathrm{rep}}^+ + \left| q_{\mathrm{rep}}^- \right|} , \tag{26}$$

where the maximum value of $\left(q_x^+ + \left| q_x^- \right| \right)$ in a given calculation area and within a specified time interval is assumed as the representative $\left(q_{\mathrm{rep}}^+ + \left| q_{\mathrm{rep}}^- \right| \right)$. Thus, it is necessary to determine the values of q_x^+ and q_x^- in a given calculation area and within a specified time interval. It is postulated to use the three layer sediment transport model introduced by Kaczmarek *et al.* (2004) for hydrodynamic equilibrium conditions.

The model assumes that the movement of sediment is carried out in three layers (Fig. 1): bedload layer, contact load layer, and outer flow region, as a result of the shear stress impact on the bottom. In the area of each layer there is a different character of the deposit movement and the momentum exchange between the water and sediment particles. Hence they are described by various equations with boundary conditions permitting matching of velocities and concentrations in considered areas as to ensure the continuity of sediment movement description.

The model assumes that all fractions in the bedload layer move with the same speed in the form of water and solid mixture. It is assumed that the interactions between the sediment fractions are so strong that finer fractions are slowed down by the coarser ones, and finally all fractions move with the same speed. Thus, this layer does not apply to the simple summation of transport flow for individual fractions treated as the homogeneous sediment.

The onshore sediment transport intensity in the bedload layer (for the duration of wave crest) q_{bx}^{+} and offshore q_{bx}^{-} (for the duration of wave trough) are defined as (see Sawczyński *et al.* 2011):

$$q_{bx}^{+} = \frac{1}{T} \int_{0}^{T_c} \left(\int_{0}^{\delta_{br}^{+}} u_{br}^{+}(z',t) c_{br}^{+}(z',t) dz' \right) dt \tag{27}$$

and

$$q_{bx}^{-} = \frac{1}{T} \int_{T_c}^{T} \left(\int_{0}^{\delta_{br}^{-}} u_{br}^{-}(z',t) c_{br}^{-}(z',t) dz' \right) dt \;, \tag{28}$$

where T_c is wave crest duration, T is wave period, $\delta_{br}^{+/-}$ represents bedload layer thickness, $u_{br}^{+/-}$ is instantaneous sediment velocity in bedload layer, $c_{br}^{+/-}$ is instantaneous sediment concentration in bedload layer, z' is elevation, while z' axis is directed vertically down.

The mathematical modelling takes into account the fact that the most intensive vertical sorting takes place in the process of raising grains in the contact load layer over the bottom. In the contact load layer, turbulent water pulsations and chaotic collisions between particles cause very strong transport differentiation of individual sediment fractions. Very close to the bottom, in a sublayer, bed slip speed is strongly revealed and there is a very strong interaction between the individual fractions. Further from the bottom, these interactions become weaker. However, at this level the concentration of the *i*-th fraction is still so high as to cause turbulence suppression which is assumed to be dependent on the grain diameter d_i. As a result of interactions described above, each *i*-th fraction moves with its own speed and is characterized by its own concentration. The interactions between fractions when coarser ones are accelerated by the finer ones result in the increase of speed of coarser fractions in the mixture. Further, coarser fraction concentrations are greater than the ones which would have been if the bottom had been homogenous and made of only one, corresponding fraction.

The onshore q_{cx}^{+} and the offshore q_{cx}^{-} sediment transport rates may be described as (see Sawczyński *et al.* 2011):

$$q_{cx}^{+} = \sum_{i=1}^{N} n_i \left(\frac{1}{T} \int_{0}^{T_c} \left(\int_{0}^{\delta_{cr}} u_{ci}^{+}(z,t) c_{ci}^{+}(z,t) dz \right) dt \right) \tag{29}$$

and

$$q_{cx}^{-} = \sum_{i=1}^{N} n_i \left(\frac{1}{T} \int_{T_c}^{T} \left(\int_{0}^{\delta_{cr}} u_{ci}^{-}(z,t) c_{ci}^{-}(z,t) dz \right) dt \right) \;, \tag{30}$$

where δ_{cr} represents the contact load layer thickness, n_i is the percentage content of the i-th fraction, $u_{ci}^{+/-}$ is instantaneous velocity of the i-th fraction, and $c_{ci}^{+/-}$ denotes instantaneous concentration of the i-th fraction in the contact load layer.

In the outer layer it is assumed that the particle size distribution of the transported sediment does not change along the distance from the bottom. The vertical distribution of concentration in this layer is described by a power function.

The sediment transport flow in the outer layer can be expressed by the following formulas (see Sawczyński *et al.* 2011):

$$q_{ox}^- = \int_{\delta_{cr}}^{H} U_{ox}^-(z) C_0(z)\, dz \ , \tag{31}$$

$$q_{oy} = \int_{\delta_{cr}}^{H} U_{oy}(z) C_0(z)\, dz \ , \tag{32}$$

where C_0 means the value of suspended sediment concentration averaged in time, and U_{ox}^- and U_{0y} are velocities of the return current (directed towards the sea) and alongshore current, respectively.

4.   LABORATORY TESTS

4.1 Measurements of eroded thickness $(z_m)_e$. Experiment of "IBW PAN 2010"

The wave flume in which the measurements of eroded thicknesses are conducted is 64.1 m long, 0.6 m wide, and 1.4 m high (Fig. 2). The vertical, mutually parallel channel walls are made of glass, which allows free observation of ongoing researches. The bottom of the channel is made of aluminium alloys.

Programmable and controlled by the computer and hydraulically driven piston-type flap generator provides the ability to produce waves with heights up to 0.6 m and any length of allocated time over 0.5 s. Wave forcing element is a rigid vertical plate of the generator, subjected to horizontal oscillatory movements similarly to the movement of the piston inside the cylinder. The generator is located independently of the supporting structure of the flume, so that the transmission of undesired vibrations is reduced.

The final part of the flume is equipped with original wave energy absorber, made in the form of porous slope with 1:7 inclination and is composed of a dozen of plastic matt layers. That kind of structure allows to reduce the effects of reflected waves from the rear wall of the flume on wave propagation.

Fig. 2. Wave flume at IBW PAN in Gdańsk – "IBW PAN 2010" experiment (photo by Ł. Sobczak).

In order to measure the thickness of eroded sediment, the model of sand trap was used. This is made of waterproof plywood and plastic plates (Sawczyński 2012). Then, the model was placed on the bottom of the flume. The model of the sand trap consisted of two slopes, both inclined at 1:18, and the central box located between the slopes. The sand trap box was composed of three smaller parts: the central one of the dimensions of $20.0 \times 57.8 \times 10.0$ cm and the two extreme parts, each one of dimensions of $57.8 \times 10.0 \times 10.0$ cm. The central box was filled with sand. Diagram of the experiment is presented in Fig. 3.
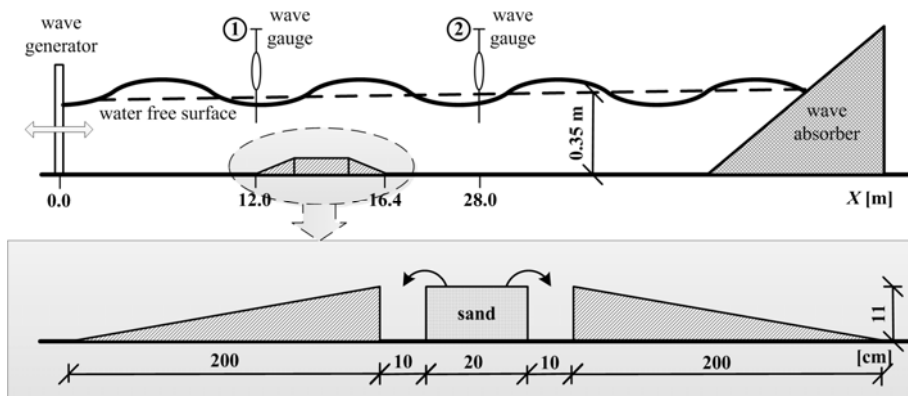
Fig. 3. Diagram of the "IBW PAN 2010" experiment.

In order to determine the grain size of the sediment, the particle laser measure device of Analysette 22 MicroTec plus was used. The device measures with two lasers. The detector captures 108 measuring channels. The meter, by the usage of laser diffraction, allowed to analyze samples of a small amount of solid fractions in the range of diameters 0.01-2000 μm. Sediment used in experiments was sand originated from the beach of seaside town Lubiatowo, situated at south Baltic coast in Poland. For this purpose, the particle size laser measuring device was used. Particle size analyses were performed using the "wet method" for the two randomly taken sand samples. Each of them was analyzed twice and identical results were obtained. On the basis of those analysis, the following representative sand parameters were acquired – the diameter $d_{50} = 0.27$ mm and $d_{90} = 0.42$ mm.

Two resistance wave probes were used for measurements of water free surface elevations. A resistive element of these probes is water layer between two vertical non-insulated electrodes. The wave height measurement is based on a variation of electrical resistance of water between two electrodes. The device accuracy is ± 1 mm.

For measurements of the flow velocity it was necessary to use the ultrasonic Acoustic Doppler Velocimeter (ADV) which enabled simultaneous measurements of the three velocity components in a configurable range of ± 0.01, 0.1, 0.3, 2.0, 4.0, and 7.0 m/s with an accuracy of ± 0.5% of measured value ± 1 mm/s. The sampling rate of the device is 1-64 Hz, while the sound frequency is 6 MHz.

Within the frame of the experiment, the two measuring tests (T4 and T5) were carried out for the same sand and different wave conditions (Sawczyński 2012). The change in wave conditions concerned only the wave height, while the wave period remained unchanged. The computer software controlled wave generation process and regular and sine waves were to produce. Wave parameters were recorded by the two wave probes placed along the flume, downstream and upstream the sand trap model (Fig. 3). In Table 1 there are shown basic parameters of the experiment. The wave height of each T4 and T5 test marked by symbol H_w determines the time averaged value taken from two wave probes.

Measurements of wave height and period took place using two wave probes arranged along the flume, as shown in Fig. 3. Because of the fact that the wave parameters were kept constant during undulation, the recording of wave parameters had been made fragmentarily at selected time intervals.

Wave height records with duration of twenty seconds of T4 and T5 tests are shown in Fig. 4 as an example while time averaged values of the wave height are shown in Table 2.

It can be seen from Fig. 4 that regular sine waves were subjected to transformation process due to their propagation along the flume. The records

Table 1

Basic data of "IBW PAN 2010" experiment

Parameter	Symbol	Value	Unit
Water depth	h_0	0.35	[m]
Wave height T4/T5	H_w	6.77 / 4.95	[cm]
Undulation duration time T4/T5	T_w	25/60	[min]
Wave peak period T4/T5	T_p	2.0 / 2.0	[s]
Water temperature	T	18.0	[°C]
Representative diameter of sand grains	d_{50}; d_{90}	0.27; 0.42	[mm]
Sand density	ρ_s	2650	[kg/m^3]
Fluid density	ρ_w	1000	[kg/m^3]
Sand porosity	n_p	0.4	[–]

Table 2

Time averaged wave heights of tests T4 and T5
– "IBW PAN 2010" experiment

Test no.	Wave height H_w [cm]		
	Probe no. 1	Probe no. 2	Average
T4	7.65	5.90	6.77
T5	5.70	4.20	4.95

of wave parameters taken from the probe 1 installed upstream the model (Fig. 3) for both tests, T4 and T5, show that amplitudes of wave crest are bigger than of the trough. However, the difference in these amplitudes is much bigger in T4 test when higher waves were generated. In addition to the amplitude difference, it is also possible to observe shape differentiation between wave crests and troughs. In both tests, the elevations during wave crests were higher and steeper than during troughs. Simultaneously, the duration of wave crests was shorter than that of troughs. Moreover, this description of waves is identical with the one of the second Stokes' wave approach theory.

The sand originally located in the central box of the sand trap was transported to the two extreme parts as a result of undulation of the free water surface (Fig. 3). The measurement diagram is shown in Fig. 6. During wave crest, sand was transported towards the right trap and during wave trough towards the left part. In order to determine the volume of sand accumulated in extreme traps, each time after completion of the undulation, the mixture of sand and water was pumped from the two extreme traps. The siphoning method was used with the silicone hoses. The samples were then dried and

(a)

(b)
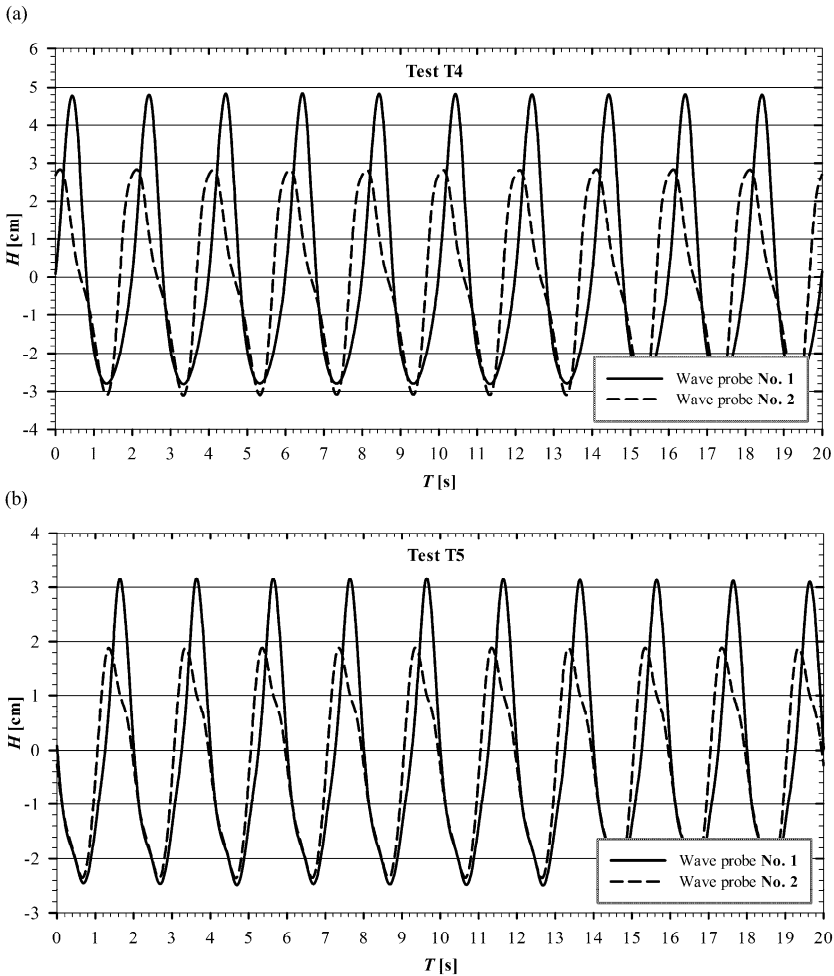
Fig. 4. "IBW PAN 2010" experiment: record of wave elevations (wave probe no. 1 – solid line, wave probe no. 2 – dashed line): (a) test T4, and (b) test T5.

the sand volume taken from each trap was determined. On this basis, the thickness of the layer eroded from the central box was calculated. The results are shown in Table 3.

It is worth emphasizing that the duration of T4 test was three times smaller than T5 test's. Recorded wave height during the test T4 was greater than the one of test T5 and it was a reason why the total amount of eroded sediment in T4 test was almost twice bigger than in T5 test. Further, it is clear that for both measurement tests, a greater amount of sediment taken from the central box was deposited in the right trap (Table 3). This means

(a)
Test T4

(b)
Test T5
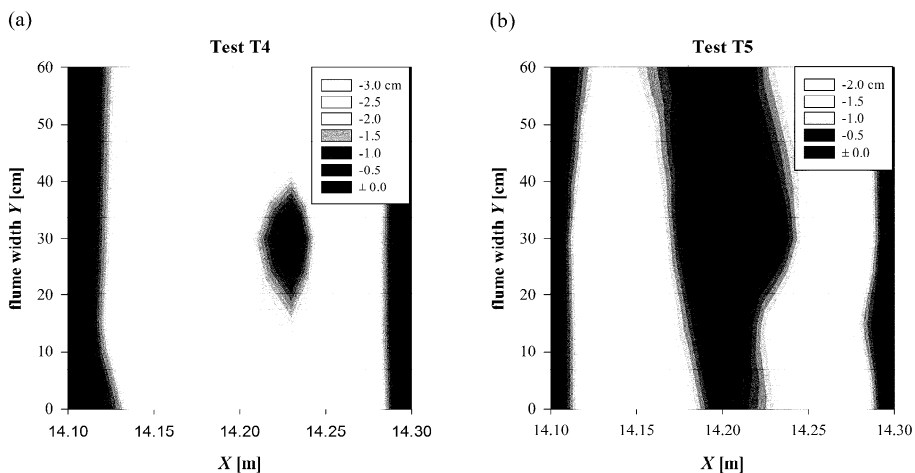
Fig. 5. Final bathymetry of "IBW PAN 2010" experiment: (a) test T4, and (b) test T5.
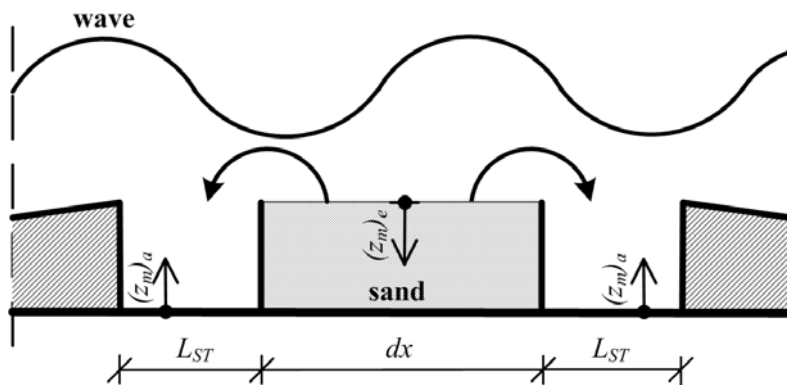
Fig. 6. Diagram of eroded thickness measurements $(z_m)_e$ – "IBW PAN 2010" experiment.

Table 3

Measured amount of eroded sand – "IBW PAN 2010" experiment

Test no.	Undulation time [min]	Amount of sediment in the left box [cm³]	Amount of sediment in the right box [cm³]	Total amount of the sediment [cm³]	Erosion thickness $(z_m)_e$ [cm]
T4	25	780	1150	1930	1.67
T5	60	460	580	1040	0.90

that the greater amount of sediment was transported during the phase of wave crest than during wave trough. Of course, it is a consequence of wave asymmetry.

After completion of the undulation, the bathymetry was carried out manually in the central box. The measurements were taken point by point every 2.0÷2.5 cm in the longitudinal direction and approximately every 10 cm in the transversal direction. Total number of measurements per test was 35. Intermediate data was obtained using linear interpolation technique. The results of final bathymetry of T4 and T5 tests are presented in Fig. 5. It is noticeable that, for both measurement tests, erosion on the left side of the central box was bigger than on the right side. This reflects the dominance of sand transport during wave crest, resulting from the asymmetry of propagating waves (see Fig. 4). Moreover, it documents the dominant impact of wave crest on the process of bathymetric changes.

4.2 Measurements of accumulated thickness $(z_m)_a$. Experiment of "IBW PAN 1996"

The experiment was conducted at the Institute of Hydroengineering, Polish Academy of Sciences in Gdańsk in 1996 (Kaczmarek 1999) in the wave flume 0.5 m wide, 25.5 m long, and maximal possible filling of up to 0.7 m. In the wave flume, in conditions of constant filling of $h = 0.5$ m, the regular (tests 1, 2, 3, 4, 11, 12) and irregular (tests 5, 6, 7, 8, 9, 10) undulations were generated. The experiment was repeated several times for each set of hydrodynamic parameters. A total number of 141 measurements were made for 12 tests. Figure 7 presents the diagram of the experiment.

At the beginning and at the end of the flume, its bottom was covered by concrete slabs with a thickness of 0.080 m. The central part of the flume was filled with sand with representative diameters $d_{50} = 0.22$ mm and $d_{90} = 0.42$ mm. The sand thickness was equal to the thickness of the concrete slab. The length of the section filled with sand was 7.0 m. At a distance of 2.0 m from the end of the section of concrete slab, the sand trap was situated. It
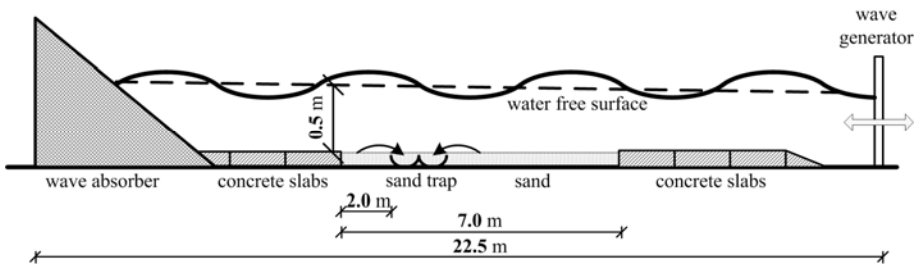
Fig. 7. Diagram of the "IBW PAN 1996" experiment.

was used to collect trapped sand and hence it was possible to measure accumulated thickness of the sediment.

The main objective of the measurements − at the contemporary stage of research − was to determine the amount of sediment transported mainly in the bedload layer, in the direction of outflow and inflow by measuring the amount of sediment which flows into the sand trap. The sand trap consisted of two separate chambers with a height of 0.063 m. During the experiment, the wave conditions were generated within a range for which Shields parameter $\Theta_{2.5}$ determined for the representative diameter $d_{50} = 0.22$ mm, contained within the range $0.085 \leq \Theta_{2.5} \leq 0.4$, which corresponded to the regime of rippled bed. Thus, measurements were made after the time required for the full development of bed forms, *i.e.*, after 25-60 minute duration of continuous undulation. By this time, the sand trap was covered with a lid. Then, the cover was removed and undulation was continued for 1.2 to 15 minutes. In this time, the trap was filled with sand. In the next stage of the experiment, the sand was sucked from the sand trap outside the flume and then it was weighed. Finally, by determining the volume of sand pumped from inside the sand trap, accumulated thickness was determined.

5. COMPARISON BETWEEN RESULTS OF NUMERICAL AND PHYSICAL MODEL

5.1 Input data for calculations

The percentage contents of different sediment fractions for both experiments are shown in Table 4. This data was used as an input one in the calculations of sediment transport rates q_x^+ and q_x^- using Eqs. 27-32. The input data maintains measured indicators of d_{50} and d_{90}.

The wave parameters (wave heights H and periods T), taken from Kaczmarek (1999), were used as an input data in calculations conducted for an experiment of "IBW PAN 1996". For all 12 measurement tests, the sine waves theory was used in modelling.

In case of the "IBW PAN 2010" experiment, the first wave probe was installed just in front of the sand trap, at a distance of 12 m from the wave generator. The second wave probe was situated 28 m from the wave generator (Fig. 3). In view of the fact that the first wave probe was situated just in front of the sand trap, it was reasonable to treat its wave record as an input data for sediment transport model. The data concerns both the wave height and the shape of the wave run. Basing on the wave first probe record, it was decided to use the second Stokes' approximation in order to reproduce shape of the wave run. The proposed wave approximation was confirmed by the second wave probe, although further wave deformation had been observed due to its propagation along the flume, including the distance over the sand trap.

Table 4

Percentage contents of sediment fractions used in calculations
performed for experiments of "IBW PAN 1996" and "IBW PAN 2010"

IBW PAN 1996		IBW PAN 2010	
$d_{50} = 0.22$ mm; $d_{90} = 0.42$ mm		$d_{50} = 0.27$ mm; $d_{90} = 0.42$ mm	
Fraction diameter [mm]	Percentage fraction content [%]	Fraction diameter [mm]	Percentage fraction content [%]
0.10	7	0.08	32
0.18	23	0.25	10
0.20	20	0.28	10
0.22	8	0.31	21
0.26	8	0.37	9
0.28	13	0.41	7
0.30	8	0.45	5
0.32	6	0.50	6
0.34	4		
0.36	3		

5.2 Calculation results in comparison to measurements

Consideration regarding sediment transport under hydrodynamic equilibrium
conditions involves the discussion of Eq. 22 linearity. This linearity seems to
be confirmed by the results from "IBW PAN 1996" and "IBW PAN 2010"
experiments.

Using the Eq. 22, with respect to the above-mentioned experiments, it is
possible to write as follows:

$$\frac{\left(q_x^+ + \left|q_x^-\right|\right)}{L_{ST}} = \frac{\text{Vol}}{B_{ST} L_{ST} dt} = \frac{\left(1 - n_p\right)\left(z_m^+ + z_m^-\right)}{dt} \,, \tag{33}$$

where B_{ST} and L_{ST} are the sand trap width [m] and length [m], respectively,
Vol/dt denotes the volume of sediment motion in time [m/s]; it is assumed
that $dx = L_{ST}$.

In order to compare the calculation results of the bed layer thicknesses
$(z_m^+ + z_m^-)_{\text{cal}}$, where sediment is in apparent rectilinear motion, with the
measurement results $(z_m^+ + z_m^-)_{\text{meas}}$, it is convenient to present Eq. 33 in the
following form:

$$\left(1 - n_p\right)\frac{\left(z_m^+ + z_m^-\right)_{\text{cal}}}{dt} = \frac{\left(q_x^+ + \left|q_x^-\right|\right)_{\text{cal}}}{L_{ST}} \,, \tag{34}$$

while

$$\left(1-n_p\right)\frac{\left(z_m^+ + z_m^-\right)_{\text{meas}}}{\Delta t_{\text{meas}}} = \frac{\left(q_x^+ + \left|q_x^-\right|\right)_{\text{meas}}}{L_{ST}} = \frac{\text{Vol}_{\text{meas}}}{B_{ST}L_{ST}\Delta t_{\text{meas}}} \ , \tag{35}$$

where $\left(q_x^+ + \left|q_x^-\right|\right)_{\text{meas}}$ denotes the measurement sediment transport rate determined on the basis of Vol_{meas} (see Eq. 35), $\left(q_x^+ + \left|q_x^-\right|\right)_{\text{cal}}$ is the sediment transport rate calculated using Eqs. 27-32, Vol_{meas} is the measured volume of the sediment stopped in (or eroded from) the sand trap in time Δt_{meas} .

The results of the comparison are presented in Fig. 8. The dashed lines shown in Fig. 8 are inclined in relations 1:0.75, 1:1, and 1:1.5. The presented cloud of points is arranged in a very narrow area, in the close vicinity of the straight line inclined at the angle of 45°. This convergence of results indicates a very good compatibility between results of calculations and results obtained from the measurements.
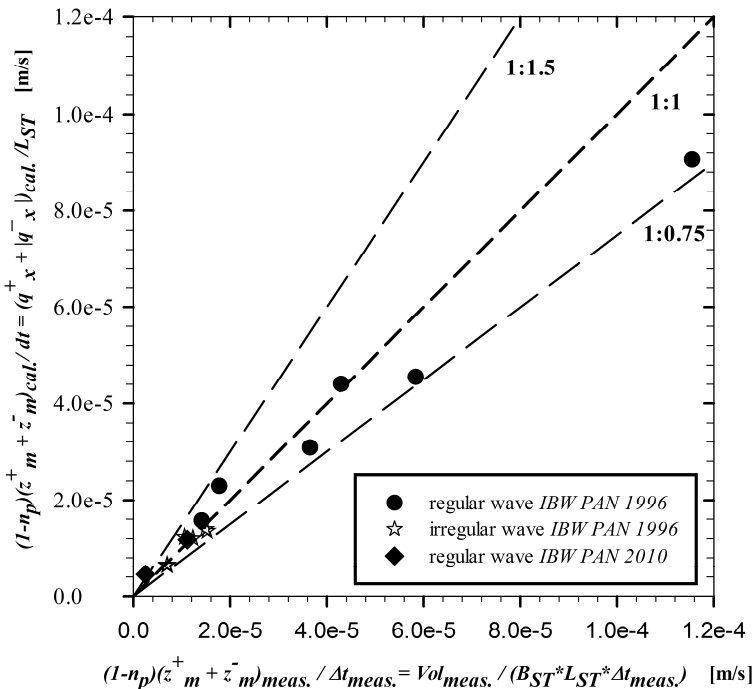
Fig. 8. Comparison between calculations and measurements of $\left(z_m^+ + z_m^-\right)$; black symbols for regular waves: circles – experiment of "IBW PAN 1996", diamonds – experiment of "IBW PAN 2010"; white star symbols for irregular waves – experiment of "IBW PAN 1996"; inclination of dashed lines: 1:0.75, 1:1, and 1:1.5.

The analysis of linearity of the relationship between the results of calculated sediment transport rates $\left(q_x^+ + \left|q_x^-\right|\right)_{cal}$ and the results of measurements $\left(1 - n_p\right)\dfrac{(z_m^+ + z_m^-)_{meas}}{\Delta t_{meas}}$ are presented in Fig. 9. Experimental values were presented as averaged values from the results obtained during measurements. The relation between calculated values of sediment transport rate $\left(q_x^+ + \left|q_x^-\right|\right)_{cal}$ and measured values $\left(1 - n_p\right)\dfrac{(z_m^+ + z_m^-)_{meas}}{\Delta t_{meas}}$ of the eroded ("IBW PAN 2010") or accumulated ("IBW PAN 1996") thicknesses, clearly expresses its linear character. The intersection of a straight line with the elevation axis (Fig. 9) seems to indicate the linear relationship validity only for such sediment transport rate that allows to establish the minimal thickness of $(z_{m0}^+ + z_{m0}^-)$ in time of $dt = \Delta t_{meas}$ in the sand trap of the length $dx = L_{ST}$. In other words, the sediment transport rate smaller than the one resulting from
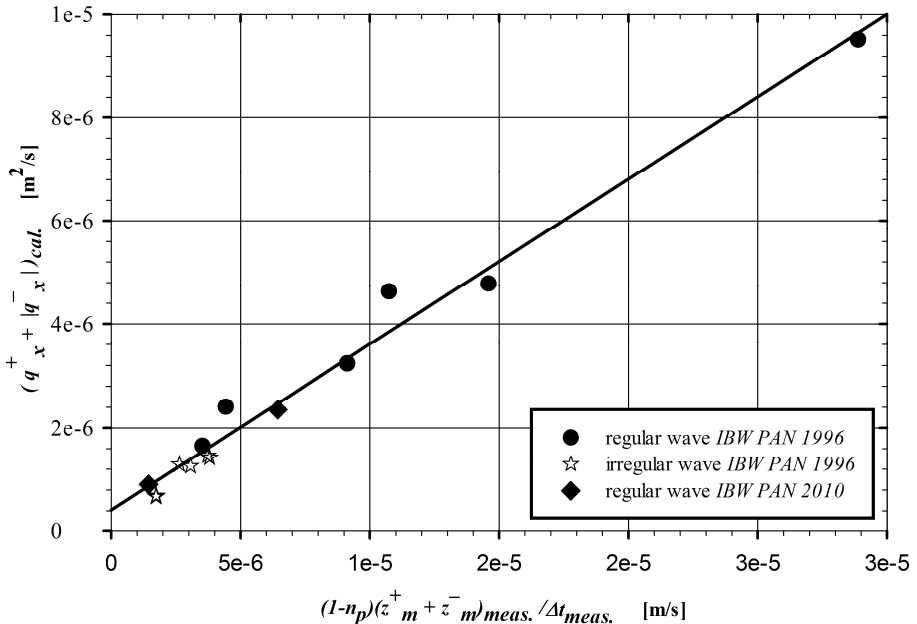
Fig. 9. The analysis of linearity of the relationship between results of sediment transport rate calculations and the results of measurements of eroded/accumulated thicknesses; black symbols for regular waves: circles – experiment of "IBW PAN 1996", diamonds – experiment of "IBW PAN 2010"; white star symbols for irregular waves – experiment of "IBW PAN 1996"; best fit – solid line.

the intersection of the best fit with the elevation axis does not fill the trap with sediment of minimal thickness along the entire length and in time Δt_{meas}.

Equation 33 may be rewritten in the following form:

$$\left(q_x^+ + \left| q_x^- \right| \right) dt = \left(1 - n_p \right) \left(z_m^+ + z_m^- \right) L_{ST} \ . \tag{36}$$

Using Eqs. 22 and 23, it is possible to transform Eq. 36 into the following form:

$$
\begin{aligned}
\left(q_x^+ + \left| q_x^- \right| \right) dt &= \left(1 - n_p \right) \left(\kappa^+ \int_{z_m^+}^{H} \left\langle C^+ \right\rangle dz + \kappa^- \int_{z_m^-}^{H} \left\langle C^- \right\rangle dz \right) L_{ST} \\
&= \left(1 - n_p \right) \left[\left(\kappa^+ - 1 \right) + 1 \right] z_{m0}^+ L_{ST} + \left(1 - n_p \right) \left[\left(\kappa^- - 1 \right) + 1 \right] z_{m0}^- L_{ST} \\
&= \left(1 - n_p \right) \left(\kappa^+ - 1 \right) z_{m0}^+ L_{ST} + \left(1 - n_p \right) \left(\kappa^- - 1 \right) z_{m0}^- L_{ST} + \left(1 - n_p \right) \left(z_{m0}^+ + z_{m0}^- \right) L_{ST}
\end{aligned}
\tag{37}
$$

from which

$$\left(q_x^+ + \left| q_x^- \right| \right) = \left(1 - n_p \right) \left(\kappa^+ - 1 \right) z_{m0}^+ \frac{L_{ST}}{dt} + \left(1 - n_p \right) \left(\kappa^- - 1 \right) z_{m0}^- \frac{L_{ST}}{dt} + \left(q_{x0}^+ + \left| q_{x0}^- \right| \right) , \tag{38}$$

where

$$\left(q_{x0}^+ + \left| q_{x0}^- \right| \right) = \left(1 - n_p \right) + \left(z_{m0}^+ + z_{m0}^- \right) \frac{L_{ST}}{dt} \ . \tag{39}$$

To clarify, it is more convenient to present Eq. 38 in a slightly different form:

$$\left(q_x^+ + \left| q_x^- \right| \right)_{\text{cal}} - \left(q_{x0}^+ + \left| q_{x0}^- \right| \right) = \left(1 - n_p \right) \frac{L_{ST}}{\Delta t_{\text{meas}}} \left(z_m^+ + z_m^- \right)_{\text{meas}} \ . \tag{40}$$

6. BED RESPONSE: DISCUSSION OF THE RESULTS

During the experiments (tests T4 and T5) concerning collecting data of eroded thickness $\left(z_m^+ + z_m^- \right)_{\text{meas}}$, the measurements of depths in the central box of trap (Fig. 6) were carried out each time after completion of wave generation process in the wave flume. Hence, the bathymetry map was completed.

The averaged bathymetry results in the longitudinal section of the central box of the trap, for T4 and T5 tests, are presented in Fig. 10. The eroded thicknesses $\left(z_m^+ + z_m^- \right)^{T4}$ and $\left(z_m^+ + z_m^- \right)^{T5}$ corresponding to them were estimated in a way that the erosion areas, determined by the bathymetry T4 and T5, are equal to rectangle ones, $\left(z_m^+ + z_m^- \right)^{T4} L_{ST}$ and $\left(z_m^+ + z_m^- \right)^{T5} L_{ST}$, respectively.
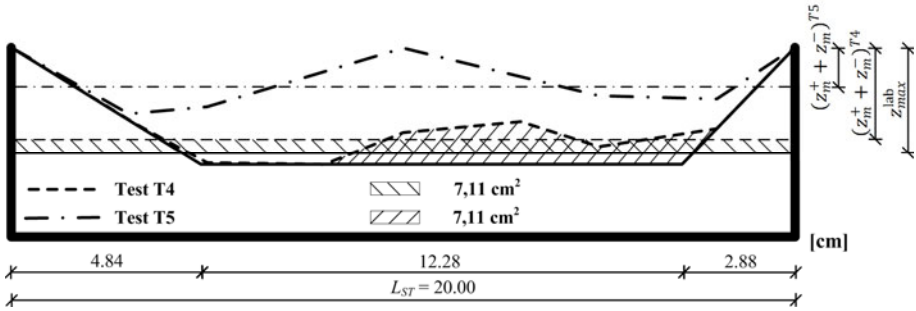
Fig. 10. Bathymetry results in the longitudinal section of the trap for T4 and T5 tests; trapezoid envelope (solid line), T4 bathymetry (dashed line), and T5 bathymetry (dash-dot line).

Additionally, Fig. 10 shows the value of z_{max}^{lab}, which was estimated from the experiments. This value was established whilst the rectangle area, expressed by the relation of $L_{ST} z_{max}^{lab}$, is equal to the trapezoid area with inclined sides, as shown in Fig. 10 and a height for which the area expressed by the relation of $L_{ST}[z_{max}^{lab} - (z_m^+ + z_m^-)^{T4}]$ is approximately equal to the one between the trapezoid envelope and bathymetry of T4. Therefore, it is postulated that while the thickness $(z_m^+ + z_m^-)^{T4}$ reaches the value of z_{max}^{lab}, the ultimate bathymetry coincides with the trapezoid envelope. It is worth noticing that the value $z_{max}^{lab} = 2.72 \, cm$ is close to $z_{max}^{field} = 2.75 \, cm$, assessed in the experiments conducted on the Atlantic updrift coast with the use of fluorescent tracers (Balouin *et al.* 2005) and close to $z_{max}^{field} = 2 - 3 \, cm$ assessed in the former experiments carried out in the Baltic coastal zone by Pruszak *et al.* (1996).

The trapezoid envelope can be also deduced from Eq. 3. This relation shows that the amount of erosion of the bed, specified by the relationship of $q_x dt$, is equal to the triangle area with the side of dx and height of $2z_m$. This triangle area is equal to the rectangle one with constant thickness erosion z_m over the entire length dx. The situation described above occurs when $dx \le (dx)_{lim}$. Boundary length $(dx)_{lim}$ does not depend on the transport rate and it is related to the size of z_{max} in the following way:

$$(dx)_{lim} = \frac{4 z_{max}}{tg\varphi} \, , \tag{41}$$

where $tg\varphi$ is the underwater slope, which must be determined from an experiment. The experiments presented above enable to assess the incline angle of the underwater slope as 38°.

When $dx \geq (dx)_{\text{lim}}$ and when $z_m = z_{\text{max}}$, it is expected that the erosion area of the bed takes the form of a trapezoid, whose area is equal to the calculated one with a constant thickness erosion $z_m = z_{\text{max}}$, along the entire length dx. When $z_m < z_{\text{max}}$, the erosion area boils down to this part of the trapezoid, which is the difference between its area and a part of it where the sediments still remain. Of course, the erosion area is equal to the rectangle one with the dimensions of dx and z_m.

7. CONCLUSIONS

The paper presents the results of laboratory measurements of the thickness of eroded bottom sediment ("IBW PAN 2010"). Within the frame of the experiment, the two tests concerning the measurement of eroded thickness have been carried out for the sand collected from the south Baltic coast in Poland. The sand originally placed in the sand trap was eroded to the outside of the trap under the influence of the oscillating movement of water. After completion of the undulations, the bathymetry data was completed in the sand trap. After that, it was possible to determine the thickness z_m of the sand trap layer from which sediments had been eroded.

The laboratory experiments were carried out under hydrodynamic equilibrium conditions, when the stream of sediments picked up from the bed with length dx, averaged in time dt, equals (at each level) the stream of sediments falling onto the bed. Because the bottom outside the sand trap was made of concrete, it is not difficult to imagine that all the sediments moving as sediment transport under hydrodynamic equilibrium conditions originates from that bed structure.

The measurements of a bed layer thickness where the densely packed grains participate in the apparent rectilinear and translational motion were compared with calculations and a good accuracy was achieved. Further, the results of thickness measurements are used to confirm the hypothesis of linear relation between sediment transport rate and the thickness z_m of a bed layer, participating in the apparent motion. Under the hydrodynamic equilibrium conditions, the apparent bed motion is caused by the actual flow of sediments over the bed in the form of jumps in saltation motion.

The linear form of the relationship between sediment transport rates and the thickness z_m of the layer in apparent flow was also documented by the results of experiments conducted in the laboratory IBW PAN in 1996. It is worth noticing, however, that the transport rate still remains a nonlinear function of the elevation of the bed level.

This linear relationship allows to use the first order "upwind" numerical scheme of FDM in the bathymetry changes calculations and, thus, it provides an accurate solution. However, in order to solve the equation, describing the changes in bathymetry in time and space, properly it is necessary to carry out

the decomposition of the sediment transport into the transport in the positive direction – onshore and negative direction – offshore. Onshore sediment transport takes place during the wave crest and offshore transport during the wave trough duration.

As a result of bed response to sediment transport, a bed erosion area is created. It is assumed that bed erosion (determined by the thickness z_m) is constant at the entire length of this area.

The laboratory experiments indicate, however, that when the thickness z_m reaches the value of z_{max}, the bathymetry of the erosion area coincides with the trapezoid envelope, whose area is equal to the one of the rectangle with the thickness of z_{max}.

When $z_m < z_{max}$ the area of erosion reduces to this part of the trapezoid which is the difference between its area and that part of it where the sediments still remain. The value of z_{max} assessed in the laboratory experiments is close to values obtained in the field experiments conducted on the Atlantic and the Baltic coasts.

References

Balouin, Y., H. Howa, R. Pedreros, and D. Michel (2005), Longshore sediment movements from tracers and models, Praia de Faro, South Portugal, *J. Coast. Res.* **21**, 1, 146-156, DOI: 10.2112/01066.1.

Bialik, R.J. (2013), Numerical study of near-bed turbulence structures influence on the initiation of saltating grains movement, *J. Hydrol. Hydromech.* **61**, 3, 202-207, DOI: 10.2478/johh-2013-0026.

Bialik, R.J., and W. Czernuszenko (2013), On the numerical analysis of bed-load transport of saltating grains, *Int. J. Sediment Res.* **28**, 3, 413-420, DOI: 10.1016/S1001-6279(13)60051-7.

Bialik, R.J., V.I. Nikora, and P.M. Rowiński (2012), 3D Lagrangian modelling of saltating particles diffusion in turbulent water flow, *Acta Geophys.* **60**, 6, 1639-1660, DOI: 10.2478/s11600-012-0003-2.

Callaghan, D.P., F. Saint-Cast, P. Nielsen, and T.E. Baldock (2006), Numerical solutions of the sediment conservation law; a review and improved formulation for coastal morphological modelling, *Coast. Eng.* **53**, 7, 557-571, DOI: 10.1016/j.coastaleng.2006.03.001.

Cayocca, F. (2001), Long-term morphological modeling of a tidal inlet: the Arcachon Basin, France, *Coast. Eng.* **42**, 2, 115-142, DOI: 10.1016/s0378-3839(00)00053-3.

Chiang, Y.-Ch., and S.-S. Hsiao (2011), Coastal morphological modeling. **In:** A.J. Manning (ed.), *Sediment Transport in Aquatic Environments*, InTech, Shanghai, 203-230, DOI: 10.5772/22076.

Chiang, Y.-Ch., S.-S. Hsiao, and M.-C. Lin (2010), Numerical solutions of coastal morphodynamic evolution for complex topography, *J. Marine Sci. Tech.* **18**, 3, 333-344.

Einstein, H.A. (1950), *The Bed-load Function for Sediment Transportation in Open Channel Flows*, Technical Bulletin, No. 1026, United States Department of Agriculture, Washington D.C., 71 pp.

Johnson, H.K., and J.A. Zyserman (2002), Controlling spatial oscillations in bed level update schemes, *Coast. Eng.* **46**, 2, 109-126, DOI: 10.1016/s0378-3839(02)00054-6.

Kaczmarek, L.M. (1999), *Moveable Sea Bed Boundary Layer and Mechanics of Sediment Transport*, IBW PAN, Gdańsk.

Kaczmarek, L.M., J. Biegowski, and R. Ostrowski (2004), Modelling cross-shore intensive sand transport and changes of bed grain size distribution versus field data, *Coast. Eng.* **51**, 5-6, 501-529, DOI: 10.1016/j.coastaleng.2004.05.007.

Kaczmarek, L.M., Sz. Sawczyński, and J. Biegowski (2011), Bathymetry changes and sand sorting during silting up of the channels: Part 1 – Conservation of sediment mass, *Tech. Sci.* **14**, 1, 153-170.

Kuroiwa, M., J.W. Kamphuis, T. Kuchiishi, and Y. Matsubara (2004), A 3D morphodynamic model with shoreline change based on quasi-3D nearshore current model. **In:** *Proc. 2nd Int. Conf. "Asian and Pacific Coasts 2003", 29 February – 4 March 2004, Makuhari, Japan.*

Long, W., J.T. Kirby, and Z. Shao (2008), A numerical scheme for morphological bed level calculations, *Coast. Eng.* **55**, 2, 167-180, DOI: 10.1016/j.coastaleng.2007.09.009.

Moreno, P.A., and F.A. Bombardelli (2012), 3D numerical simulation of particle-particle collisions in saltation mode near stream beds, *Acta Geophys.* **60**, 6, 1661-1688, DOI: 10.2478/s11600-012-0077-x.

Plumb, R.A. (1979), Eddy fluxes of conserved quantities by small-amplitude waves, *J. Atmos. Sci.* **36**, 9, 1699-1704, DOI: 10.1175/1520-0469(1979)036<1699: EFOCQB>2.0.CO;2.

Pruszak, Z., R. Wierzchnicki, A. Owczarczyk, and R.B. Zeidler (1996), Nearbed sediment concentration from tracer studies, *Coast. Eng. Proc.* **25**, 3901-3912, DOI: 10.9753/icce.v25.%25p.

Sawczyński, Sz. (2012), Bathymetry changes and sediment sorting within coastal structures: a case of the silting-up of navigation channels, Ph.D. Thesis, University of Technology, Koszalin, Poland (in Polish).

Sawczyński, Sz., L.M. Kaczmarek, and J. Biegowski (2011), Bathymetry changes and sand sorting during silting up of the channels: Part 2 – Modelling versus laboratory data, *Tech. Sci.* **14**, 2, 171-192.

Struiksma, N., K.W. Olesen, C. Flokstra, and H.J. de Vriend (1985), Bed deformation in curved alluvial channels, *J. Hydraul. Res*. **23**, 1, 57-79, DOI: 10.1080/00221688509499377.

Szymkiewicz, R. (1999), Similarity of kinematic and diffusive waves: a comment on accuracy criteria for linearised diffusion wave flood routing: By K. Bajracharya and D.A. Barry [Journal of Hydrology, vol. 195 (1997), 200-217], *J. Hydrol.* **216**, 3-4, 248-251, DOI: 10.1016/s0022-1694(98)00276-5.

van Rijn, L.C. (1984), Sediment transport, part I: Bed load transport, *J. Hydraul. Eng.* **110**, 10, 1431-1456, DOI: 10.1061/(ASCE)0733-9429(1984)110:10 (1431).

van Rijn, L.C., R.L. Soulsby, P. Hoekstra, and A.G. Davies (eds.), (2005), *SANDPIT, Sand Transport and Morphology of Offshore Mining Pits*, Aqua Publications, The Netherlands.

Walstra, D.J.R., T. Chesher, A.G. Davies, J. Ribberink, P. Sergent, P. Silva, G. Vittori, R. Walther, and L.C. van Rijn (2005), Intercomparison of the state of the morphological models. **In**: L.C. van Rijn, R.L. Soulsby, P. Hoekstra, and A.G. Davies (eds.), *SANDPIT, Sand Transport and Morphology of Offshore Mining Pits,* Aqua Publications, The Netherlands, AY1–AY23.

Watanabe, A. (1988), Modeling of sediment and beach evolution. **In:** K. Horikawa (ed.), *Nearshore Dynamics and Coastal Processes*, University of Tokyo Press, Tokyo, 292-302.

Wiberg, P.L., and J.D. Smith (1985), A theoretical model for saltating grains in water, *J. Geophys. Res.* **90**, C4, 7341-7354, DOI: 10.1029/JC090iC04p07341.

Yalin, M.S., and A.M.F. da Silva (2001), *Fluvial Processes*, IAHR Monograph, International Association for Hydraulic Research, Delft.

Scales of Turbulent Eddies
in a Compound Channel

Adam KOZIOŁ

Faculty of Civil and Environmental Engineering,
Warsaw University of Life Sciences, Warszawa, Poland
e-mails: adam_koziol@sggw.pl

A b s t r a c t

Experimental research was undertaken to investigate the changes in scales of turbulent eddies (macro- and microeddies) in a compound channel and the influence of rigid, emergent floodplain vegetation on scales of turbulent eddies. The results of eight tests for different roughness conditions (smooth bed, rough bed) and with a tree system on the floodplains from two earlier studies are presented. The increase of the channel roughness resulted in a decrease of longitudinal sizes of macroeddies in the whole channel. Trees on the floodplains resulted in disintegration of the sizes of macroeddies, making values of sizes more uniform. A more significant decreasing influence on sizes of macroeddies in the whole channel was exerted by an increase of the main channel sloping bank roughness, having a higher effect than a twofold decrease in the floodplain trees density. The microeddies' sizes are larger in the main channel centreline than on the floodplains and the smallest ones were present in the main channel/floodplain interface.

Key words: compound channel, eddies, emergent vegetation, floodplains.

1. INTRODUCTION

Natural rivers during flood events are often characterized by a compound cross-section and unsteady flow. In these channels, flood conditions lead to

a complex, 3D flow situation with intensive mass and momentum exchange between the deep main channel and the shallow floodplains. This produces a transverse shear layer influencing the flow in both the main channel and the floodplains. The momentum transfer takes place not only by the bed generated turbulence but also by free shear turbulence and velocity fluctuations associated with perturbations in the secondary currents (Czernuszenko *et al.* 2007). The shape of the cross section varies and the roughness of the main channel and the floodplains is often different and the flow structure of a compound channel is usually very complex. The momentum transfer between the main channel and the floodplain generally enhances the discharge on the floodplain, decreasing it in the main channel and thus decreasing the total discharge capacity. The proper description of this so-called "interaction mechanism" is crucial for a reliable prediction of the flow field and related processes, such as flooding, spreading of pollutants, and transport of solids due to sedimentation and erosion. The structure of turbulence in such flows is extremely complex and many experimental works have been performed to investigate it (*e.g.*, Knight and Shiono 1990, 1996, Knight *et al.* 1994, Shiono and Knight 1990, 1991, Tominaga *et al.* 1989). The presence of emergent riparian vegetation on river floodplains makes hydraulic processes even more complex. Vegetation in a compound channel results in an increase in a flow resistance as well changes of velocity distributions, water depth, and turbulent flow characteristics. Research on flow conditions in channels with flexible and rigid vegetation has been performed for a wide range of vegetation density changes: from a dense vegetation distribution to individual plants in a cross-section for both the submerged and emergent vegetation (Ben Meftah *et al.* 2006, Rowiński and Mazurczyk 2006, Yang *et al.* 2007, Nepf and Ghisalberti 2008, Sanjou *et al.* 2010, Terrier *et al.* 2010, Siniscalchi *et al.* 2012, Kubrak *et al.* 2013, Kozioł 2013).

The turbulence generation produces fluctuations of the flow velocity associated with big vortices, and the turbulent energy is transferred in an energy cascade to smaller-scale eddies until it is dissipated into heat by the molecular viscosity (Nezu and Nakagawa 1993). The largest eddies (macroeddies) are impermanent and disintegrate into structures of smaller sizes (microeddies), but simultaneously new, large structures are generated. As a result, the whole and continuous range of sizes of eddies exists in the flow. The scales of eddies in flows are crucial for determining sediment transport and deposition, bed formation, and other processes in rivers. Therefore, the turbulence structure in vegetated river sections has been in the focus of many studies (*e.g.*, Nepf and Ghisalberti 2008).

This paper presents the analysis of an influence of channel roughness and floodplain trees on the longitudinal sizes of the macro- and microeddies, calculated based on the measurements of instantaneous, longitudinal water

velocities in a compound channel from two previous experimental investigations. The main aim of this paper is an analysis of the longitudinal sizes of macro- and microeddies distributions in a channel of a compound cross-section, without and with the impact of floodplain trees.

2. SCALES OF TURBULENT EDDIES

The external scale of turbulence is determined by the sizes of macroeddies. Evaluation of sizes of macroeddies must be preceded by the determination of time-scales of macroeddies. To this end, autocorrelation functions $R(t)$ can be used for this evaluation. The functions exhibit very similar forms of decaying curves with an alteration of the domains of the positive and negative values. Basing on autocorrelation functions, Euler time-macroscales T_E are derived according to Nikora *et al*. (1994):

$$T_E = \int_0^\infty R(t)dt \ , \tag{1}$$

which are a measure of the slowest changes in the turbulent flow caused by macroeddies. For a steady and uniform turbulent flow, when mean velocity in a given point significantly exceeds the velocity of fluctuations, there exists, according to the Taylor's hypothesis, a direct relationship of temporal and spatial Eulerian autocorrelation functions. Referring to Taylor's relationships between the spatial L and temporal T_E turbulence macroscale, the following formula for mean longitudinal sizes can be derived:

$$L = UT_E \ , \tag{2}$$

where U is the time-averaged point velocity. Most often the sizes of macroeddies are related to the water depths h in an analysed measurement vertical for easier comparison.

The size of microeddies is a measure for the smallest size of eddies which are present in the turbulent flow of water. Kolmogoroff and Taylor proposed the spatial scale of microeddies (Nezu and Nakagawa 1993, Mazurczyk 2007). The Taylor's microscale depends both on the macroscopic motion by means of the fluctuating velocity and on dissipative characteristics, whereas the Kolmogoroff's microscale depends only on dissipative and viscous characteristics (Nezu and Nakagawa 1993). In order to identify the size of the Kolmogoroff's microscale η (microeddies), the spectrum density function $S(\omega)$ is calculated for instantaneous, longitudinal velocities and next the proper frequency subrange of velocities is determined, and for that subrange the rate of energy dissipation ε can be obtained (Kozioł 2012). The spectrum density function $S(\omega)$ expresses kinetic energy of eddies for the frequency range $(\omega, \omega + d\omega)$ and this function is not uniform *versus* fre-

quency. This function takes the largest values for the frequencies of the averaged motion, while the lowest values are reached for the largest frequencies. The analysis of the spectrum density function involves, among others, the determination of the frequency subrange (inertial subrange), whose existence is assumed by the Kolmogoroff's hypothesis, and where the local anisotropy hypothesis is also valid as well as the "–5/3" power law (Nikora 1999). Such a subrange is called the inertial subrange of a constant energy stream, coming from all scales of eddies: from the largest eddies to the smallest ones. The determination of the inertial subranges enables to find turbulence scales and energy dissipation rates. The aforementioned values can be calculated by applying following formulas (Nikora *et al.* 1994):

$$S(\omega) = C_1 \varepsilon^{2/3} U^{2/3} (\omega)^{-5/3} \ , \tag{3}$$

$$\eta = \left(\frac{v^3}{\varepsilon} \right)^{1/4} \ , \tag{4}$$

where $C_1 = 0.48$ is a constant (Monin and Yaglom 1975), η is the size of microeddies (Kolmogoroff's microscale), and v is the molecular (kinematic) viscosity. Equations 3 and 4 are valid on the condition that the Taylor hypothesis of "frozen turbulence" is applied.

3.   EXPERIMENTAL SETUP AND PROCEDURE

In this paper, an analysis of sizes of macro- and microeddies distributions in the channel, including all eight tests from two studies without and with the impact of floodplain trees, is presented (Fig. 1). Three measurements were conducted with the use of an electromagnetic meter (tests 1-3, Fig. 1a-c), and five measurements were continued with the use of the 10-MHz acoustic Doppler velocity meter (ADV) (tests 4-5.1, Fig. 1d-e). Results of the first study (tests 1-3) were reported by Kozioł (2000) and Rowiński *et al.* (2002), while partial results of the second study were reported by Kozioł (2008, 2012) (tests 4-5), Czernuszenko *et al.* (2007) (test 4), and Mazurczyk (2007) (test 5).

The experiments considered herein were carried out in the Hydraulic Laboratory of the Department of Hydraulic Engineering, Faculty of Civil and Environmental Engineering at the Warsaw University of Life Sciences – SGGW. A straight open channel (16 m long and 2.10 m wide) with a symmetrically trapezoidal cross-section was used for the laboratory tests. The main channel width was 30 cm with a floodplain width of 60 cm. The banks were inclined at a slope of 1:1. The channel bed slope of the channel was 0.5‰. Water discharge values were recorded with the use of the 540 mm diameter circular weir. A row of PCV pipes was installed in the upstream ap-
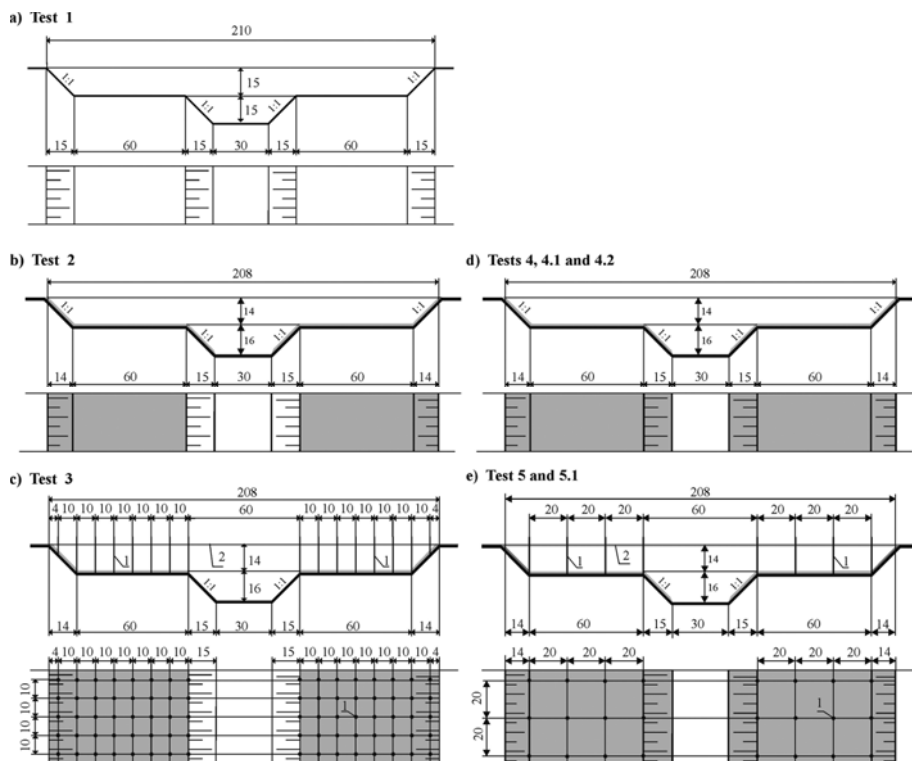
Fig. 1. Scheme of a laboratory cross-section for two considered studies. The first study: (a) test 1 in a channel with the smooth bed, (b) test 2 in a channel with the smooth bed of the main channel and rough floodplains, and (c) test 3 in a channel with the smooth bed of the main channel and rough floodplains vegetated with trees (0.1 × 0.1 m trees system). The second study: (d) tests 4, 4.1, and 4.2 in a channel with the smooth bed of the main channel and rough sloping banks of the main channel, rough floodplains, and (e) tests 5 and 5.1 in a channel with the smooth bed of the main channel and rough sloping banks of the main channel, rough floodplains vegetated with trees (0.2 × 0.2 m trees system). Explanations: 1 – pipes imitating trees, and 2 – wooden strips supporting the trees (dimensions in cm).

proach channel reach to straighten the flow. During the experiments the following parameters were measured: water levels in the main channel and on the floodplains, velocities, water temperature, and water discharge. Before general measurements were started, some trial velocity measurements had been performed in a few cross-sections at distances of 4 and 12 m from the channel entrance. These tests showed that the Reynolds number was sufficiently large, in order to create the state characterized by local isotropy and homogeneity and associated universal behaviour of statistical properties

(*e.g.*, velocity, standard deviation, and skewness). The cross-section half way down of the channel length was selected for velocity measurements. Instantaneous longitudinal velocities were measured with two probes.

In the first study, in tests 1, 2, and 3, instantaneous longitudinal velocities were measured with the use of the programmable electromagnetic liquid velocity meter manufactured by Delft Hydraulics. The accuracy of the measurement of velocity is ±0.01 m/s. The measurements of instantaneous longitudinal velocities were carried out at 48 points at 10 verticals (verticals 6-36 and 39, Fig. 2). The time of velocity recording at each point was roughly 40 min with a sampling interval of 5 Hz. The analysis of the velocity data showed that a sampling period slightly shorter than 27 min was enough to obtain fully steady statistical characteristics of the flow and, therefore, the time series of 8000 elements (corresponding to the time period of about 26.7 min) were further elaborated. More detailed information about the methodology can be found in Rowiński *et al*. (2002).

In the second study, in tests 4, 4.1, 4.2, 5, and 5.1, instantaneous longitudinal velocities were measured with the use of a three-component acoustic Doppler velocity meter (ADV) manufactured by Sontek Inc. The ADV proved to yield a good description of the turbulence characteristics when certain conditions related to the flow itself and the configuration of the instrument were satisfied. The measurements were conducted with the maximum frequency of 25 Hz in the velocity range of 0 to 1.0 m/s with an accuracy of 0.25 cm/s. The measurements of instantaneous velocities were carried out at 250 points at 23 verticals – six on each floodplain and eleven in the main channel (Fig. 2). The time of velocity recorded in several dozen points was roughly 20 min; it turned out that a sampling time, slightly shorter than 6 min, is enough to obtain fully steady statistical characteristics of the flow and, therefore, a time series of 9000 elements (corresponding to the time period of about 360 s) has been used subsequently. More detailed information about the methodology can be found in the paper by Czernuszenko *et al*. (2007). Before the second study measurements were started, some trial velocity measurements had been performed by two probes in a few points in
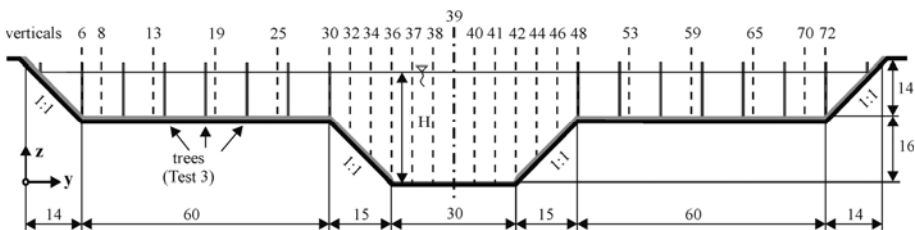
Fig. 2. Scheme of the experimental cross-section with the arrangement of measuring verticals (dimensions in cm).

the main channel and on the floodplains. The calculated values of velocity, standard deviations, skewness, and kurtosis coefficients were compared for two probes, and the comparison confirmed that the values obtained from the two instruments were similar.

In the first study, in test 1, the surface of the channel bed was smooth (the smooth channel) and made of concrete. In tests 2 and 3 the floodplains were covered by cement mortar composed of terrazzo with grains of diameter within the range from 0.5 to 1 cm. The values of an average Manning coefficient and absolute roughness of the channel surface were determined from the Manning's equation and the Colebrook–White's equation on the basis of the average velocity values of the flow measured in the main channel and on both the floodplains. For the smooth channel, Manning's roughness coefficient was equal to about $n = 0.011$ m$^{-1/3}$s, $n = 0.018$ m$^{-1/3}$s for the left rough surface of the floodplain, and 0.025 m$^{-1/3}$s for the right surface. The obtained roughness amounted to $k_s = 0.00005$ m for the smooth surfaces, $k_s = 0.0074$ m for the rough surface of the left floodplain, and $k_s = 0.0124$ m for the rough surface of the right floodplain. In test 3, rigid emergent stems of riparian trees growing on the floodplains were modelled by aluminium pipes of 0.8 cm diameter, placed with both longitudinal and lateral spacings of 10 cm (100 stems per m^2, Fig. 1c). The type of arrangement of stems is illustrated in Fig. 1. The stems simulate basically only the rigid stem of trees. There were 16 pipes in each of 161 cross-sections.

In the second study (tests 4, 4.1, 4.2, 5, and 5.1), both banks of the main channel were covered by cement mortar composed with terrazzo with grains of diameter within the range 0.5 to 1 cm (Fig. 1d-e). In tests 5 and 5.1 rigid emergent stems of riparian trees growing on the floodplains were modelled

Table 1

The hydraulic parameters

Parameter	Test							
	1	2	3	4	4.1	4.2	5	5.1
Discharge Q [l/s]	69.3	50.0	38.9	95.2	81.1	61.5	65.7	58.9
Depth in the main channel H [cm]	20.2	21.9	21.2	28.3	26.4	24.1	28.0	26.3
Depth on the floodplain h_f [cm]	5.2	5.9	5.2	12.3	10.4	8.1	12.0	10.3
Relative depth $Dr = h_f/H$	0.26	0.27	0.24	0.43	0.39	0.34	0.43	0.39
Vegetation on the floodplains	–	–	Trees	–	–	–	Trees	
The sloping banks [–]	1:1							
The bed slope i [‰]	0.5							

by aluminium pipes of 0.8 cm diameter, placed with both longitudinal and lateral spacings of 20 cm (25 stems per m^2, Fig. 1e). There were 8 stems in tests 5 and 5.1 in 80 cross-sections. The hydraulic parameters of two studies are given in Table 1.

4.   RESULTS AND ANALYSIS

4.1 Temporal scale of macroeddies in the stream

The ranges of the calculated values of time macroscales on the floodplains and in the main channel for all analysed tests are given in Table 2. The calculated Euler time-macroscale values ranged from 0.08 to 0.95 s on the floodplains without trees, while in the main channel they varied from 0.08 to 1.97 s (Table 2). On the other hand, in scenarios with floodplain trees, the time-macroscale values ranged from 0.1 to 0.98 s on the floodplains, while in the main channel they ranged from 0.1 to 0.82 s (Table 2). In the main channel the largest values of time-macroscale were reached for a smooth bed (test 1), and the lowest were found for tests 5 and 5.1. Introducing rigid, emergent vegetation causes generation of wakes behind each stem; therefore, trees generated a decrease of time-macroscale in test 5, in comparison with test 4, both for the main channel and the floodplains.

Table 2

The calculated values of the macro time-scale of turbulence T_E [s]

Area	Test 1	Test 2	Test 3 (trees)	Test 4
The left floodplains	0.15-0.62	0.25-0.95	0.30-0.98	0.14-0.59
The main channel	0.10-1.97	0.23-0.76	0.50-0.82	0.15-0.96
The right floodplains	–	–	–	0.13-0.62
Area	Test 4.1	Test 4.2	Test 5 (trees)	Test 5.1 (trees)
The left floodplains	0.09-0.74	0.15-0.68	0.10-0.51	0.14-0.38
The main channel	0.08-0.63	0.07-0.64	0.10-0.46	0.10-0.50
The right floodplains	0.08-0.65	0.08-0.58	0.14-0.42	0.18-0.38

4.2 Spatial longitudinal scales of macroeddies in the stream

The calculated mean longitudinal sizes of macroeddies and relative sizes of macroeddies (L/h) related to the water depths h at location of y (Fig. 2) are presented in Table 3.

Figure 3 illustrates the contour lines for the longitudinal relative sizes of macroeddies (L/h). The results show that for all tests with rough floodplains, the relative sizes of macroeddies are larger on the floodplains than in the main channel. The largest sizes appear on the floodplains in the interaction

The calculated longitudinal sizes of macroeddies L (spatial macro-scale)

Area	Test 1		Test 2		Test 3 (trees)		Test 4	
	L [cm]	L/h	L [cm]	L/h	L [cm]	L/h	L [cm]	L/h
The left floodplains	6-24	1.1-4.2	3-17	0.5-2.9	2-11	0.4-2.1	4-19	0.3-1.6
The main channel	4-89	0.2-4.4	8-30	0.4-1.6	13-25	0.6-1.7	3-26	0.1-1.2
The right flood plains	–	–	–	–	–	–	1-23	0.1-1.2

Area	Test 4.1		Test 4.2		Test 5 (trees)		Test 5.1 (trees)	
	L [cm]	L/h	L [cm]	L/h	L [cm]	L/h	L [cm]	L/h
The left floodplains	2-24	0.2-2.3	4-20	0.5-2.5	2-9	0.2-0.8	1-8	0.1-0.8
The main channel	2-21	0.1-1.2	2-17	0.1-1.0	1-13	0.2-0.7	1-15	0.1-0.8
The right floodplains	2-24	0.1-2.3	2-13	0.2-1.6	2-9	0.2-0.8	1-8	0.1-0.7

zone, close to the main channel, and the smallest ones occur over the main channel bed.

Figure 4 illustrates the changes of relative sizes of macroeddies (L/h) as a function of the relative depth z/h (z is the distance from the bed), at selected measurement verticals on the floodplains and in the main channel. The distributions of longitudinal sizes of macroeddies in verticals are not similar within the whole channel cross-section.

Macroeddies in the main channel

For the smooth channel (test 1) the largest relative sizes of macroeddies were achieved for the main channel centreline (V39, Fig. 4), and the smallest size was reached in test 4, with the rough surface of main channel sloping banks and floodplains (Fig. 4) (Kozioł 2000). In the main channel in test 1, when for a relative depth of $z/h = 0.5$, the largest relative sizes of macroeddies were equal to 4.4 times the stream depth (V39, Fig. 4). The increase of the roughness of the floodplains (test 2) and the main channel sloping banks (tests 4-4.2) resulted in a decrease of water velocity and also reduced the sizes of macroeddies. The increase of roughness on the floodplains in test 2 resulted in the decrease in sizes of macroeddies at the bed and at the water surface only, but in the middle depth zone ($0.25 < z/h < 0.75$, V39, Fig. 4) it almost did not change and amounted to about one time stream depth. However, the increase of main channel sloping banks roughness in tests 4-4.2 resulted in the decrease in sizes of macroeddies in the main channel to 0.5 of the stream depth (V39, Fig. 4). Floodplain trees (tests 3, 5, and 5.1) resulted in the decrease of water velocities in the main channel, which are identical in verticals for three tests (V39, Fig. 4); however, sizes of macroeddies almost did not change in the smooth main channel (tests 2-3, Fig. 3, V39 in Fig. 4),
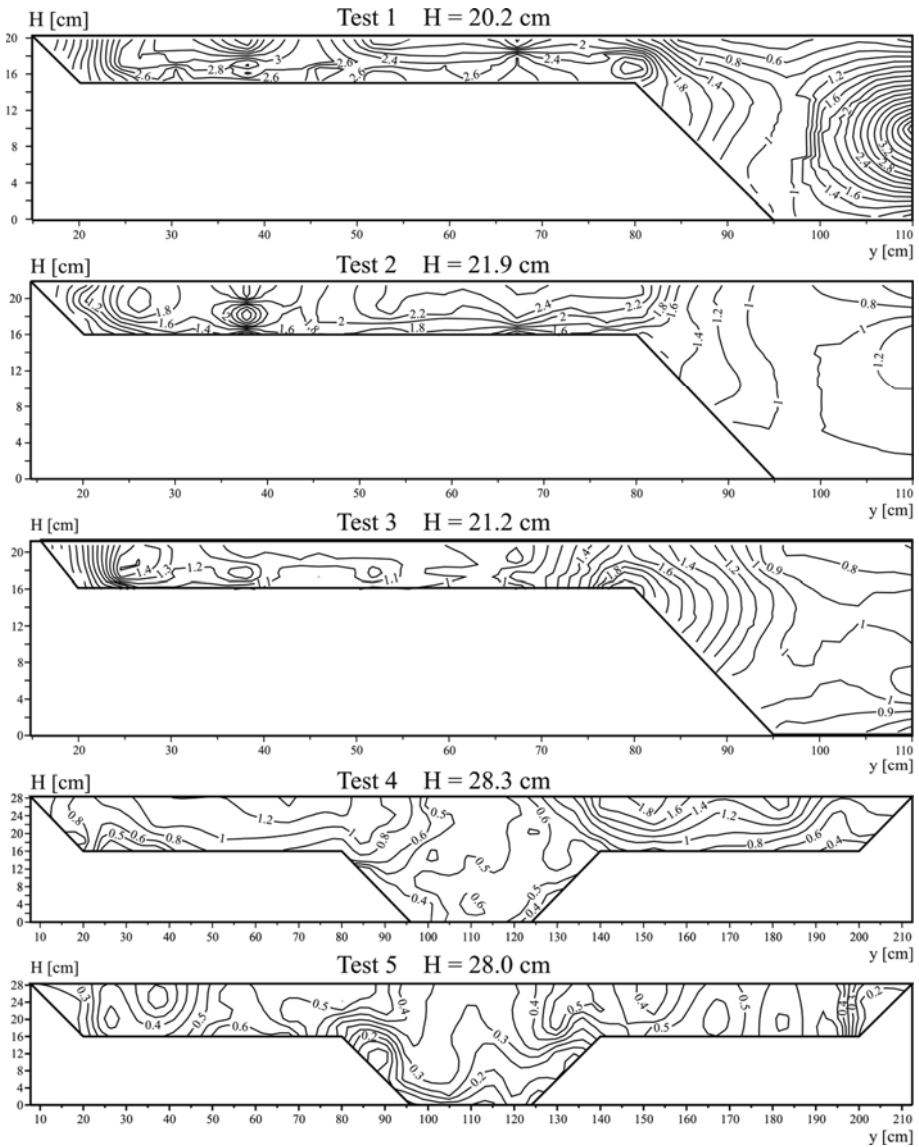
Fig. 3. Contour lines for longitudinal relative sizes of macroeddies (L/h).

but they decreased from $L/h = 0.6$ (test 4) to 0.3 (test 5) in the main channel with rough banks. The comparison of the main channel to rough banks (tests 5 and 5.1) shows that sizes of macroeddies are over three times larger in the smooth main channel (test 3, Fig. 3, V39 in Fig. 4).
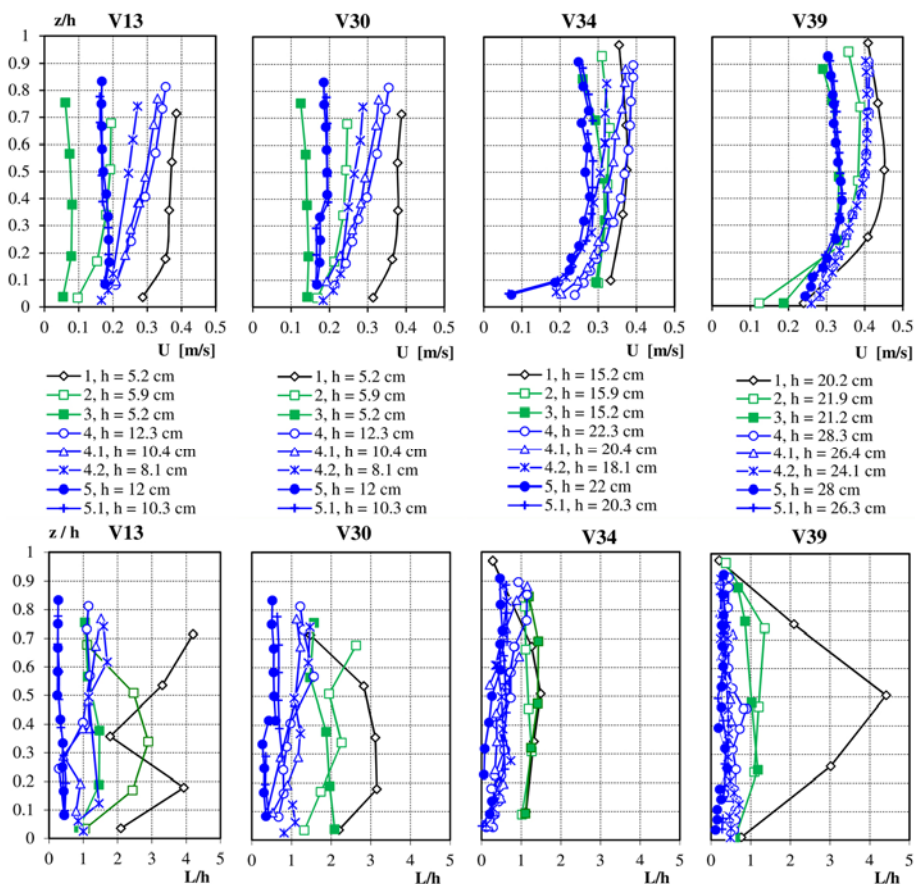
Fig. 4. Example vertical distributions of longitudinal velocities and relative sizes of macroeddies on floodplains.

Macroeddies over the sloping banks of the main channel

The sizes of macroeddies are the largest nearby the floodplain over the sloping banks of the main channel (Fig. 3, V30 in Fig. 4) and the smallest in the vertical located between the sloping bank and the main channel (Fig. 3). In the vertical located between the sloping bank and the bed of the main channel, the largest sizes of macroeddies occur in the smooth channel (test 1) and in the channel with rough floodplains (test 2), and the sizes are of the order of the stream depth (Fig. 3). The smallest sizes of macroeddies appeared in the channel with rough floodplains and rough sloping banks (tests 4-4.2), and they decreased downwards, towards the bed, from 1 to 0.1 times the stream depth (Fig. 3, V34 in Fig. 4). Over the sloping banks of the main channel close to the floodplains, where the greatest interaction between the

main channel and the floodplain exists, the largest sizes of macroeddies appear only in the smooth channel (test 1), and at the bed they reach 2 stream depths (Fig. 3, V30 in Fig. 4). The smallest sizes of macroeddies occur in the case of rough floodplains and rough sloping banks of the main channel (tests 4-4.2), and here the size decreases downwards the bed from 1.2 to 0.1 times the stream depth (Fig. 3, V30 in Fig. 4). In the smooth channel, sizes of macroeddies increase downwards the bed, and in the channel with rough floodplains they remain constant with depth ($L/h = 1.5$ approximately). Floodplain trees (tests 3, 5, and 5.1) resulted in a decrease of water velocity over smooth sloping banks of the main channel only above $z/h = 0.4$, but over the rough sloping banks the velocity decreased in the whole depth (V34, Fig. 4). Sizes of macroeddies are the largest over the sloping banks of the main channel, close to the floodplain in case when the sloping banks are smooth (test 3). The farther from the floodplains, the smaller the impact of trees on sizes of macroeddies over the main channel sloping banks.

Macroeddies on the floodplains

On the floodplains, the largest relative sizes of macroeddies were achieved in the smooth channel, and reached 4 times the stream depth (Table 3, Fig. 4). The increase of the roughness of floodplains and the main channel sloping banks resulted in the lowest sizes of macroeddies in tests 4-4.2, which appear in a vertical near the bed and close to the sloping bank of the floodplain. The largest, relative sizes of macroeddies on the floodplain are observed at the floodplain/main channel interface. On both floodplains, sizes of macroeddies decreased downwards the bed, from 2.5 to 0.1 times the stream depth; however, close to the floodplain sloping banks the macroeddies were almost constant with depth, and are equal to one stream depth.

Trees on the floodplains (tests 3, 5, and 5.1; Fig. 4) resulted in the relative scales of macroeddies' values being more uniform. It is clear that the emergent vegetation generates turbulence along all verticals by creation of a wake behind a row of parallel rods. In other words, some large local gradients of a mean velocity exist behind the "trees", thereby making the relative scales of macroeddies much smaller and the vertical distributions of macroeddies more uniform. In tests 5 and 5.1 on both floodplains, relative sizes of macroeddies are lower than one time the stream depth, ranging from 0.1 to 0.8 of the stream depth. In some verticals, almost behind a tree, relative sizes of macroeddies slightly increase downwards the bed (V13, Fig. 4).

In tests 3, 5, and 5.1, relative sizes of macroeddies were slightly different close to the floodplain sloping bank only, but in other areas of the channel cross-section, the macroeddies were several times larger in test 3 than in other cases. Lower relative sizes of macroeddies in tests 5 and 5.1 were achieved for rough main channel sloping banks and the two times smaller

number of floodplain trees. It can be concluded that a more significant influence on the decreasing macroeddies' size is exerted by an increasing roughness of the main channel than by 2 times smaller density of floodplain trees.

Additional proof for this can be found in research by Nepf and Vivoni (2000), Carollo *et al.* (2005), and Mazurczyk (2007). Carollo *et al.* (2005) found that the increase of an absolute roughness resulted in a decrease in longitudinal macroeddies size. These findings were based on the gravel-bottom channels research, considering a different grain size. However, Nepf and Vivoni (2000) and also Mazurczyk (2007) performed research on channels with emergent vegetation and proved that the increase of vegetation density considerably decreases sizes of macroeddies, so the decrease of the density should lead to an increase of the macroeddies size. A literature review of laboratory and field measurement results (*e.g.*, McQuivey *et al.* 1971, Nikora *et al.* 1994, Nikora and Smart 1997, Mazurczyk 2007) proved that sizes of macroeddies usually ranged from 0.5 to 4 times the stream depth, so the results achieved in this paper are consistent with the research results by other authors.

4.3 Spatial longitudinal scales of microeddies in the stream

The analysis of the spectrum density function enabled us to determine an inertial subrange and the subsequent determination of the energy dissipation rates (Eq. 3), and the sizes of Kolmogorov's microeddies η (Eq. 4) were calculated. Figure 5 presents the frequency spectra $S(f)$ of instantaneous longitudinal velocities in analysed tests. The presence of floodplain trees resulted in the decrease of values of the spectrum density function (*e.g.*, comparing test 2 and 4 or 4 and 5, Fig. 5).

Figure 6 presents vertical distribution of longitudinal velocities and longitudinal sizes of microeddies in tests 1, 2, and 4 for non-vegetated floodplains. The calculated values of microeddies are very small in comparison with the value of velocity (Fig. 6) and with the sizes of macroeddies L which most often are presented in relation to the water depth (Fig. 4). All calculated sizes of microeddies are in the order of decimal parts of a millimetre, and they vary from 0.13 to 0.44 mm. Most of calculated values do not vary significantly in the channel, and they are approximately 0.2 mm, while the increase of sloping banks roughness in the main channel resulted in a growth of microeddies size in the main channel that they reached, from 0.24 to 0.44 mm in length (test 4, V39, Fig. 6b). It was concluded that the increase of the floodplain roughness and main channel sloping banks roughness did not result in vital changes of the microeddies size on the floodplain. However, the increase of floodplain roughness and the increase of the main channel sloping banks roughness resulted in a growth of microeddies size in the main channel (Fig. 6b).
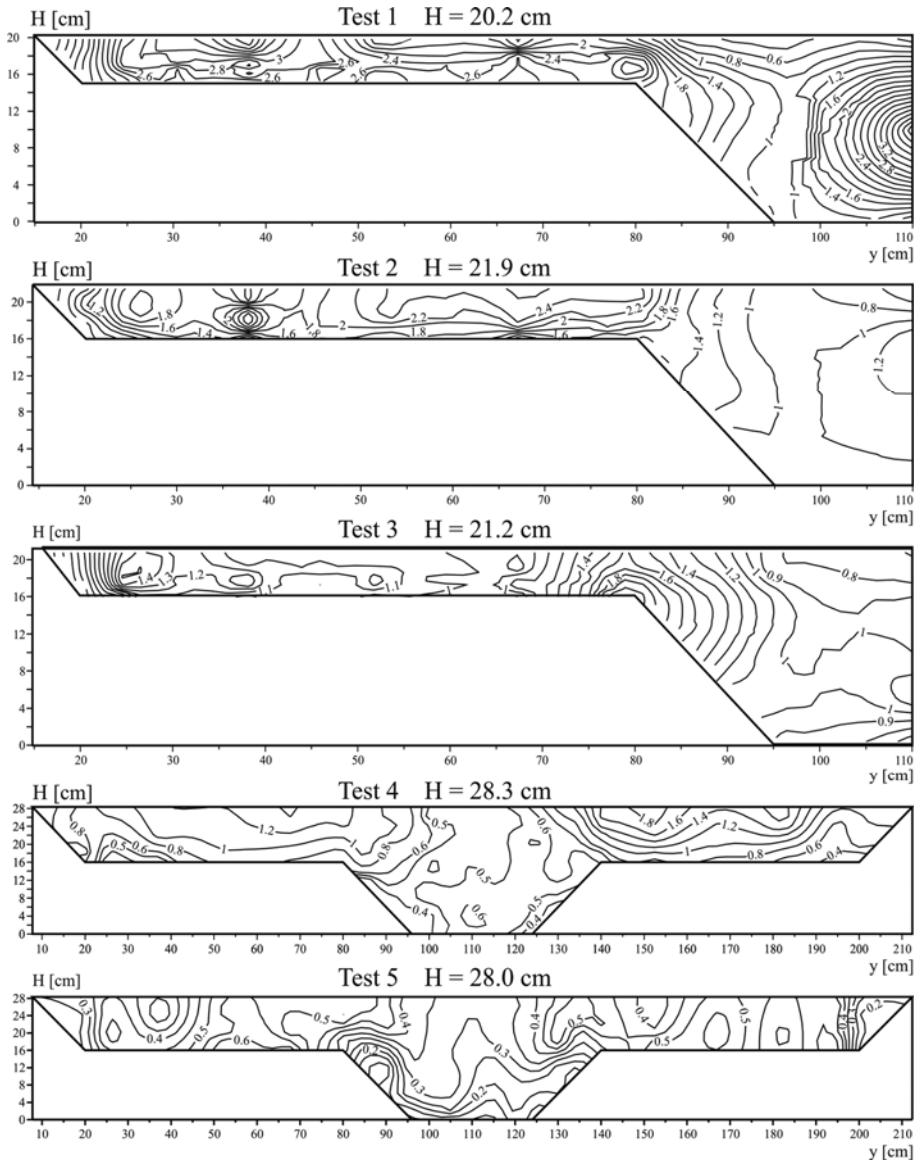
Fig. 5. Frequency spectra *S(f)* of instantaneous longitudinal velocities in tests 1-5 (*z* is the distance from the bed).

Figure 7 presents vertical distributions of longitudinal velocities and sizes of Kolmogoroff's microeddies η in the compound channel with emergent vegetation on the floodplains (tests 3 and 5). Trees did not result in large changes of sizes of microeddies, and the calculated values are also very simi-
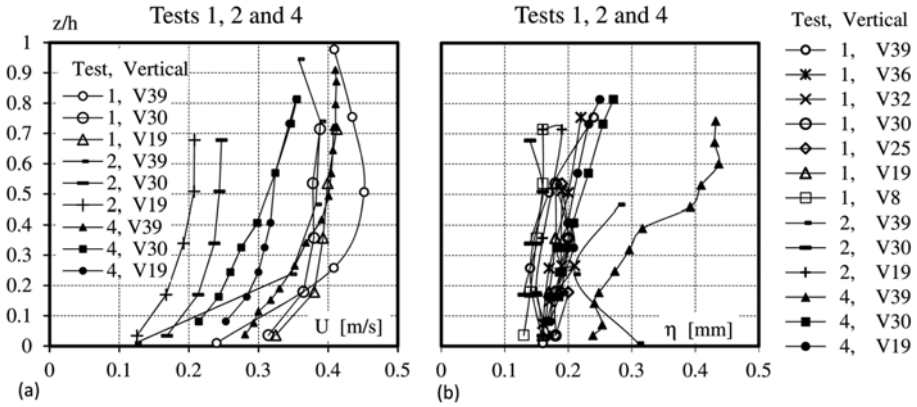
Fig. 6. Vertical distribution of longitudinal velocities and sizes of microeddies in tests 1, 2, and 4.
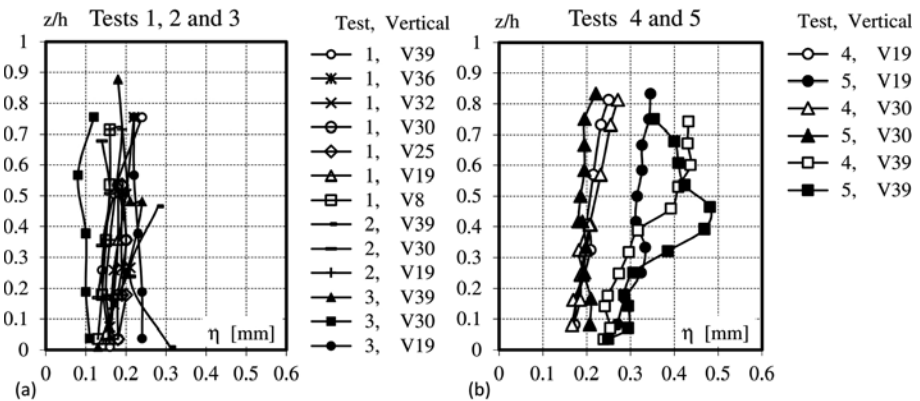
Fig. 7. Distribution of longitudinal sizes of microeddies in tests 1-5.

lar and range from 0.08 to 0.48 mm. In test 3, the sizes of microeddies did not change on the floodplain and in the main channel, but only in the interaction zone, close to the surface that divides the main channel and the floodplain, the eddies disintegrated into structures of smaller sizes (V30, Fig. 7a). In test 5, floodplain trees contributed to the disintegration into structures of bigger sizes on the floodplain, bigger by about 50% in comparison with test 3, but had no influence on the vertical distribution of sizes of microeddies (V19, Fig. 7b). However, in the plane dividing the main channel and the floodplain, the sizes of microeddies did not vary considerably, and in the test with floodplain trees, the sizes of microeddies did not change with depth (V30, Fig. 7b). Therefore, the sizes of microeddies are slightly bigger near the bed and smaller above the relative depth $z/h = 0.4$ in test 5. In the main channel the sizes of microeddies almost did not change in test 5, but flood-

plain trees resulted in changes of the vertical distribution of microeddies size (V39, Fig. 7b). The longest sizes of microeddies are at the relative depth $z/h = 0.45$. The sizes of microeddies are significantly bigger in the middle zone of the main channel ($z/h = 0.3$-0.7) than on the floodplain.

Trees on the floodplains did not lead to considerable changes of sizes of microeddies, and usually their values were very similar. The analysed microscale was within the range from 0.1 to 10 mm. This statement is consistent with the results achieved for free streams and near the wall regions (Czernuszenko and Lebiecki 1989, Nikora *et al.* 1994, Nikora and Smart 1997, Mazurczyk 2007).

5. CONCLUSIONS

Investigations of the longitudinal sizes of macro- and microeddies, presented in the paper, are based on measurements of instantaneous velocities in a hydraulic laboratory in a compound trapezoidal channel with different roughness conditions (smooth bed, rough bed) and with a trees system on the floodplains. The increase of a channel bottom roughness resulted in a considerable decrease in longitudinal sizes of macroeddies in the whole channel. In case of rough floodplains and rough sloping banks of the main channel, sizes of macroeddies become larger on the floodplains than in the main channel; in verticals the size increased upwards the water surface on the floodplains and over the main channel sloping banks, but over the bed of the main channel it was almost constant with depth. In all tests with and without floodplain trees, the largest macroeddies appeared on the floodplain/main channel interface, and the lowest were found over the main channel bed. Floodplain trees resulted in a decrease in water velocity and sizes of macroeddies; the most significant decrease took place on the floodplains, and the less significant concerns the main channel. Trees located in the active flow zone on the floodplains resulted in macroeddies disintegrating and significant decreasing their size, but also contributed to their constant value *versus* depth. However, trees did not influence the values and vertical distribution of macroeddies over the main channel bed. More significant influence on sizes of macroeddies decreasing in the whole channel is exerted by the increase of the main channel sloping banks roughness than the influence of twice decreased density of floodplain trees.

The calculated values of sizes of microeddies varied within the ranges of 0.08 to 0.48 mm. The changes of longitudinal sizes of microeddies on the floodplains were not influenced by the increase of floodplain roughness (test 2), the increase of main channel sloping banks roughness (test 4), or the presence of trees, but only for the smooth main channel (test 3). However, when the roughness of the floodplains and sloping banks of the main channel

were identical (test 5), the occurrence of trees resulted in disintegration of the sizes of macroeddies into bigger sizes of microeddies. The increase of the main channel sloping banks roughness (test 4) and the occurrence of trees (test 5) resulted in an increase of sizes of the microeddies in the main channel. The sizes of the microeddies are greater in the centreline of the main channel than on the floodplains, the smallest being in the main channel/floodplain interface. The calculated longitudinal sizes of microeddies were very small in comparison to the value of velocity (0.05-0.45 m/s) and to the scale of macroeddies which most often are presented in relation to the water depth. The values of the microeddies were in the order of decimal parts of a millimetre, and that is why it is difficult to draw conclusions about significant changes of their values as a result of an influence of bed roughness and floodplain trees.

R e f e r e n c e s

Ben Meftah, M., F. De Serio, D. Malcangio, and A.F. Petrillo (2006), Experimental study of flexible and rigid vegetation in an open channel. **In:** *Proc. Int. Conf. on Fluvial Hydraulics "River Flow 2006", 6-8 September 2006, Lisbon, Portugal*, 603-611.

Carollo, F.G., V. Ferro, and D. Termini (2005), Analyzing turbulence intensity in gravel bed channels, *J. Hydraul. Eng. ASCE* **131**, 12, 1050-1061, DOI: 10.1061/(ASCE)0733-9429(2005)131:12(1050).

Czernuszenko, W., and P. Lebiecki (1989), The turbulence in river-flows, *Arch. Hydrotech*. **36**, 1-2, 17-34 (in Polish).

Czernuszenko, W., A. Kozioł, and P.M. Rowiński (2007), Measurements of 3D turbulence structure in a compound channel, *Arch. Hydro-Eng. Environ. Mech.* **54**, 1, 55-73.

Knight, D.W., and K. Shiono (1990), Turbulence measurements in a shear layer region of a compound channel, *J. Hydraul. Res.* **28**, 2, 175-195, DOI: 10.1080/00221689009499085.

Knight, D.W., and K. Shiono (1996), River channel and floodplain hydraulics. **In:** M.G. Anderson, D.E. Walling, and P.D. Bates (eds.), *Floodplain Processes*, Ch. 5, John Wiley and Sons, Chichester, 139-181.

Knight, D.W., K.W.H. Yuen, and A.A.I. Alhamid (1994), Boundary shear stress distributions in open channel flow. **In:** K. Beven, P. Chatwin, and J. Millbank (eds.), *Physical Mechanisms of Mixing and Transport in the Environment*, John Wiley and Sons, New York, 51-87.

Kozioł, A.P. (2000), Longitudinal sizes of the largest eddies in the compound channel, *Prz. Nauk. Wydz. Inż. Kształt. Środow.* **18**, 151-159 (in Polish).

Kozioł, A.P. (2008), Investigation of the time and spatial macro-scale of turbulence in a compound channel, *Acta Sci. Pol. – Architectura* **7**, 4, 15-23 (in Polish).

Kozioł, A.P. (2012), The Kolmogorov's microscale eddies in a compound channel, *Ann. Warsaw Univ. Life Sci. – SGGW, Land Reclamation* **44**, 2, 121-132, DOI: 10.2478/v10060-011-0068-7.

Kozioł, A.P. (2013), Three-dimensional turbulence intensity in a compound channel, *J. Hydraul. Eng*. **139**, 8, 852-864, DOI: 10.1061/(ASCE)HY.1943-7900. 0000739.

Kubrak, E., J. Kubrak, and P.M. Rowiński (2013), Application of one-dimensional model to calculate water velocity distributions over elastic elements simulating Canadian waterweed plants (Elodea Canadensis), *Acta Geophys*. **61**, 1, 194-210, DOI: 10.2478/s11600-012-0051-7.

Mazurczyk, A. (2007), Scales of turbulence in compound channels with trees on floodplains, *Publs. Inst. Geophys. Pol. Acad. Sc.* **E-7**, 401, 169-176.

McQuivey, R.S., T.N. Keefer, and M.A. Shirazi (1971), Basic data report on the turbulent spread of heat and matter, Open-file Report, USA Dept. of Interior Geological Survey, Fort Collins, USA, 166 pp.

Monin, A.S., and A.M. Yaglom (1975), *Statistical Fluid Mechanics: Mechanics of Turbulence*, Vol. 2, MIT Press, Cambridge.

Nepf, H., and M. Ghisalberti (2008), Flow and transport in channels with submerged vegetation, *Acta Geophys*. **56**, 3, 753-777, DOI: 10.2478/s11600-008-0017-y.

Nepf, H.M., and E.R. Vivoni (2000), Flow structure in depth-limited, vegetated flow, *J. Geophys. Res*. **105**, C12, 28547-28557, DOI: 10.1029/2000JC 900145.

Nezu, I., and H. Nakagawa (1993), *Turbulence in Open-Channel Flows*, IAHR Monographs, Balkema, Rotterdam, 293 pp.

Nikora, V.I. (1999), Origin of the "–1" spectral law in wall-bounded turbulence, *Phys. Rev. Lett*. **83**, 4, 734-736, DOI: 10.1103/PhysRevLett.83.734.

Nikora, V.I., and G.M. Smart (1997), Turbulence characteristics of New Zealand gravel-bed rivers, *J. Hydraul. Eng*. **123**, 9, 764-773, DOI: 10.1061/(ASCE) 0733-9429(1997)123:9(764).

Nikora, V.I., P. Rowiński, A. Sukhodolov, and D. Krasuski (1994), Structure of river turbulence behind warm-water discharge, *J. Hydraul. Eng*. **120**, 2, 191-208, DOI: 10.1061/(ASCE)0733-9429(1994)120:2(191).

Rowiński, P.M., and A. Mazurczyk (2006), Turbulent characteristics of flows through emergent vegetation. **In:** *Proc. Int. Conf. on Fluvial Hydraulics "River Flow 2006", 6-8 September 2006, Lisbon, Portugal*, 623-630.

Rowiński, P.M., W. Czernuszenko, A.P. Kozioł, and J. Kubrak (2002), Properties of a streamwise turbulent flow field in an open two-stage channel, *Arch. Hydro-Eng. Environ. Mech*. **49**, 2, 37-57.

Sanjou, M., I. Nezu, S. Suzuki, and K. Itai (2010), Turbulence structure of compound open-channel flows with one-line emergent vegetation, **In:** *J. Hydrodyn. B* **22**, 5, Suppl. 1, 577-581, DOI: 10.1016/S1001-6058(09)60255-9.

Shiono, K., and D.W. Knight (1990), Mathematical models of flow in two or multi stage straight channels. **In:** *Proc. Int. Conf. on River Flood Hydraulics, Wallingford, United Kingdom*, 229-238.

Shiono, K., and D.W. Knight (1991), Turbulent open-channel flows with variable depth across the channel, *J. Fluid Mech.* **222**, 7, 617-646, DOI: 10.1017/S0022112091001246.

Siniscalchi, F., V.I. Nikora, and J. Aberle (2012), Plant patch hydrodynamics in streams: Mean flow, turbulence, and drag forces, *Water Resour. Res.* **48**, 1, W01513, DOI: 10.1029/2011WR011050.

Terrier, B., S. Robinson, K. Shiono, A. Paquier, and T. Ishigaki (2010), Influence of vegetation to boundary shear stress in open channel for overbank flow. **In:** Dittrich, Koll, Aberle, and Geisenhainer (eds.), *Proc. Int. Conf. River Flow 2010, Bruansweig, Germany*, Vol. 1, 285-292.

Tominaga, A., I. Nezu, K. Ezaki, and H. Nakagawa (1989), Three-dimensional turbulent structure in straight open channel flows, *J. Hydraul. Res.* **27**, 1, 149-173, DOI: 10.1080/00221688909499249.

Yang, K., S. Cao, and D.W. Knight (2007), Flow patterns in compound channels with vegetated floodplains, *J. Hydraul. Eng.* **133**, 2, 148-159, DOI: 10.1061/(ASCE)0733-9429(2007)133:2(148).

Discriminating Between Different Streamflow Regimes by Using the Fisher–Shannon Method: An Application to the Colombia Rivers

Jorge O. PIERINI[1], Juan C. RESTREPO[2], Michele LOVALLO[3], and Luciano TELESCA[4]

[1]Centro Científico y Tecnológico de Bahía Blanca (CCT-BB, CIC, UNS), CRIBABB, Bahía Blanca, Argentina; e-mail: jpierini@criba.edu.ar

[2] Grupo de Física Aplicada: Océano y Atmósfera, Departamento de Fisica, Universidad del Norte, Barranquilla, Colombia

[3]Agenzia Regionale per la Protezione dell'Ambiente di Basilicata (ARPAB), Potenza, Italy

[4]National Research Council of Italy, Institute of Methodologies for Environmental Analysis (CNR-IMAA), Tito, Italy

A b s t r a c t

The Fisher–Shannon (FS) information plane, defined by the Fisher information measure (FIM) and the Shannon entropy power (N_X), was robustly used to investigate the complex dynamics of eight monthly streamflow time series in Colombia. In the FS plane the streamflow series seem to aggregate into two different clusters corresponding to two different climatological regimes in Colombia. Our findings suggest the use of the statistical quantity defined by the FS information plane as a tool to discriminate among different hydrological regimes.

Key words: Caribbean streamflow, Fisher information measure, Shannon entropy.

1.   INTRODUCTION

The analysis of the time dynamics of a river streamflow has gained a great importance because it can be considered as the integral of the annual or

interannual climatic fluctuations, which characterize its basin. In fact, the literature dealing with river streamflows is mainly focused on the role of a river streamflow as a sort of climatic indicator to identify and characterize different climatic periods (Pekarova *et al.* 2003, Milliman *et al.* 2008). Indeed, the use of streamflow changes for detecting significant trends, identifying major oscillations periods, and determining relationships between hydrological responses and climatic forcings was performed by several authors. Probst and Tardy (1987), analyzing the annual streamflow data of fifty major rivers distributed all around the world deduced that during the first half of the last century Europe and Asia were affected by a significant humid regime that affected Africa, North and South America in the last half of that century. Pekarova *et al*. (2003) analyzing many streamflows worldwide identified alternating phases of wet and dry periods and extreme cycles of high-low discharge with periods from years to tens of years. But Milliman *et al*. (2008) noted significant changes just in individual rivers and at regional levels. Some authors have also analyzed the relationship between streamflow of Colombian rivers and El Niño-Southern Oscillation (ENSO) phenomenon (Mesa *et al.* 1997, Restrepo and Kjerfve 2000, Gutiérrez and Dracup 2001, Restrepo *et al.* 2014).

From all these studies it is possible to argue that investigating the time structure of streamflows was principally aimed at identifying and possibly quantifying climate-related long-term trends, cycles, scaling, or also anomalous patterns with respect to a certain background.

Streamflows can be considered as the output of systems whose complexity can be measured by its organization and order. In order to get such a knowledge, we need to use appropriate methodological approaches.

In the present paper, we investigate the time dynamics of eight streamflows measured by gauging stations located along the Caribbean plain of Colombia using the Fisher–Shannon (FS) method. The FS approach is based on the information content of the time series using the statistical measures of the Fisher information measure and the Shannon entropy (described in Section 2.2), and is a powerful statistical method to gain into insight the inner dynamics of a complex system.

2.   DATA AND METHODS

2.1 Hydrological data

The Caribbean plain of Colombia is located in the northernmost South America. It extends from the Darien tropical rainforest in the Colombia–Panama border, to the Peninsula de La Guajira in the east, and the slopes of the Cordillera de los Andes in the south (Fig. 1). It comprises extensive lowlands with heights below 100 m, plateaus with heights between 200 and
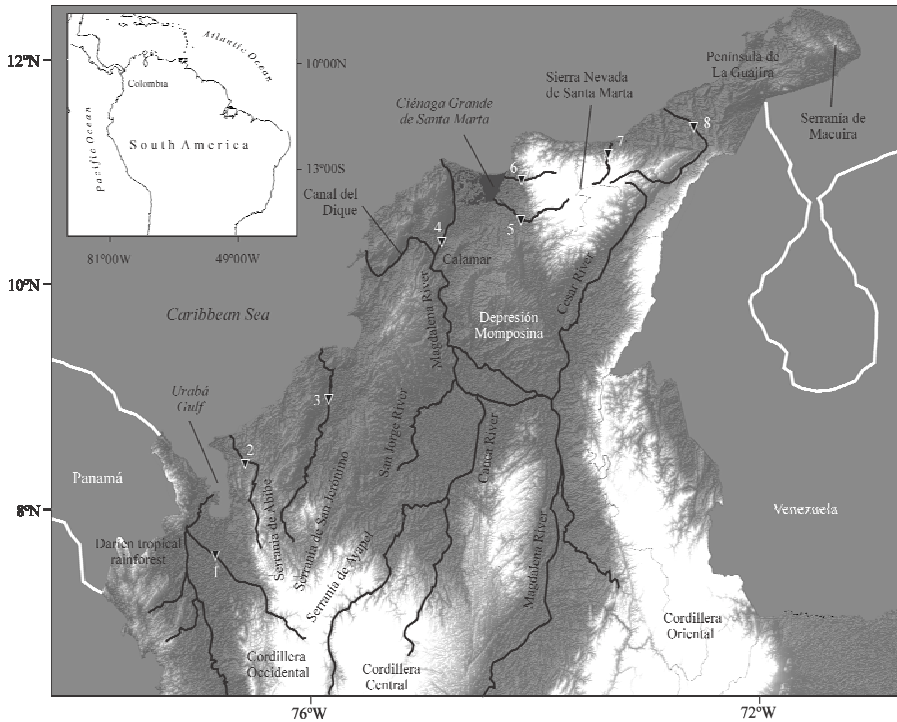
Fig. 1. Caribbean plain of Colombia in northwest South America. Major topographic features, selected rivers and gauge stations: 1 – Sucío River, 2 – Mulatos River, 3 – Sinú River, 4 – Magdalena River (Calamar), 5 – Fundación River, 6 – Frío River, 7 – Palomino River, and 8 – Ranchería River.

Table 1

Drainage basin, river length, headwater, and mean monthly streamflow
of selected rivers in the Caribbean plain of Colombia.

River	Drainage basin [$\times 10^3$ km^2]	River length [km]	Headwater [m a.s.l.]	Mean monthly streamflow [m^3 s^{-1}]
Andean Rivers				
Sucío	4.5	169	4080	278.5
Mulatos	0.01	115	2120	4.6
Sinú	14.7	415	3960	398.1
Magdalena	257.4	1540	3650	6497.1
Sierra Nevada Rivers				
Fundación	1.87	161	2586	28.2
Frío	0.32		3975	13.8
Palomino	0.68		4325	25.7
Ranchería	4.23	151	3725	12.8

Name, location, and historic record period of studied discharge stations

River	Gage station	Location			Historic record
		Elevation [m a.s.l.]	Longitude	Latitude	
Andean Rivers					
Sucio	Mutata	132	76°26W	7°13N	1976-2010
Mulatos	Pueblo Bello	84	76°31W	8°12N	1977-2010
Sinu	Cotoca Abajo	5	75°51W	9°13N	1970-2010
Magdalena	Calamar	8	74°55W	10°15N	1941-2010 (Q_{min}) 1969-2010 (Q_{max})
Sierra Nevada Rivers					
Fundacion	Fundacion	55	74°11W	10°31N	1958-2010
Frio	Rio Frio	30	74°09W	10°34N	1967-2009
Palomino	Puente Carretera	30	73°34W	11°14N	1973-2010
Rancheria	Hacienda Guamito	76	72°37W	11°10N	1979-2007

1000 m in the southwest (Serranías de Abibe, San Jerónimo, and Ayapel) and northeast (Serranía de Macuira), and one of the highest coastal mountains of the world, named Sierra Nevada de Santa Marta, with heights up to 5000 m.

The Instituto de Hidrología, Meteorología y Estudios Ambientales – Colombia (IDEAM) provided monthly maximum and minimum streamflow data from the main rivers of the Caribbean plain of Colombia (Fig. 1). The selection of rivers and corresponding gauging stations was based on two important conditions: (1) the station is located close to the downstream part of the basin; and (2) its hydrological data record is longer than thirty years (see Tables 1 and 2 for the characteristics of the selected rivers).

2.2 The Fisher–Shannon information plane

The Fisher–Shannon (FS) information plane represents an efficient tool to investigate the complex temporal fluctuations of nonstationary signals. It is constructed with coordinate axes given by the Fisher information measure (FIM) and the Shannon entropy power (N_X) that are both well known in the context of information theory. The FIM quantifies the amount of organization or order in a system, while N_X measures its degree of uncertainty or disorder. The FIM was developed by Fisher (1925) in the context of statistical estimation. Then, it was utilized for different aims. Frieden (1990) used the FIM to describe the evolution laws of physical systems. Martin *et al.* (1999, 2001) applied it to characterize the temporal fluctuations of electroencephalograms (EEG) and to detect significant dynamical changes. Complex geophysical and environmental phenomena, like volcano-related signals, earthquake-related electromagnetic signals, and atmospheric particulate mat-

ter, benefited of the application of the FIM methodology to gain insight into their inner time dynamics and the mechanisms underlying their temporal fluctuations and to reveal precursory signatures of critical phenomena (Lovallo and Telesca 2011, Telesca and Lovallo 2011, Telesca *et al.* 2009, 2010, 2011).

Shannon entropy is used to quantify the uncertainty of the prediction of the outcome of a probabilistic event (Shannon 1948a, b); in fact, it is zero for deterministic events. For continuous distributions the Shannon entropy can take any real positive and negative value. In order to avoid the difficulty arising with negative information measures, the so-called Shannon power entropy N_X (defined below) can be used instead of the Shannon entropy.

Let $f(x)$ be the probability density of a signal x, then its FIM I is given by

$$I = \int_{-\infty}^{+\infty} \left(\frac{\partial}{\partial x} f(x) \right)^2 \frac{dx}{f(x)}, \tag{1}$$

and its Shannon entropy is defined as (Shannon 1948a, b):

$$H_X = -\int_{-\infty}^{+\infty} f_X(x) \log f_X(x) dx. \tag{2}$$

As specified above, the notion of Shannon entropy power will be used

$$N_X = \frac{1}{2\pi e} e^{2H_X}. \tag{3}$$

The two measures satisfy the so-called "isoperimetric inequality" $IN_X \geq D$, where D is the dimension of the space (Esquivel *et al.* 2010). The isoperimetric inequality indicates that FIM and Shannon entropy power are linked to each other, and the so-called Fisher–Shannon (FS) information plane represents a tool for the characterization of signals. The product IN_X can also be employed as a statistical measure of complexity (Romera and Dehesa 2004), and for 1-dimensional space, the line $IN_X = 1$ separates the FS plane in two parts, one allowed ($IN_X > 1$), and the other not allowed ($IN_X < 1$).

The calculation of the FIM and the Shannon entropy depends on the calculation of the probability density function $f(x)$ (*pdf*). The *pdf* can be estimated by means of the kernel density estimator technique (Devroye 1987, Janicki and Weron 1994) that approximates the density function as:

$$\hat{f}_M(x) = \frac{1}{Mb} \sum_{i=1}^{M} K\left(\frac{x - x_i}{b} \right), \tag{4}$$

with b the bandwidth, M the number of data, and $K(u)$ the kernel function, a continuous non-negative and symmetric function satisfying the two following conditions:

$$K(u) \geq 0 \text{ and } \int_{-\infty}^{+\infty} K(u)du = 1. \tag{5}$$

In our study, we estimated the *pdf* $f(x)$ by means of the algorithm developed by Troudi *et al.* (2008) combined with that developed in Raykar and Duraiswami (2006), that uses a Gaussian kernel with zero mean and unit variance:

$$\hat{f}_M(x) = \frac{1}{M\sqrt{2\pi b^2}} \sum_{i=1}^{M} e^{-\frac{(x-x_i)^2}{2b^2}}. \tag{6}$$

3. RESULTS AND DISCUSSION

We analyzed eight monthly streamflow time series recorded in Caribbean plains of Colombia. Figure 1 shows the geographical location of the gauging stations. Figure 2 shows the raw data. Selection was based on record length and location of the stations in order to cover a long range of years and a wide Colombian coast. Data cover the instrumental period, are proved to contain good quality measurements and correspond to different basins of Caribbean Colombia. We firstly applied the standard well-known power spectral density analysis to detect periodicities. We can identify annual and seasonal oscillations (as it was shown by the peaks in the power spectra) (Fig. 3 shows the power spectrum of Rancheria Q_{min} data as an example), which we removed before applying the Fisher–Shannon method. Having filtered out the annual and seasonal cycles and normalized the monthly maximum ($Q_{max,N}$) and minimum ($Q_{min,N}$) streamflows (in order to avoid any dependence on the real amount of water volume per second), for each site we constructed a residual time series as ($Q_{max,N} + Q_{min,N}$)/2   (Fig. 4).

Figure 5 shows the Fisher–Shannon (FS) information plane: the y-axis represents the FIM and the x-axis represents the Shannon entropy power N_X; each symbol represents a residual streamflow series. The analysis of the FS information plane allows to discriminate two clusters, the one comprising the north-eastern sites (Rancheria, Palomino, Frio, and Fundación) and the other comprising the south-western ones (Sucio, Sinu, and Magdalena); Mulatos, even though it belongs to the second group, in the FS plane is closer to the first group. The two clusters reflect the different hydrological conditions that characterize the Andean rivers and the Sierra Nevada ones: the Andean rivers drain rather large basins, with Magdalena being the largest fluvial system
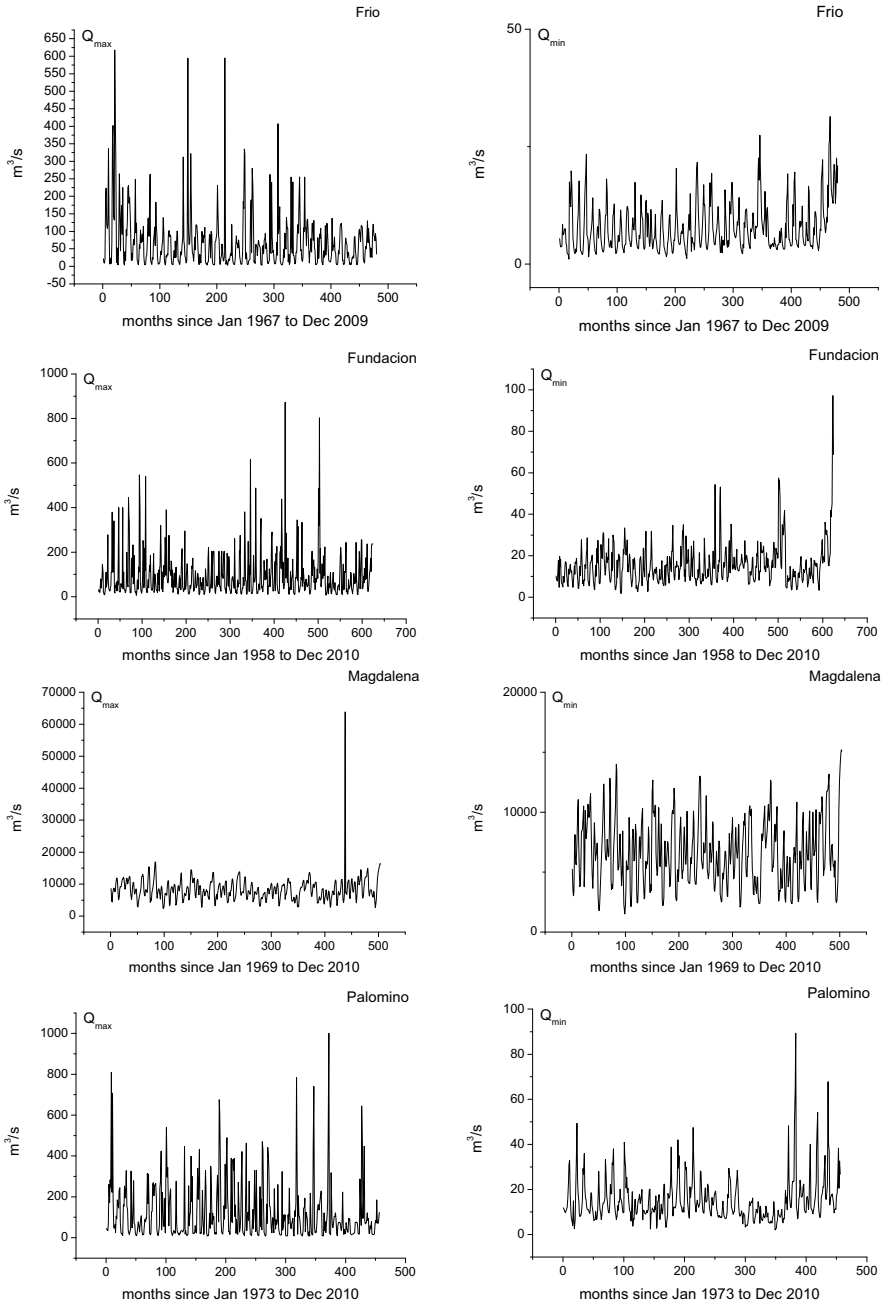
Fig. 2. Continued on the next page.
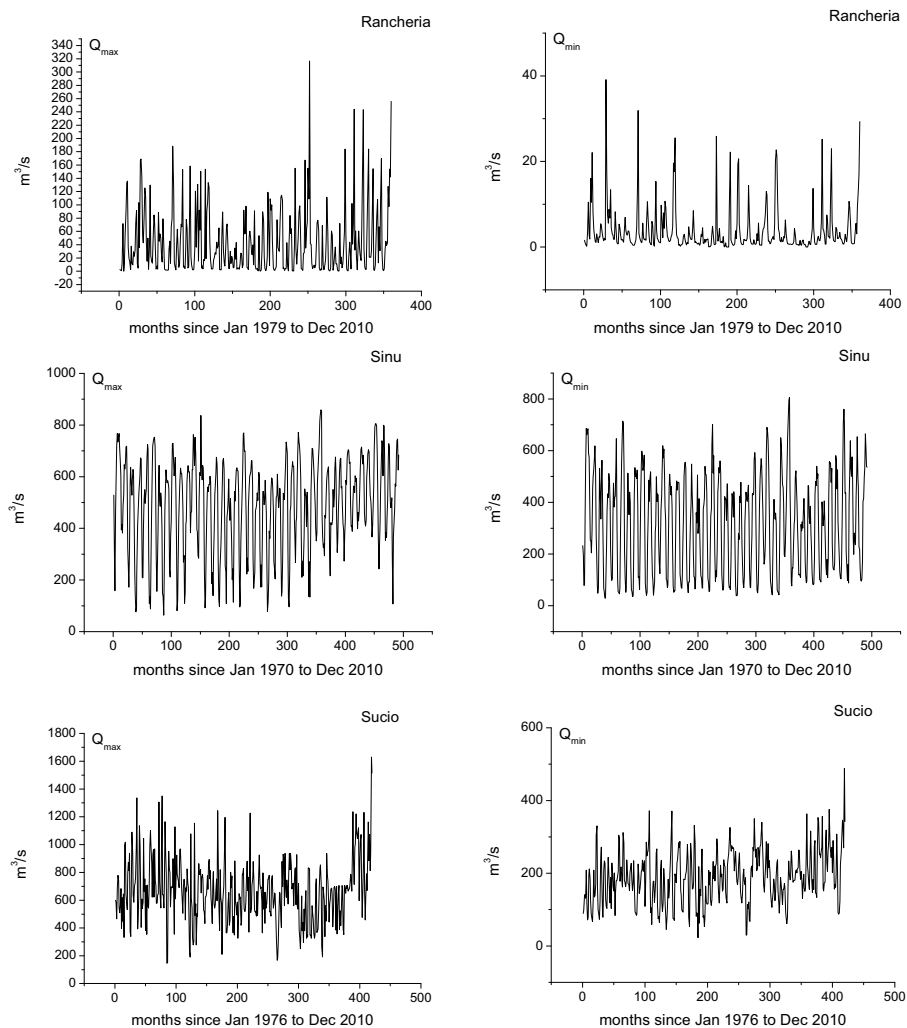
Fig. 2. Mean monthly maximum and minimum streamflows at the selected sites shown in Fig. 1.

in the Caribbean plain; while the rivers of the Sierra Nevada de Santa Marta drain small mountainous basins ($< 5000 \text{ km}^2$) with steep gradients and limited alluvial floodplains. The Andean Mulatos river has a local pattern very likely influenced by the local orographic effects, with a small drainage basin that makes it closer to the Sierra Nevada cluster.

The Fisher–Shannon analysis shows that the Andean rivers are characterized by lower FIM (higher N_X) and the Sierra Nevada ones by higher FIM
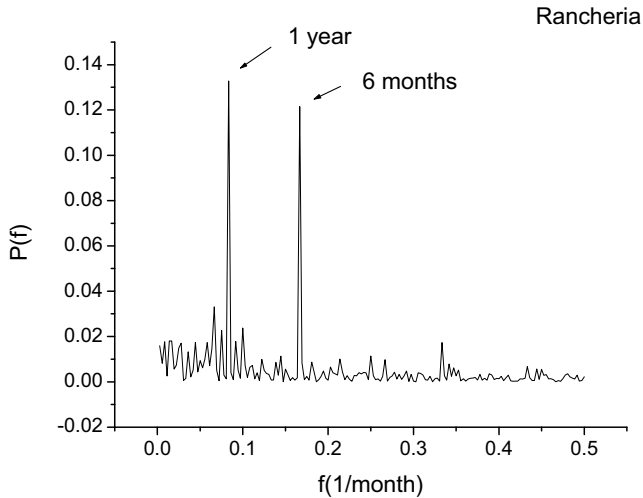
Fig. 3. Power spectrum of Rancheria Q_{min} data.

(lower N_X). This indicates that the Andean rivers are featured by lower system organization and order, while the Sierra Nevada ones by lower system disorder and higher organization. Hence, the Andean rivers undergo a disordered regime, with no preference for any particular set of states, while the Sierra Nevada ones undergo an ordered regime, in which a particular set of states is preferred. In fact, since the fluvial systems of the Sierra Nevada have drainage areas smaller than 5000 km^2 in mountainous zones, the topographical setting is a primary factor controlling streamflow variability. Instead, the Andean rivers, since they drain extensive plateaus or low-lying alluvial valleys, are subject to the coexistence of several controlling factors that enhance their disorder degree and lower their organization structure. For instance, the Magdalena river has a basin formed by 151 subcatchments, 42 of which are second-order watersheds, several main tributaries, like Cauca river (second largest river in Colombia), and, furthermore, characterized by high tectonic activity, hillslopes commonly exceeding 45°, landslides, steep gradients, and high relief tributary basins (Restrepo and Restrepo 2005, Restrepo *et al*. 2006). Fath *et al*. (2003) asserted that the Shannon entropy can be considered as a measure of the degrees of freedom of a system; and, hence, a system with few degrees of freedom has a high information content. In our case, the Sierra Nevada rivers, being less extensive and with a basin smaller and more concentrated than those of Andean rivers, are characterized by streamflow time variability with a lower number of degrees of freedom and high information content, as clearly indicated by
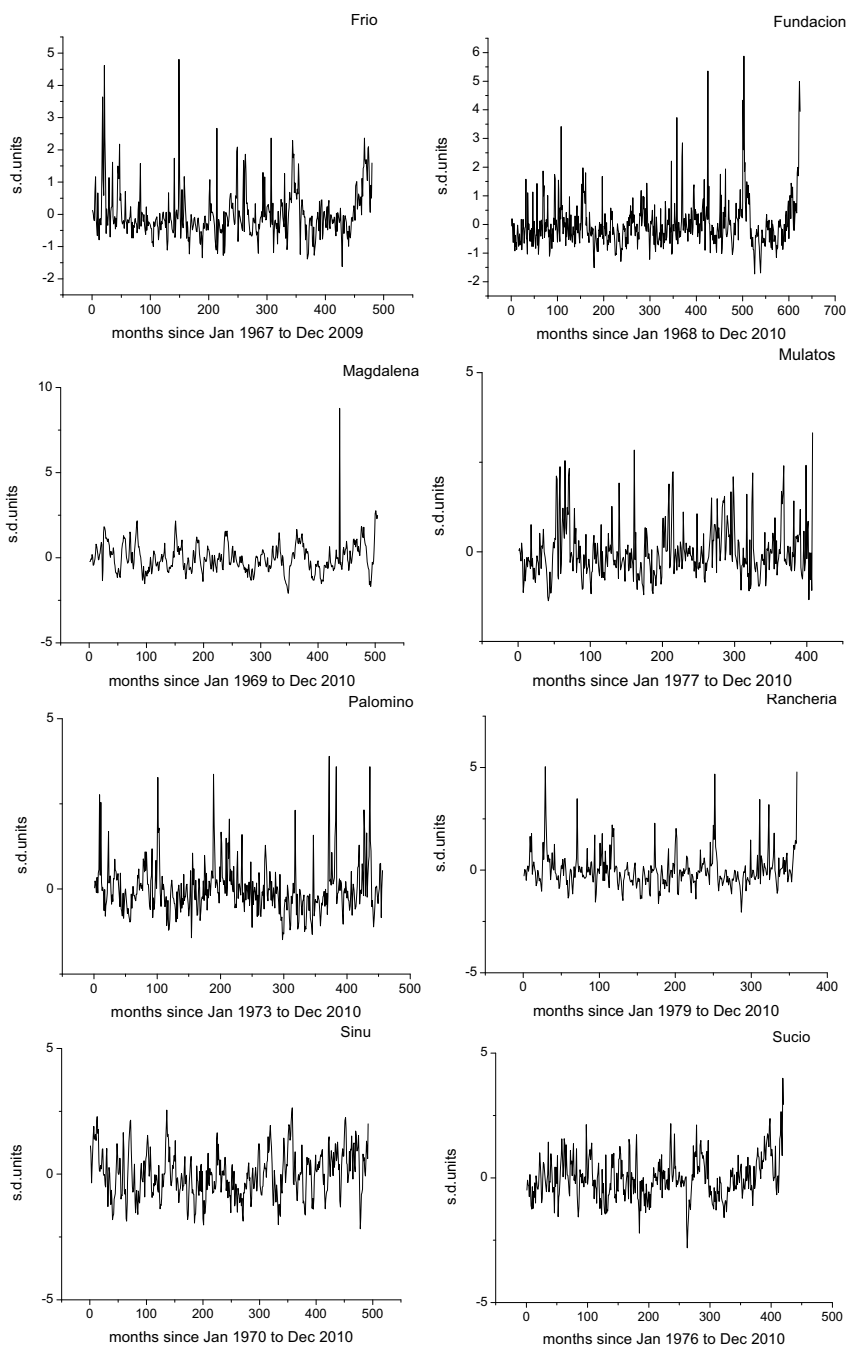
Fig. 4. Residual time series given by the average between the filtered normalized monthly maximum and minimum streamflows at the selected sites shown in Fig. 1.
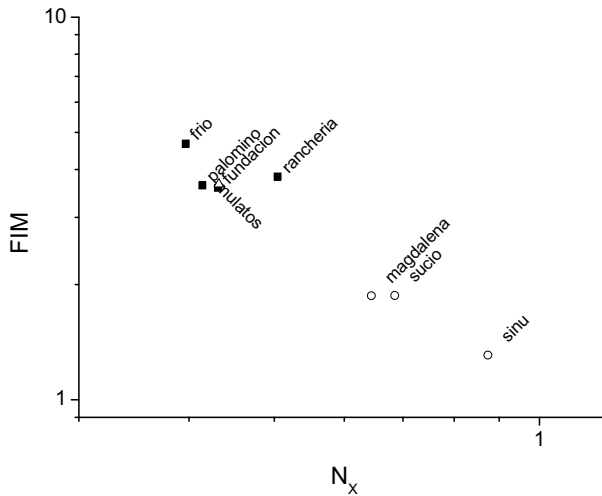
Fig. 5. Fisher–Shannon (FS) information plane: the *y*-axis represents the FIM and the *x*-axis represents the Shannon entropy power N_X; each symbol represents a residual streamflow series.

the FS information plane. As it can be argued, such a number of streamflow controlling factors, larger for the Andean rivers than Sierra Nevada ones, may organize and order the streamflow data, as indicated by the significantly different FIM and Shannon entropy power values.

4. CONCLUSIONS

We analyzed eight monthly streamflow time series recorded in Caribbean plain of Colombia by using the Fisher–Shannon statistical method. The streamflow time series aggregate into two different clusters: one corresponds to Andean Rivers Group (except of Mulatos) and the other corresponding to the Sierra Nevada Group. The Andean Group is characterized by higher Shannon entropy power and lower FIM with respect to the Sierra Nevada Group. This indicates that the Sierra Nevada rivers are featured by higher organization and order than the Andean rivers that, on the contrary, are characterized by greater disorder and uncertainty. A connection between these dynamical properties and the particular characteristics of the topography and drainage area of the two groups was suggested.

R e f e r e n c e s

Devroye, L.A. (1987), *Course on Density Estimation*, Birkhauser, Boston.

Esquivel, R.O., J.C. Angulo, J. Antolin, J.S. Dehesa, S. López-Rosa, and N. Flores-Gallegos (2010), Analysis of complexity measures and information planes of selected molecules in position and momentum spaces, *Phys. Chem. Chem. Phys.* **12**, 7108-7116, DOI: 10.1039/B927055H.

Fath, B.D., H. Cabezas, and C.W. Pawlowski (2003), Regime changes in ecological systems: an information theory approach, *J. Theor. Biol.* **222**, 4, 517-530, DOI: 10.1016/S0022-5193(03)00067-5.

Fisher, R.A. (1925), Theory of statistical estimation, *Math. Proc. Cambridge Philos. Soc.* **22**, 5, 700-725, DOI: 10.1017/S0305004100009580.

Frieden, B.R. (1990), Fisher information, disorder, and the equilibrium distributions of physics, *Phys. Rev. A* **41**, 4265-4276, DOI: 10.1103/PhysRevA.41.4265.

Gutiérrez, F., and J.A. Dracup (2001), An analysis of the feasibility of long-range streamflow forecasting for Colombia using El Niño–Southern Oscillation indicators, *J. Hydrol.* **246**, 1-4, 181-196, DOI: 10.1016/S0022-1694(01)00373-0.

Janicki, A., and A. Weron (1994), *Simulation and Chaotic Behavior of alpha-Stable Stochastic Processes*, Marcel Dekker Inc., New York.

Lovallo, M., and L. Telesca (2011), Complexity measures and information planes of x-ray astrophysical sources, *J. Stat. Mech.* **2011**, 3, P03029, DOI: 10.1088/1742-5468/2011/03/P03029.

Martin, M.T., F. Pennini, and A. Plastino (1999), Fisher's information and the analysis of complex signals, *Phys. Lett. A* **256**, 2-3, 173-180, DOI: 10.1016/S0375-9601(99)00211-X.

Martin, M.T., J. Perez, and A. Plastino (2001), Fisher information and nonlinear dynamics, *Physica A* **291**, 1-4, 523-532, DOI: 10.1016/S0378-4371(00)00531-8.

Mesa, O., G. Poveda, and L. Carvajal (1997), *Introducción al Clima de Colombia,* Universidad Nacional de Colombia, Bogotá, Colombia, 390 pp. (in Spanish).

Milliman, J.D., K.L. Farnsworth, P.D. Jones, K.H. Xu, and L.C. Smith (2008), Climatic and anthropogenic factors affecting river discharge to the global

ocean, 1951-2000, *Global Planet. Change* **62**, 3-4, 187-194, DOI: 10.1016/j.gloplacha.2008.03.001.

Pekarova, P., P. Miklanek, and J. Pekar (2003), Spatial and temporal runoff oscillation analysis of the main rivers of the world during the 19th-20th centuries, *J. Hydrol.* **274**, 1-4, 62-79, DOI: 10.1016/S0022-1694(02)00397-9.

Probst, J.L., and Y. Tardy (1987), Long range streamflow and world continental runoff fluctuations since the beginning of this century, *J. Hydrol.* **94**, 3-4, 289-311, DOI: 10.1016/0022-1694(87)90057-6.

Raykar, V.C., and R. Duraiswami (2006), Fast optimal bandwidth selection for kernel density estimation. **In**: J. Ghosh, D. Lambert, D. Skillicorn, and J. Srivastava (eds.), *Proc. Sixth SIAM Int. Conf. Data Mining, 20-22 April 2006, Bethesda, USA*, 524-528.

Restrepo, J.C., and J.D. Restrepo (2005), Efectos naturales y antrópicos en la producción de sedimentos de la Cuenca del Río Magdalena, *Rev. Acad. Col. Cien. Exact. Fís. Nat.* **29**, 111, 239-254 (in Spanish).

Restrepo, J.C., J.C. Ortíz, J. Pierini, K. Schrottke, M. Maza, L. Otero, and J. Aguirre (2014), Freshwater discharge into the Caribbean Sea from the rivers of Northwestern South America (Colombia): Magnitude, variability and recent changes, *J. Hydrol.* **509**, 266-281, DOI: 10.1016/j.jhydrol.2013.11.045.

Restrepo, J.D., and B. Kjerfve (2000), Magdalena river: interannual variability (1975-1995) and revised water discharge and sediment load estimates, *J. Hydrol.* **235**, 1-2, 137-149, DOI: 10.1016/S0022-1694(00)00269-9.

Restrepo, J.D., B. Kjerfve, M. Hermelin, and J.C. Restrepo (2006), Factors controlling sediment yield in a major South American drainage basin: the Magdalena River, Colombia, *J. Hydrol.* **316**, 1-4, 213-232, DOI: 10.1016/j.jhydrol.2005.05.002.

Romera, E., and J.S. Dehesa (2004), The Fisher–Shannon information plane, an electron correlation tool, *J. Chem. Phys.* **120**, 19, 8906-8912, DOI: 10.1063/1.1697374.

Shannon, C.E. (1948a), A mathematical theory of communication, *Bell Syst. Tech. J.* **27**, 3, 379-423, DOI: 10.1002/j.1538-7305.1948.tb01338.x.

Shannon, C.E. (1948b), A mathematical theory of communication, *Bell Syst. Tech. J.* **27**, 4, 623-656, DOI: 10.1002/j.1538-7305.1948.tb00917.x.

Telesca, L., and M. Lovallo (2011), Analysis of the time dynamics in wind records by means of multifractal detrended fluctuation analysis and the Fisher–Shannon information plane, *J. Stat. Mech.* **2011**, 7, P07001, DOI: 10.1088/1742-5468/2011/07/P07001.

Telesca, L., M. Lovallo, A. Ramirez-Rojas, and F. Angulo-Brown (2009), A nonlinear strategy to reveal seismic precursory signatures in earthquake-related self-potential signals, *Physica A* **388**, 10, 2036-2040, DOI: 10.1016/j.physa.2009.01.035.

Telesca, L., M. Lovallo, and R. Carniel (2010), Time-dependent Fisher Information Measure of volcanic tremor before the 5 April 2003 paroxysm at Stromboli volcano, Italy, *J. Volcano. Geoth. Res.* **195**, 1, 78-82, DOI: 10.1016/ j.jvolgeores.2010.06.010.

Telesca, L., M. Lovallo, H.-L. Hsu, and C.-C. Chen (2011), Analysis of dynamics in magnetotelluric data by using the Fisher-Shannon method, *Physica A* **390**, 7, 1350-1355, DOI:. 10.1016/j.physa.2010.12.005.

Troudi, M., A.M. Alimi, and S. Saoudi (2008), Analytical plug-in method for kernel density estimator applied to genetic neutrality study, *EURASIP J. Adv. Sig. Proc.* **2008**, 739082, DOI: 10.1155/2008/739082.

Testing Texture of VHR Panchromatic Data as a Feature of Land Cover Classification

Stanisław LEWIŃSKI, Sebastian ALEKSANDROWICZ,
and Marek BANASZKIEWICZ

Space Research Centre, Polish Academy of Sciences, Warszawa, Poland;
e-mails: stlewinski@cbk.waw.pl (corresponding author),
saleksandrowicz@cbk.waw.pl, marekb@cbk.waw.pl

A b s t r a c t

While it is well-known that texture can be used to classify very high resolution (VHR) data, the limits of its applicability have not been unequivocally specified. This study examines whether it is possible to divide satellite images into two classes associated with "low" and "high" texture values in the initial stage of processing VHR images. This approach can be effectively used in object-oriented classification. Based on the panchromatic channel of KOMPSAT-2 images from five areas of Europe, datasets with down-sampled pixel resolutions of 1, 2, 4, 8, and 16 m were prepared. These images were processed using different texture analysis techniques in order to discriminate between basic land cover classes. Results were assessed using the normalized feature space distance expressed by the Jeffries–Matusita distance. The best results were observed for images with the highest resolution processed by the Laplacian filter. Our research shows that a classification approach based on the idea of "low" and "high" textures can be effectively applied to panchromatic data with a resolution of 8 m or higher.

Key words: object-oriented classification, land cover classification, texture, Laplacian filter.

1. INTRODUCTION

Texture is expressed by the frequency of tonal values (brightness or colour) of neighbouring points of an image and is one of the most important features for the interpretation of data gathered through remote sensing. In visual classification, texture is as important as the tone, shape, and size of the identified object. This is also the case in automatic classification, which is based on the digital numbers of pixels that make up an image (Tuceryan and Jain 1999). Image processing systems for the interpretation of satellite data have functions for the calculation of texture parameters such as: variance, contrast, entropy, homogeneity or correlation. The spatial extent of the analyses is determined by the size of the moving window. Computations are usually performed on square windows of 3×3 or 5×5 pixels and the resulting value is output to a layer that corresponds to the position of the window's central pixel.

More complex calculations are based on the Grey Level Co-occurrence Matrix (GLCM). With this method, texture can be determined not only within a pre-defined window, but also for an object of any shape. The first step is to make the co-occurrence matrix symmetrical around the diagonal. Then it is normalized and at this point it represents the probability of occurrence of data values. Next, it is processed in order to describe the texture of the object (Hall-Beyer 2000). Based on the GLCM, Haralick *et al.* (1973) proposed 14 features for image classification. Another method of determining texture is to use image filtration methods. Although these are usually based on the moving window technique, their goal is not specifically to measure texture but rather to increase and highlight it. There are many different high pass and edge detection filters (including adaptive statistical filters) that can be used for this purpose. Similarly, filters that smooth the image by removing anomalies can be applied to highlight textural features.

Texture plays a particularly important role in object-oriented algorithms (Blaschke 2010). In contrast to traditional pixel-based approaches, these algorithms work at the level of entire objects rather than individual pixels. Objects are formed through a segmentation process that groups pixels that meet given similarity conditions. Objects can be then analysed on the basis of a variety of distinctive features, including texture. The computation of the texture of the area identified as the object is based on spectral channels processed by texture functions or filters. Haralic functions are implemented in object-oriented software such as eCognition (eCognition Developer 2011) and are often used despite very time-consuming calculations. Texture is applied not only for class identification but also as an additional data source for image segmentation (Ryherd and Woodcock 1996, Hofmann *et al.* 1998, Hu *et al.* 2005).

Textural features are mainly applied in order to identify built-up areas on VHR satellite images. GLCM parameters are used to discriminate urban and non-urban areas and to extract urban areas from other land cover classes (De Martinao *et al.* 2003, Morales *et al.* 2003, Puissant *et al.* 2005, Wang *et al.* 2003, He *et al.* 2005, Su *et al.* 2008, Lewinski and Bochenek 2009). Kit *et al.* (2012) used VHR satellite images to identify informal settlements. De Kok (2012) proposed the simulation of a large 49×49 Laplacian filter for testing neighbourhood contrast in order to enhance the detection of artificial areas. Esch *et al.* (2010) provide examples of urban classification based on the analysis of the texture of synthetic aperture radar (SAR) images. Berberoglu *et al.* (2007) assessed usefulness of texture measures for maximum likelihood and artificial neural networks classification routines. They tested texture measures calculated for three different windows (3×3, 5×5, 7×7) over the Landsat TM image of Mediterranean area. Textural analysis is not only useful for typical land cover classification; other examples are forest biomass estimation (Eckert 2012) or the classification of sub-Antarctic vegetation (Murray *et al.* 2010).

The algorithm proposed by De Kok and Wezyk (2008) was based on the hypothesis that it is possible to divide an image into so-called "low" and "high" textures from calculations based on high-resolution panchromatic channels. Low texture values are associated with agricultural land and bodies of water. Other land cover classes (primarily various species of forests and built-up areas) are characterized by higher texture. The strategy is based on a sequential classification – from general to more detailed. The division into two textural groups is very helpful because the algorithm can be targeted at the initial stage of the classification. In general, it can be assumed that once all the water bodies have been identified, agricultural land (the most complex land cover class) will be "automatically" recognized because it belongs to the rest of the "low" texture area. Other land cover classes belong to high texture areas and are classified in separate processes. A simplified work-flow is presented in Fig. 1.

A threshold value between low and high textures can be defined interactively using designated sets of training fields for agricultural areas, water bodies, forests, and buildings. De Kok and Wezyk (2008) proposed using the distribution of histogram quantiles to express the threshold value. Similarly, Nussbaum *et al.* (2006) developed statistical tools to determine the threshold (the crossing point) between the distributions of textures in sets of training fields.

The idea of initializing the classification using "low" and "high" textures was implemented in the SATChMo-K2 algorithm. The algorithm was developed as one of the Core Mapping Services of the Geoland2 project SATChMo (Seasonal and Annual Change Monitoring). The Area Frame Sam-
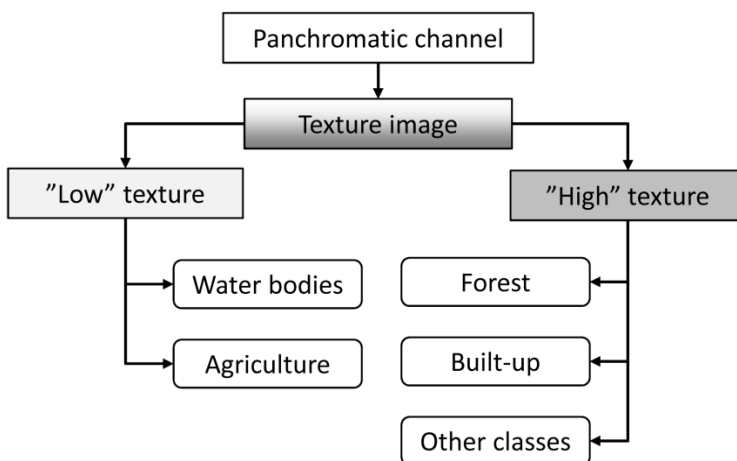
Fig. 1. Land cover classification work-flow based on assumptions of "low" and "high" texture.

pling Europe (AFS Europe) subproject generated generic maps of land cover. This approach was designed for the classification of KOMPSAT-2 and FORMOSAT-2 VHR satellite images (Lewinski *et al.* 2014). Texture was calculated based on panchromatic channels processed by sigma filters. The algorithm has also been used to generate land cover maps based on other VHR datasets (Ikonos, QuickBird, and SPOT). The use of texture as a basic classifier made it possible to achieve similar results irrespective of the season (spring, summer or autumn) and to use the same classification features for images of similar European regions.

The project ran extensive experiments. Seven teams classified the basic land cover classes of more than 120 images scattered across Europe. This work showed that the classification approach is sensitive to spatial resolution. Both KOMPSAT-2 and FORMOSAT-2 images belong to the group of VHR data, but the spatial resolution of KOMPSAT-2 (PAN – 1 m, MS – 4 m) is four times better than that of FORMOSAT-2 (PAN – 2 m, MS – 8 m); it is therefore easier to divide KOMPSAT-2 images into "low" and "high" textures. Moreover, the effectiveness of the algorithm depended on the geographic area. The most difficult area to classify was Southern Europe, especially Spain. This was mainly due to dispersed vegetation, and not seasonal differences.

The work of the SATChMo team forms the basis for the study described here. The initial texture-based classification was a crucial step and more insight is needed into the correct selection of textural parameters. Parameter selection mainly depends on the sensor resolution and the geographic loca-

tion. The aim of this study was therefore to determine whether it would be possible to use the "low" and "high" texture classification approach for the panchromatic images of different spatial resolution.

2. SATELLITE DATA AND TEXTURE PREPROCESSING

The work was performed using KOMPSAT-2 satellite images provided by the European Space Agency (ESA), which were intended for land cover classification in the context of the SATChMo project. Five images were selected from over one hundred datasets, scattered across Europe. They included typical types of land cover found in central, northern, and southern Europe. Analyses were performed on a subset of panchromatic data defined for each image. The selected images are specified in Table 1 and shown in Fig. 2. The Spanish image was representative of the Mediterranean countries; it was included because of the difficulty of classifying this type of land cover. The selected image subsets do not have an equal area. This is because we were trying to find image fragments that included all types of land cover.

Table 1

Selected subsets of satellite images

No.	Satellite data	Date of acquisition	Resolution of PAN data	Coordinates of a subset (centre) φ, λ	Country	Area [km^2]
1	KOMPSAT-2	10 Sep 2009	1 m	52°10′ 18.6″ N 17°24′ 02.8″ E	Poland	85.5
2	KOMPSAT-2	19 Sep 2009	1 m	53°44′ 31.8″ N 8°40′ 20.0″ E	Germany	73.0
3	KOMPSAT-2	9 Aug 2009	1 m	63°40′ 25.5″ N 26°31′ 30.3″ E	Finland	83.0
4	KOMPSAT-2	21 Jul 2009	1 m	45°05′ 55.2″ N 0°46′ 12.0″ W	France	70.8
5	KOMPSAT-2	26 Aug 2009	1 m	43°19′ 41.3″ N 4°06′ 55.6″ W	Spain	91.7

Lower-resolution images were simulated using the 1 m resolution panchromatic channel of KOMPSAT-2. We applied the "degradation" function of the ERDAS Imagine software and generated images with pixel sizes of 2, 4, 8, and 16 m. Using this technique, a panchromatic dataset of 1-16 m resolution was collected for each test area. Figure 3 shows subsets at resolutions of 1 and 8 m.

Fig. 2. Selected subsets of satellite images: (a) France, (b) Finland, (c) Spain,
(d) Germany, and (e) Poland.

Fig. 3. Panchromatic channel of KOMPSAT-2 image at 1 m (a) and 8 m (b) resolu-
tion (German test site).

In order to obtain textural information, panchromatic channels were processed using six image texture analysis techniques: the PanBF transformation, the sum of two Lee Sigma filters, the Sobel filter, the Laplacian filter, Gray-Level Co-occurrence Matrix (GLCM) homogeneity, and GLCM entropy. These functions were selected because they are the most useful for land cover classification based on the SATChMo-K2 algorithm.

The PanBF transformation was proposed by Wezyk and De Kok (2006). It is the ratio of panchromatic data and the sum of two edge-detection images processed inside a 3×3 moving window by the Lee Sigma filter:

$$\text{PanBF} = \frac{\text{PAN}}{\text{LeeSigma} + \text{LeeSigma}\_n} \ , \tag{1}$$

where PanBF is the PanBF transformation, PAN is panchromatic data, Lee-Sigma is panchromatic data processed using the Lee Sigma filter, Lee-Sigma$\_n$ is inversed (negative) panchromatic data processed using the Lee Sigma filter.

The Lee Sigma filter first computes a mean and standard deviation for a given window. Next the number of pixels that satisfy the criteria defined by the sigma value is tested and the output is calculated based on this value (Jensen 1996).

The second feature texture test also used the Lee Sigma filter. This is a sum of two Lee Sigma filters and it was applied in the SATChMo-K2 algorithm. eCognition software was used to calculate the sum of Lee Sigma filters with "dark" and "bright" options (eCognition Developer 2011). Edge detection was performed inside a moving 3×3 window for dark and bright objects, respectively. Results were similar to LeeSigma and LeeSigma$\_n$.

$$\text{Sum\_Sigma} = \text{LeeSigma(dark)} + \text{LeeSigma(bright)} \ , \tag{2}$$

where Sum_Sigma is the sum of Lee Sigma filters, LeeSigma is panchromatic data processed using eCognition's Lee Sigma filter with "dark" and "bright" options.

The Sobel filter is an edge detection filter. The output value is computed based on two 3×3 kernels (Jain 1989):

$$\text{Sobel} = \sqrt{X^2 + Y^2} \ , \tag{3}$$

where X and Y represent the convolution mask:

$$X = \begin{matrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{matrix} \qquad Y = \begin{matrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{matrix} \ . \tag{4}$$

Laplacian filters were designed for edge enhancement (Jensen 1996). Image processing is based on elements of the convolution mask. The 3×3 mask that was applied is presented below:

$$\text{Laplacian mask} = \begin{matrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{matrix} . \tag{5}$$

GLCM texture functions were proposed by Haralick *et al.* (1973) and they are regularly used in object-oriented classification. Homogeneity and entropy were selected from among the various GLCM functions because they are not well correlated. The calculation consists of two steps. First, the co-occurrence matrix is defined based on the pixels that make up the object, and then this matrix is transformed using the following expressions:

$$\text{Homogenity} = \sum_{i,j}^{n} \frac{c_{i,j}}{1 + (i - j)^2} , \tag{6}$$

$$\text{Entropy} = \sum_{i,j}^{n} C_{i,j} * \log(C_{i,j}) , \tag{7}$$

where C is the co-occurrence matrix; n, i, and j are the dimensions and elements of the matrix C.

Two examples of texture images are presented in Fig. 4. They were obtained from the image subsets shown in Fig. 3 with resolutions of 1 and 8 m, processed by a Laplacian filter.

(a)   (b)

Fig. 4. Panchromatic channel processed by a Laplacian 3×3 filter; 1 m (a) and 8 m (b) resolution (German test site).
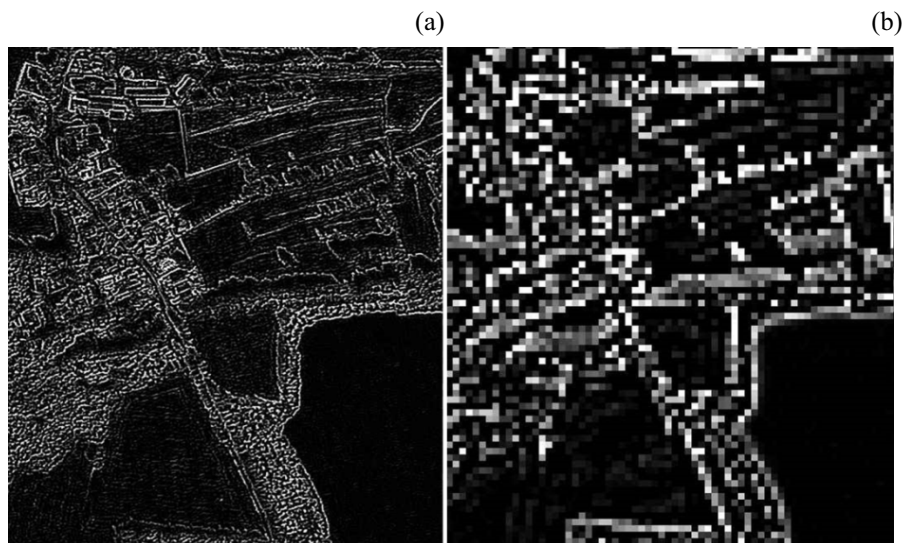
3. CLASS SEPARATION

The aim of the analysis was to determine the ability of the various texture measurement functions to define the threshold between "low" and "high" texture. The functions were tested on panchromatic data at different resolutions. For each of the images, training sets of four basic land cover classes were selected: water bodies, agricultural areas, forest and built-up areas. Next, we calculated separation measures between land cover classes based on texture. It is usually assumed that "water" and "agricultural areas" are associated with low values of texture while "forest" and "built-up areas" are associated with high values. For the Spanish image, a new class called "shrubland" was included in the analysis in order to represent the dispersed vegetation that occurs in the South of Europe.

Analyses were based on an object-oriented classification and the objects to be analysed were defined in the segmentation process. The segmentation procedure depends on the type of data (its spectral, spatial, and radiometric resolution) and the Earth's land cover structure. As segmentation defines the shapes of objects, it strongly influences the classification results. A "chessboard" segmentation was applied; this was both to avoid the complex process of finding optimal segmentation parameters for each case and also to minimize the impact of segmentation on the final results. Consequently, each image was divided into a square object measuring 128×128 m. Next, objects that were completely filled with one of the chosen land cover classes were selected as training sets. This process was made possible by the results of the land cover classification performed in the context of the SATChMo project.

The extent of class discrimination was determined using the Jeffries–Matusita (J-M) distance. The J-M distance is a widely-used measure for feature selection. Calculations were based on the Bhattacharyya distance, which was defined assuming that the distributions of the analysed classes were Gaussian. Jensen (1996) provides the following formulas:

$$\text{Bhat}_{cd} = \frac{1}{8}\left(M_c - M_d\right)^T \left(\frac{V_c + V_d}{2}\right)^{-1} \left(M_c - M_d\right) + \frac{1}{2}\ln\left[\frac{\left|\dfrac{V_c + V_d}{2}\right|}{\sqrt{|V_c||V_d|}}\right], \qquad (8)$$

where Bhat_{cd} is the Bhattacharyya distance between classes c and d, M_c and M_d are the means of classes c and d, V_c and V_d are the covariance matrices of classes c and d.

$$\text{JM}_{cd} = \sqrt{2\left(1 - e^{\text{Bhat}_{cd}}\right)}, \qquad (9)$$

where JM_{cd} is the Jeffries–Matusita distance between classes c and d.

The spectral separability can be described by the squared J-M distance and in this case the values are squared to range between 0 and 2.

4. RESULTS

Our work aimed to determine the possibility of dividing images into "low" and "high" texture values. We calculated the J-M distance between pairs of land cover (LC) classes: water, agricultural areas, forest, and built-up areas. Five test sites in Poland, Germany, Finland, France, and Spain were analysed. For each site, six textures obtained from panchromatic data with resolution 1, 2, 4, 8, and 16 m were analysed. Results were obtained for a total of 150 cases. The full set of results for the German test image is presented in Tables 2-6. These tables show the J-M distances for different image resolutions. Columns represent texture features, while rows list pairs of land cover classes. Similar results were obtained for test sites in Poland, Finland, and France. The results were a little different for the Spanish test site, and they are presented in Tables 7-11.

We assumed that two classes were completely separate if the squared J-M distance was 2, while a value of 1.4 or less indicated potential overlapping of classes (Intergraph 1999). Thomas *et al.* (2003) defined good and poor separability for squared J-M distances as > 1.9 and < 1.0, respectively. We assumed that it was possible to define the threshold between "low" and "high" texture if the value was over 1.7, and these values are marked bold in the tables that follow.

In Tables 12-16 the squared J-M distances are calculated for the combined classes: (i) "water bodies" and "agricultural areas", and (ii) "forest" and "built-up areas". Information about the separation of overlapping classes is particularly important from the point of view of the definition of the threshold between low and high texture. Like the previous tables, results are divided according to spatial resolution (1, 2, 4, 8, and 16 m), and values over 1.7 are highlighted in bold.

As expected, the most promising results – indicating the ability to divide an image into low and high textures – were obtained with a resolution of 1 m (Table 12). This classification was possible for images from Poland, Germany, Finland, and France. Almost identical results were obtained using Laplace and Sum_Sigma filtration. Sobel filtering made it possible to obtain slightly better results for images from Germany and Finland, slightly worse for Poland and much worse for France. However, for the Spanish image the J-M distances indicated that distinguishing between low and high texture was impossible.

The results shown in Table 13 for a 2 m resolution are similar to those shown previously. The Sum_Sigma and Laplace filters performed best, while the Sobel filter did not perform as well. Other transformation tech-

Table 2

PAN data – 1 m, Germany. Squared J-M distance between basic LC classes

Land cover classes	Texture functions					
	PanBF	Laplace 3 × 3	Sobel	Sum_ Sigma	GLCM homo- geneity	GLCM entropy
agriculture/forest	**1.78**	**1.91**	**1.96**	**1.91**	1.24	**1.78**
forest/built-up	0.14	0.09	0.22	0.09	1.26	0.56
built-up/water	**2.00**	**2.00**	**2.00**	**2.00**	**1.94**	**2.00**
agriculture/built-up	**1.72**	**1.97**	**1.99**	**1.97**	0.71	**1.74**
forest/water	**2.00**	**2.00**	**2.00**	**2.00**	**1.99**	**2.00**
agriculture/water	1.34	**2.00**	**2.00**	**2.00**	0.46	1.58

Table 3

PAN data – 2 m, Germany. Squared J-M distance between basic LC classes

Land cover classes	Texture functions					
	PanBF	Laplace 3 × 3	Sobel	Sum_ Sigma	GLCM homo- geneity	GLCM entropy
agriculture/forest	1.69	**1.95**	**1.95**	**1.94**	1.37	**1.75**
forest/built-up	0.18	0.11	0.82	0.14	1.13	0.15
built-up/water	**2.00**	**2.00**	**2.00**	**2.00**	**2.00**	**2.00**
agriculture/built-up	1.63	**1.99**	**2.00**	**1.98**	0.92	**1.77**
forest/water	**2.00**	**2.00**	**2.00**	**2.00**	**2.00**	**2.00**
agriculture/water	1.00	**2.00**	**2.00**	**2.00**	0.43	1.56

Table 4

PAN data – 4 m, Germany. Squared J-M distance between basic LC classes

Land cover classes	Texture functions					
	PanBF	Laplace 3 × 3	Sobel	Sum_ Sigma	GLCM homo- geneity	GLCM entropy
agriculture/forest	1.62	1.92	1.54	1.89	1.50	1.64
forest/built-up	0.03	0.84	1.78	0.98	0.70	0.77
built-up/water	2.00	2.00	2.00	2.00	2.00	2.00
agriculture/built-up	1.61	2.00	2.00	1.99	1.24	1.80
forest/water	2.00	2.00	2.00	2.00	2.00	2.00
agriculture/water	1.34	2.00	1.97	2.00	0.39	1.32

Table 5

PAN data – 8 m, Germany. Squared J-M distance between basic LC classes

Land cover classes	Texture functions					
	PanBF	Laplace 3 × 3	Sobel	Sum_ Sigma	GLCM homo-geneity	GLCM entropy
agriculture/forest	1.42	1.33	0.39	1.17	1.49	1.25
forest/built-up	0.53	1.72	1.68	1.77	0.01	1.46
built-up/water	2.00	2.00	2.00	2.00	1.98	2.00
agriculture/built-up	1.56	1.99	1.93	1.99	1.46	1.80
forest/water	2.00	1.99	1.68	1.98	1.98	2.00
agriculture/water	1.26	1.93	1.38	1.91	0.41	0.78

Table 6

PAN data – 16 m, Germany. Squared J-M distance between basic LC classes

Land cover classes	Texture functions					
	PanBF	Laplace 3 × 3	Sobel	Sum_ Sigma	GLCM homo-geneity	GLCM entropy
agriculture/forest	0.70	0.33	0.09	0.24	1.08	0.58
forest/built-up	0.58	1.56	0.71	1.55	0.64	1.14
built-up/water	2.00	2.00	1.73	2.00	1.92	2.00
agriculture/built-up	1.27	1.85	0.50	1.81	1.48	1.54
forest/water	2.00	1.49	0.77	1.54	1.87	1.97
agriculture/water	0.50	1.20	0.36	1.18	0.50	0.30

Table 7

PAN data – 1 m, Spain. Squared J-M distance between basic LC classes

Land cover classes	Texture functions					
	PanBF	Laplace 3 × 3	Sobel	Sum_ Sigma	GLCM homo-geneity	GLCM entropy
agriculture/forest	1.93	1.97	1.95	1.97	1.89	1.72
forest/built-up	1.41	1.01	1.00	1.02	0.96	1.30
built-up/water	1.68	1.98	1.99	1.99	1.03	1.70
agriculture/built-up	0.86	1.84	1.79	1.84	0.59	0.79
forest/water	2.00	2.00	2.00	2.00	2.00	2.00
agriculture/water	1.70	2.00	2.00	2.00	1.30	1.49

Table 8

PAN data – 2 m, Spain. Squared J-M distance between basic LC classes

Land cover classes	Texture functions					
	PanBF	Laplace 3x3	Sobel	Sum_ Sigma	GLCM Homo- geneity	GLCM Entropy
agriculture/forest	1.86	1.97	1.81	1.96	1.89	1.73
forest/built-up	1.57	1.02	0.99	1.07	1.06	1.41
built-up/water	1.64	1.99	1.99	1.99	1.19	1.70
agriculture/built-up	0.92	1.83	1.78	1.84	0.69	0.83
forest/water	2.00	2.00	2.00	2.00	2.00	2.00
agriculture/water	1.15	2.00	2.00	2.00	1.20	1.35

Table 9

PAN data – 4 m, Spain. Squared J-M distance between basic LC classes

Land cover classes	Texture functions					
	PanBF	Laplace 3×3	Sobel	Sum_ Sigma	GLCM homo- geneity	GLCM entropy
agriculture/forest	**1.75**	**1.92**	1.36	**1.90**	**1.85**	1.63
forest/built-up	1.57	1.00	1.18	1.04	1.24	1.56
built-up/water	1.69	**1.99**	**2.00**	**1.99**	1.34	1.68
agriculture/built-up	1.02	**1.81**	**1.83**	**1.83**	0.83	0.87
forest/water	**2.00**	**2.00**	**2.00**	**2.00**	**2.00**	**2.00**
agriculture/water	0.96	**2.00**	**2.00**	**2.00**	0.90	0.91

Table 10

PAN data – 8 m, Spain. Squared J-M distance between basic LC classes

Land cover classes	Texture functions					
	PanBF	Laplace 3×3	Sobel	Sum_ Sigma	GLCM homo- geneity	GLCM entropy
agriculture/forest	1.53	**1.78**	1.09	**1.77**	1.66	1.51
forest/built-up	1.40	1.03	1.34	1.09	1.27	1.63
built-up/water	1.60	**1.99**	**2.00**	**2.00**	1.33	1.65
agriculture/built-up	1.04	**1.85**	1.80	**1.87**	0.89	1.06
forest/water	**2.00**	**2.00**	**2.00**	**2.00**	**2.00**	**2.00**
agriculture/water	0.66	**2.00**	1.99	**2.00**	0.66	0.48

Table 11

PAN data – 16 m, Spain. Squared J-M distance between basic LC classes

Land cover classes	Texture functions					
	PanBF	Laplace 3 × 3	Sobel	Sum_ Sigma	GLCM homo- geneity	GLCM entropy
agriculture/forest	0.87	1.62	0.45	1.58	0.86	1.12
forest/built-up	1.19	1.22	1.22	1.22	1.14	1.53
built-up/water	1.60	2.00	1.99	2.00	1.26	1.58
agriculture/built-up	0.82	1.91	1.47	1.89	0.73	1.22
forest/water	2.00	2.00	2.00	2.00	1.99	2.00
agriculture/water	0.38	2.00	1.85	2.00	0.41	0.57

Table 12

PAN data – 1 m. Squared J-M distance between the combined classes:
(i) "water" and "agriculture areas", and (ii) "forest" and "built-up areas"

Country	Texture functions					
	PanBF	Laplace 3 × 3	Sobel	Sum_ Sigma	GLCM homo- geneity	GLCM entropy
Poland	**1.90**	**1.92**	**1.92**	**1.92**	1.38	**1.83**
Germany	**1.74**	**1.92**	**1.96**	**1.92**	0.91	**1.72**
Finland	1.59	**1.91**	**1.95**	**1.91**	1.05	1.54
France	**1.84**	**1.82**	**1.79**	**1.81**	0.95	**1.73**
Spain	1.60	1.63	1.62	1.62	1.18	1.35

Table 13

PAN data – 2 m. Squared J-M distance between the combined classes:
(i) "water" and "agriculture areas", and (ii) "forest" and "built-up areas"

Country	Texture functions					
	PanBF	Laplace 3 × 3	Sobel	Sum_ Sigma	GLCM homo- geneity	GLCM entropy
Poland	**1.81**	**1.92**	**1.82**	**1.91**	1.44	**1.78**
Germany	1.65	**1.95**	**1.92**	**1.95**	1.10	**1.72**
Finland	1.57	**1.95**	**1.87**	**1.95**	1.15	1.54
France	**1.73**	**1.81**	1.68	**1.77**	1.06	1.68
Spain	1.61	1.64	1.48	1.60	1.06	1.38

Table 14

PAN data – 4 m. Squared J-M distance between the combined classes:
(i) "water" and "agriculture areas", and (ii) "forest" and "built-up areas"

Country	Texture functions					
	PanBF	Laplace 3 × 3	Sobel	Sum_ Sigma	GLCM homo-geneity	GLCM entropy
Poland	1.60	**1.84**	1.22	**1.79**	1.51	1.48
Germany	1.61	**1.89**	1.32	**1.84**	1.36	1.61
Finland	1.60	**1.80**	1.18	**1.77**	1.45	1.44
France	1.57	**1.75**	1.38	1.67	1.25	1.61
Spain	1.60	1.60	1.13	1.54	1.03	1.48

Table 15

PAN data – 8 m. Squared J-M distance between the combined classes:
(i) "water" and "agriculture areas", and (ii) "forest" and "built-up areas"

Country	Texture functions					
	PanBF	Laplace 3 × 3	Sobel	Sum_ Sigma	GLCM homo-geneity	GLCM entropy
Poland	1.27	1.21	0.25	1.09	1.34	0.76
Germany	1.40	1.14	0.57	1.03	1.47	1.23
Finland	1.40	1.03	0.54	0.97	1.59	1.23
France	1.32	1.65	1.18	1.56	1.25	1.56
Spain	1.45	1.45	0.93	1.43	1.10	1.57

Table 16

PAN data – 16 m. Squared J-M distance between the combined classes:
(i) "water" and "agriculture areas", and (ii) "forest" and "built-up areas"

Country	Texture functions					
	PanBF	Laplace 3 × 3	Sobel	Sum_ Sigma	GLCM homo-geneity	GLCM entropy
Poland	0.39	0.33	0.02	0.24	0.77	0.15
Germany	0.75	0.52	0.03	0.45	1.08	0.64
Finland	1.06	0.54	0.17	0.49	1.51	1.11
France	0.94	1.41	0.91	1.37	0.96	1.36
Spain	1.26	1.26	0.56	1.23	1.09	1.53

niques decreased the ability to distinguish land types. As before, in the case of the image from Spain, the division into low and high texture was impossible.

The results for the 4 m resolution are presented in Table 14. By far the best results were obtained for the Laplace filter followed by the Lee Sigma. In the case of France the J-M distance obtained from the Lee Sigma filter was below the accepted threshold of 1.7. As before, the Spanish image could not be analysed, and a further increase in pixel size to 8 m, followed by 16 m, resulted in a complete lack of distinction (see Tables 15 and 16, respectively).

The results of the analysis of the images from Poland, Germany, and Finland were similar although none were as good as the image from France. As we have highlighted, the worst results were obtained for Spain. Therefore, we argue that a classification approach based on low and high texture can be applied in Northern, Central, and Eastern Europe. However, this approach is not effective in Southern Europe because the land cover structure is different.

Figure 5 illustrates the distinction between land cover classes based on German test site images processed by a Laplacian 3×3 filter. The training sets are represented by the corresponding normal distributions. In the case of the 1 m resolution (Fig. 5a), the values for "water bodies" and "agricultural areas" completely overlap. The same is seen for "forest" and "built-up areas". We can therefore assume that a threshold between low and high texture can be defined at a value of about 45. In the case of images with a resolution of 8 m (Fig. 5b), such a division is not possible, because "agricultural areas" overlap with "forest" and (unlike at 1 m resolution) it is impossible to divide
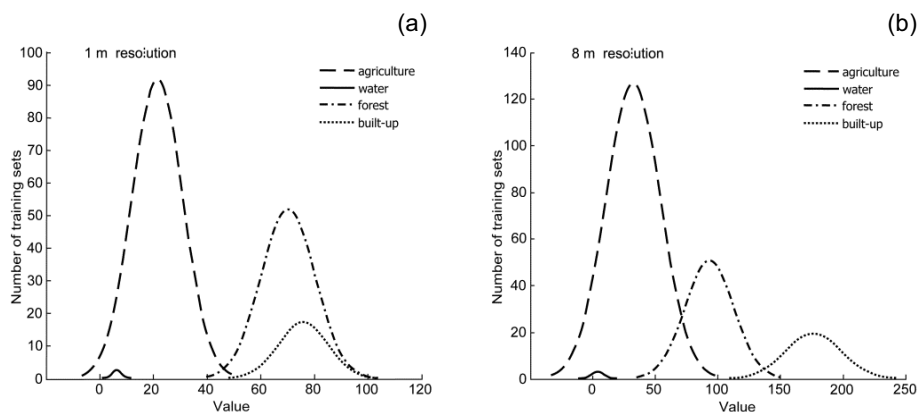
Fig. 5. Graphical analyses of the separation of land cover classes (German test site) at 1 m (a) and 8 m (b) resolutions.
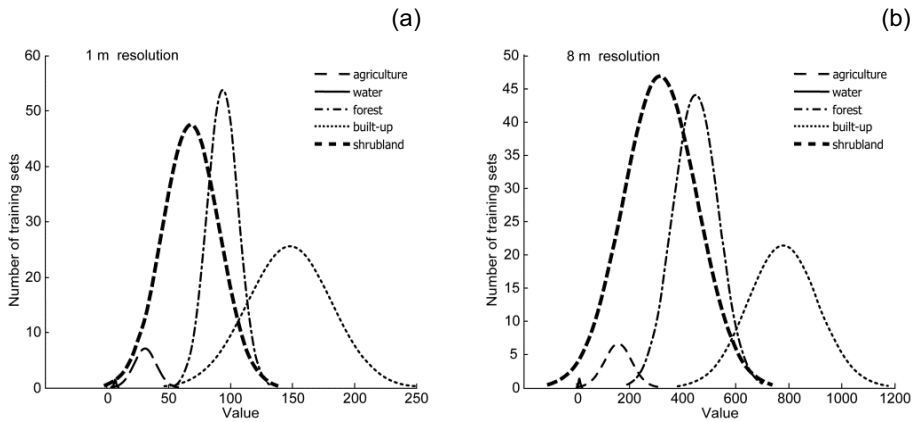
Fig. 6. Graphical analyses of the separation of land cover classes (Spanish test site) at 1 m (a) and 8 m (b) resolutions.

the image into two parts. Figure 6 shows the distribution of classes in Spain, and highlights the marked presence of the "shrubland" land cover class. This class occurs frequently in the South of Europe and must be taken into account in the classification process. "Shrubland" overlaps with "water bodies", "agricultural areas", "forest", and large parts of "built-up areas". The result is that an unequivocal division into low and high texture is impossible.

5. CONCLUSIONS

The basic division into low and high textures in the initial stage of the classification procedure enables a qualitative comparison of features. The feature space distances were calculated from the J-M distance between pairs of land cover classes: water bodies, agricultural areas, forest, built-up areas, and shrubland (dispersed vegetation). These classes were selected as the best representatives of low and high texture. An evaluation of the quality of results was performed at an early stage of the classification process, which helped to accurately trace the exact relationships between land cover classes.

Texture values depend on the function used, the spatial resolution of the image (pixel size) and the size of the land cover feature. We assumed that in the European context the size of land cover objects at the test sites were comparable. Any differences were primarily due to the density of vegetation and construction. Although similar results were obtained for the sites located in Poland, Germany, Finland, and France, the results from the Spanish test site were different.

Of the six texture feature functions we analysed, the best results were obtained with the Laplacian 3×3 filter, followed by the Sum of Lee Sigma and

PanBF functions. These texture functions made it possible to separate "water bodies" and "agricultural areas" from "forest" and "built-up areas". Much poorer results were obtained using GLCM homogeneity and entropy.

The classification approach based on low and high texture can be applied to panchromatic images with resolutions of 1, 2, and 4 m. The correctness of results decreases as the pixel size increases. Because of this, the classification should not be attempted on images with a resolution equal to or higher than 8 m. For such large pixels the threshold cannot be defined properly because the basic land cover classes cannot be properly distinguished.

The results from the Spanish site show that when land cover classes representing low and high texture overlap, it is practically impossible to divide them. This can be due to landscape features such as vegetation, topography, and climate. Even at a 1 m resolution, the distribution of "forest" that is not overlapped by "built-up areas" is misclassified as "agriculture". In addition to these areas, a highly dispersed "shrubland" class occurs. This class fully overlaps with both "agricultural areas" and "forest", which makes it difficult to perform a proper classification. Therefore, algorithms based on low and high texture can be used successfully mainly in Northern, Central, and Eastern Europe. Applying the technique to satellite images of Southern Europe is problematic if dispersed vegetation dominates, or if it affects the classification results.

R e f e r e n c e s

Berberoglu, S., P.J. Curran, C.D. Lloyd, and P.M. Atkinson (2007), Texture classification of Mediterranean land cover, *Int. J. Appl. Earth Observ. Geoinf.* **9**, 3, 322-334, DOI: 10.1016/ j.jag.2006.11.004.

Blaschke, T. (2010), Object based image analysis for remote sensing, *ISPRS J. Photogramm. Remote Sens.* **65**, 1, 2-16, DOI: 10.1016/j.isprsjprs.2009. 06.004.

De Kok, R. (2012), Spectral difference in the image domain for large neighborhoods, a GEOBIA pre-processing step for high resolution imagery, *Remote Sens.* **4**, 8, 2294-2313, DOI: 10.3390/rs4082294.

De Kok, R., and P. Wezyk (2008), Principles of full autonomy in image interpretation. The basic architectural design for a sequential process with image objects. **In:** Th. Blaschke, S. Lang, and G.J. Hay (eds.), *Object-Based Image Analysis*: *Spatial Concepts for Knowledge-Driven Remote Sensing Applications*, Lecture Notes in Geoinformation and Cartography, Springer, Berlin Heidelberg, 697-710, DOI: 10.1007/978-3-540-77058-9_38.

De Martinao, M., F. Causa, and S.B. Serpico (2003), Classification of optical high resolution images in urban environment using spectral and textural information. **In:** *Proc. IEEE Int. Symp. Geoscience and Remote Sensing IGARSS'03, 21-25 July 2003, Toulouse, France*, DOI: 10.1109/IGARSS. 2003.1293811.

Eckert, S. (2012), Improved forest biomass and carbon estimations using texture measures from WorldView-2 satellite data, *Remote Sens*. **4**, 4, 810-829, DOI: 10.3390/rs4040810.

eCognition Developer (2011), eCognition Developer, Reference book, Trimble documentation, Munich, Germany.

Esch, T., M. Thiel, A. Schenk, A. Roth, A. Muller, and S. Dech (2010), Delineation of urban footprints from TerraSAR-X data by analyzing speckle characteristics and intensity information, *IEEE Trans. Geosci. Remote Sens.* **48**, 2, 905-916, DOI: 10.1109/TGRS.2009.2037144.

Hall-Beyer, M. (2000), GLCM texture: a tutorial, Department of Geography, University of Calgary, Calgary, Canada, http://www.fp.ucalgary.ca/mhallbey/tutorial.htm.

Haralick, R.M., K. Shanmugan, and I. Dinstein (1973), Textural features for image classification, *IEEE Trans. Syst. Man Cybernetics* **SMC-3**, 6. 610-621, DOI: 10.1109/TSMC.1973.4309314.

He, C., J. Li, J. Zhang, Y. Pan, and Y. Chen (2005), Dynamic monitor on urban expansion based on a object-oriented approach. **In:** *Proc. IEEE Int. Symp. Geoscience and Remote Sensing IGARSS'05, 25-29 July 2005, Seoul, Korea*, 2850-2853, DOI: 10.1109/IGARSS.2005.1525662.

Hofmann, T., J. Puzicha, and J.M. Buhmann (1998), Unsupervised texture segmentation in a deterministic annealing framework, *IEEE Trans. Pattern Anal. Mach. Intellig.* **20**, 8, 803-818, DOI: 10.1109/34.709593.

Hu, X., C.V. Tao, and B. Prenzel (2005), Automatic segmentation of high-resolution satellite imagery by integrating texture, intensity, and colour features, *Photogramm Eng. Rem. Sens.* **71**, 12, 1399-1406, DOI: 10.14358/PERS. 71.12. 1399.

Intergraph (1999), Image Analyst User's Guide for Windows, Intergraph Corporation, Huntsville, USA.

Jain, A.K. (1989), *Fundamentals of Digital Image Processing*, Prentice-Hall, Englewood Cliffs, 569 pp.

Jensen, J.R. (1996), *Introductory Digital Image Processing: A Remote Sensing Perspective*, 2nd ed., Prentice Hall, Upper Saddle River, 316 pp.

Kit, O., M. Lüdeke, and D. Reckien (2012), Texture-based identification of urban slums in Hyderabad, India using remote sensing data, *Appl. Geogr.* **32**, 2, 660-667, DOI: 10.1016/j.apgeog.2011.07.016.

Lewinski, S., and Z. Bochenek (2009), Rule-based classification of SPOT imagery using object-oriented approach for detailed land cover mapping. **In:** *Proc. 28th EARSeL Symp. "Remote Sensing for a Changing Europe", 2-5 June 2008, Istanbul, Turkey*.

Lewinski, S., Z. Bochenek, and K. Turlej (2014), Application of an object-oriented method for classification of VHR satellite images using rule-based approach and texture measures. **In:** I. Manakos and M. Braun (eds.), *Land Use and Land Cover Mapping in Europe: Practices and Trends*, Remote Sensing and Digital Image Processing, Vol. 18, 193-201, Springer Science + Business Media, Dordrecht, DOI: 10.1007/978-94-007-7969-3_12.

Morales, D.I., M. Moctezuma, and F. Parmiggiani (2003), Urban and non urban area classification by texture characteristics and data fusion. **In:** *Proc. IEEE Int. Symp. Geoscience and Remote Sensing IGARSS'03, 21-25 July 2003, Toulouse, France*, 3504-3506, DOI: 10.1109/IGARSS.2003.1294835.

Murray, H., A. Lucieer, and R. Williams (2010), Texture-based classification of sub-Antarctic vegetation communities on Heard Island, *Int. J. Appl. Earth Observ. Geoinform.* **12**, 3, 138-149, DOI: 10.1016/j.jag.2010.01.006.

Nussbaum, S., I. Niemeyer, and M.J. Canty (2006), SEaTH – A new tool for automated feature extraction in the context of object-based image analysis. **In:** *Proc. 1st Int. Conf. on Object-based Image Analysis (OBIA), 4-5 July 2006, Salzburg, Austria*.

Puissant, A., J. Hirsch, and C. Weber (2005), The utility of texture analysis to improve per-pixel classification for high to very high spatial resolution imagery, *Int. J. Remote Sens.* **26**, 4, 733-745, DOI: 10.1080/01431160512331316838.

Ryherd, S., and C. Woodcock (1996), Combining spectral and texture data in the segmentation of remotely sensed images, *Photogramm. Eng. Remote Sens.* **62**, 2, 181-194.

Su, W., J. Li, Y. Chen, Z. Liu, J. Zhang, T.M. Low, I. Suppiah, and S.A.M. Hashim (2008), Textural and local spatial statistics for the object-oriented classifi-

cation of urban areas using high resolution imagery, *Int. J. Remote Sens*. **29**, 11, 3105-3117, DOI: 10.1080/01431160701469016.

Thomas, V., P. Treitz, D. Jelinski, J. Miller, P. Lafleur, and J.H. McCaughey (2003), Image classification of a northern peatland complex using spectral and plant community data, *Remote Sens Environ*. **84**, 1, 83-99, DOI: 10.1016/S0034-4257(02)00099-8.

Tuceryan, M., and A.K. Jain (1999), Texture analysis. **In:** C.H. Chen, L.F. Pau, and P.S.P. Wang (eds.), *The Handbook of Pattern Recognition and Computer Vision*, 2nd ed., World Scientific Publ. Co., Singapore, 207-248.

Wang, Y.W., Y.F. Wang, Y. Xue, and W. Gao (2003), A new algorithm for remotely sensed image texture classification and segmentation. **In:** *Proc. IEEE Int. Symp. Geoscience and Remote Sensing IGARSS'03, 21-25 July 2003, Toulouse, France*, 3534-3536, DOI: 10.1109/IGARSS.2003.1294845.

Wezyk, P., and R. De Kok (2006), Automatic mapping of the dynamics of forest succession on abandoned parcels in south Poland. **In:** J. Strobl, T. Blaschke, and G. Griesebner (eds.), *Angewandte Geoinformatik 2005*, Wichman Verlag, Heidelberg, 774-779 (in German).

DE GRUYTER
OPEN

Estimation of Apollo Lunar Dust Transport using Optical Extinction Measurements

John E. LANE[1]   and   Philip T. METZGER[2,3]

[1]Easi-ESC, Granular Mechanics and Regolith Operations, Kennedy Space Center, FL, USA; e-mail: john.e.lane@nasa.gov (corresponding author)

[2]NASA Granular Mechanics and Regolith Operations, Kennedy Space Center, FL, USA

[3]Florida Space Institute, University of Central Florida, Orlando, FL, USA e-mail: Philip.Metzger@ucf.edu

A b s t r a c t

A technique to estimate mass erosion rate of surface soil during landing of the Apollo Lunar Module (LM) and total mass ejected due to the rocket plume interaction is proposed and tested. The erosion rate is proportional to the product of the second moment of the lofted particle size distribution $N(D)$, and third moment of the normalized soil size distribution $S(D)$, divided by the integral of $S(D) \cdot D^2 / v(D)$, where D is particle diameter and $v(D)$ is the vertical component of particle velocity. The second moment of $N(D)$ is estimated by optical extinction analysis of the Apollo cockpit video. Because of the similarity between mass erosion rate of soil as measured by optical extinction and rainfall rate as measured by radar reflectivity, traditional NWS radar/rainfall correlation methodology can be applied to the lunar soil case where various $S(D)$ models are assumed corresponding to specific lunar sites.

Key words: Mie scattering, efficiency factor for extinction, particle size distribution, mass erosion rate, shear stress, shape factor.

1. INTRODUCTION

Previous work has focused on particle trajectory analysis and computational fluid dynamics (CFD) simulations of rocket plume interactions with the lunar surface (Morris *et al*. 2011, Immer *et al*. 2011a, Lane *et al*. 2010). An important component that may be missing from a pure trajectory simulation is intensity of dust dispersal, or more precisely, the "soil mass erosion rate". In previous work, an erosion rate was estimated from the optical extinction of a few ideal image features (Immer *et al*. 2011b, Metzger *et al*. 2010). The drawback of those methods is the limited data that is available for analysis, usually only a few frames from an entire landing video. Recently a new approach was taken, following the methodology used by the National Weather Service (NWS) in measuring rainfall intensity (hydrometeor mass accumulation and intensity rate) using Weather Surveillance Radar (Wexler and Atlas 1963, Rosenfeld *et al.* 1993).

The key to this approach is to assume a particle size distribution. Even though it may appear to be a risky assumption, this methodology has been in operational use by the NWS since the advent of weather radar. Part of this methodology relies on a strategy of substituting appropriate drop size distribution (DSD) functions for specific meteorological regimes (tropical, continental, *etc.*). This is then similar to substituting different soil size distributions for various areas on the lunar surface where surface operations and landings are planned, such as highlands, mare, and permanently shadowed craters (highland and mare samples were returned by Apollo missions, but return of permanently shadowed craters samples will be a goal of some future missions). Once the erosion rate is determined for a specific rocket engine, it can then be used to correlate those predictions with CFD simulations which also predict surface shear stress due to the rocket plume interaction with the surface (Metzger *et al*. 2011).Note that CFD simulations are needed to navigate from engine design specifications to predictions of surface shear stress predictions and dispersed particle velocities.

Another result of this study is that it becomes obvious that optical extinction due to scattering of light from hydrometeors can be used to estimate rainfall rate, just as microwave radar may be used to measure soil erosion rate (Lane *et al*. 2014b). This resemblance is a consequence of the similarity of size range of the particle distributions of hydrometeors and lunar soil and the fact that the index of refraction which determines the details of electromagnetic scattering is similar. Before this equivalence is taken too far however, it must be recognized that 10 cm weather radar does not detect fog size particles, which are comparable in size to the smallest lunar dust particles that may contribute to soil erosion. Therefore, to measure dust particles effectively, millimetre wave radar would need to be utilized to correctly quantify this analogy.

2. MASS EROSION RATE

Measuring lunar soil erosion rate $\dot{m}$ from optical extinction α during a rocket landing is analogous to measuring terrestrial rainfall rate R using NWS radar reflectivity Z. The similarity in both cases is the dependence on some knowledge of the particle size distribution functions, $N(D)$. For weather radar, the quantity of interest R is usually reported in $mm\,h^{-1}$. For soil erosion, the quantity of interest $\dot{m}$ is measured in units of $kg\,s^{-1}\,m^{-2}$. Note that the product of rainfall rate and water density also has units of $kg\,s^{-1}\,m^{-2}$.

The weather radar Z-R relation is a simple power-law and correlates measured reflectivity Z and rainfall rate R using two parameters, a and b (see Fig. 1). Radar reflectivity Z is the sixth moment of the DSD. By parameterizing and fitting $N(D)$ so that its integral is a Gamma Function, $\Gamma(D)$, equations for Z and R can be combined to form a power law of the form:

$$Z = a\,R^b \ , \tag{1}$$

where a and b are related to the parameters that describe the drop size distribution and drop terminal velocity functions. For operational use, these parameters are determined empirically from weather radar data and networks of rain gauges. Note that the size distribution curves shown in the right hand side of Figs. 1 and 2 are for illustration only and do not depict actual size distributions.

Similarly, the soil erosion rate $\dot{m}$ is found by integrating the product of the particle mass and velocity times the size distribution $N(D)$. The optical extinction is found from the second moment of the size distribution (Atlas 1953). Shipley *et al.* (1974) demonstrated empirically that optical extinction has a power law relationship with the rainfall rate and therefore with the suspended hydrometeor mass when the drop size parameter $x \equiv \pi D/\lambda \gg 1$, where λ is the illumination wavelength. This condition applies to particles such as lunar dust when $D \gg \lambda/\pi$.

$R = \dot{m}/\rho$, $\dot{m} =$ (water) mass accumulation rate, $\rho =$ density of water

radar reflectivity, Z

rainfall rate, R

$Z = a\,R^b$ $R = \left(a^{-1}Z\right)^{1/b}$

Rosenfeld Tropical Convective Systems ($a = 250, b = 1.2$)

WSR-88D Convective Summer Deep Convection / Other Non-tropical ($a = 300, b = 1.4$)

Marshall-Palmer General Stratiform precipitation ($a = 200, b = 1.6$)
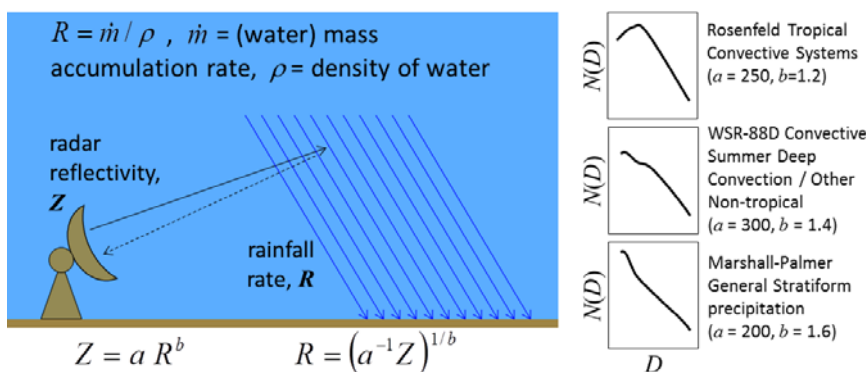
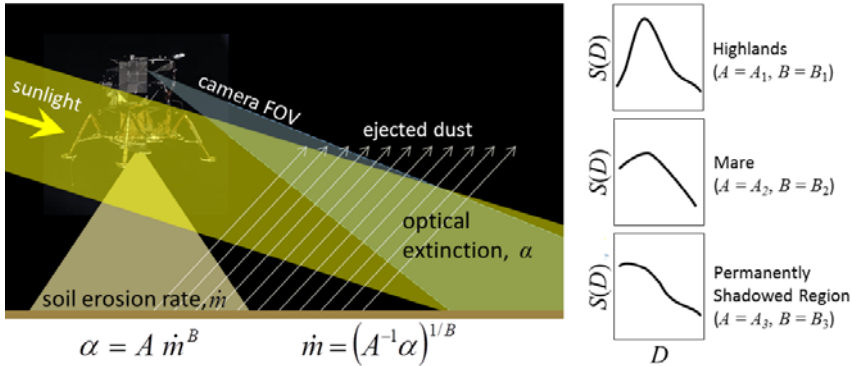Fig. 1. Radar measurement of rainfall rate is affected by the local DSD, $N(D)$.

Fig. 2. Optical measurement of dust erosion rate is affected by the local soil size distribution $S(D)$. Note that A and B are found from $S(D)$ associated with soil properties at a specific region on the lunar surface, as well as the properties of the gas exiting the rocket nozzle used to generate $v(D)$, based on the specific engine design.

At this point the similarity diverges since the size distribution $N(D)$ is the lofted distribution of particles and in the lunar soil case is an unknown. Rainfall DSDs can be measured directly using ground based or aircraft based disdrometers. However, the normalized soil size distribution $S(D)$ can be measured (using lunar samples returned to Earth) and is related to $N(D)$ and the CFD simulated particle velocity $v(D)$ according to:

$$S(D) = \frac{v(D)\,N(D)}{\int_0^\infty v(D)\,N(D)\,dD} \, . \tag{2}$$

Now the relationship between optical extinction α and mass erosion rate $\dot{m}$ can be approximated by a power-law, analogous to the hydrometeor case of Eq. 1:

$$\alpha = A\dot{m}^B \, , \tag{3}$$

where the parameters A and B are found using the soil size distribution $S(D)$ associated with soil properties at a specific region on the lunar surface, as well as the properties of the gas exiting the rocket nozzle used to generate $v(D)$, based on the specific engine design.

Calculation of soil mass erosion from optical extinction can be approximated using the following equation:

$$\dot{m}(t) = \frac{\rho_L \pi \phi}{6s(f)} M_2(t) \; \frac{\int_0^\infty S(D)D^3\,dD}{\int_0^\infty S(D)\left(D^2/v(D)\right)dD} \, , \tag{4}$$

where $M_2(t)$ is the second moment of the lofted size distribution, which is indirectly measured by video camera analysis. A DSD moment is defined as:

$$M_x = \int_0^\infty D^x N(D)dD \quad .$$

(5)

The bulk density of lunar soil ρ_L is approximated as 3100 kg m^{-3}. $S(D)$ is the normalized soil size distribution as measured by an image analysis based particle size analyzer, particle sieves, or some equivalent method. The parameter φ is a geometry factor accounting for the divergence of the dust ejecta, spreading radially outwards from the nozzle centerline. It can be shown that $\varphi \approx 2$ using a simple model of the dust ejection pattern (see Appendix A). For the case of zero divergence, such as terrestrial rainfall, $\varphi = 1$.

The particle shape factor $s(f)$ is a function of the particle aspect ratio $f = r_b/r_a$, where r_a is the short radius and r_b is the long radius. In the idealized case of the "prolate spheroid", all quantities involving D are computed as usual with a diameter $D = (r_a^2 r_b)^{1/3} = r_a f^{1/3}$. The shape factor for particles modeled as a prolate spheroid with aspect ratio f is $s(f) = (\pi + 2(f-1))/(\pi f^{2/3})$ (see Appendix B).

The optical extinction factor α is related to the second moment as $\alpha = (\pi/4) Q_e M_2$, where Q_e is the scattering efficiency factor for extinction (Berg $et\ al.$ 2011). In general, Q_e is a function of the size parameter x and, for a narrowband fixed spectrum of light, it can be approximated as a function of only particle size D. As shown by van de Hulst (1957), the minimum size factor x that determines the boundary between classical and Mie scattering is a function of refractive index of the scattering particle. The larger the refractive index n, the smaller the threshold value of x. In the case of hydrometeor scatterers, $n = 1.33$, so classical scattering with $Q_e = 2$ applies to drop sizes for $x > 6$, and for a spectrum centered about green light ($\lambda_G = 532$ nm), $D > 6\lambda_G/\pi$. Therefore, visible light can be used to measure the extinction factor of hydrometeors greater than 1 µm using a constant $Q_e = 2$. In this case, the second moment of the size distribution is related to the extinction factor α by a factor of $2/\pi$. The same is true in the case of lunar dust particles, with the exception of a larger index of refraction. In the lunar dust case with $n = 1.75$ (ignoring the small imaginary component), Fig. 24 of van de Hulst (1957) shows $x > 3$, so that $Q_e = 2$ can be used down to a particle size of 0.5 µm, again using the visible light spectrum centered about λ_G.

The soil size distribution $S(D)$ for Apollo 11 sample 10084 and Apollo 17 sample 70051 are shown to have a peak around 0.030 µm (Metzger $et\ al.$ 2010). This would seem to violate the assumption $x > 3$ and $Q_e = 2$. However, it is not the peak of $S(D)$ that is relevant to the issue of Q_e, it is the

moments of the soil size distributions as expressed by the numerator and denominator of the right side of Eq. 4. It can be shown that the peaks occur around $D = 30$ μm, well within the assumption that $x > 3$ and $Q_e = 2$.

2.1 Extinction factor and effective radius of erosion

Figure 3 shows the time dependence of the LM cockpit video frame sequence for the final 60 s of descent and for 60 s after landing and engine cutoff. The dotted line represents the video frame number at 12 fps rate, while the line with open circles plots the LM height according to the voice recording of altitude radar callout. The upper black line is the histogram average of each video frame, while the lower green line is the corresponding standard deviation. The histogram plots of Figs. 18 and 19 (Appendix C) show a bimodal characteristic. Crater shadows and a low sun illumination angle produce the low end histogram peak (dark shadows) while the high end peak is due to everything else. Dust in the image has the effect of forcing these two peaks together. This process is revealed primarily by a reduction of the standard deviation, caused by the peaks converging as the dust cloud density increases. Region A of Fig. 3 can be used as a reference since it is a clear image of a typical surface scene. Other regions are described in Table 1.

Reiterating from the last section, $\alpha = \pi Q_e M_2/4$, where Q_e is the scattering efficiency, assumed to be equal to 2. The second moment in Eq. 4 is then the product of the measured extinction factor and $2/\pi$. The extinction factor α can be estimated from the Apollo videos by adding dust to a clear reference image (before dust appears) and comparing to the dusty image of interest. By matching histograms of the two images, the extinction factor can be estimated. The details of the histogram matching method (HMM) (Lane and
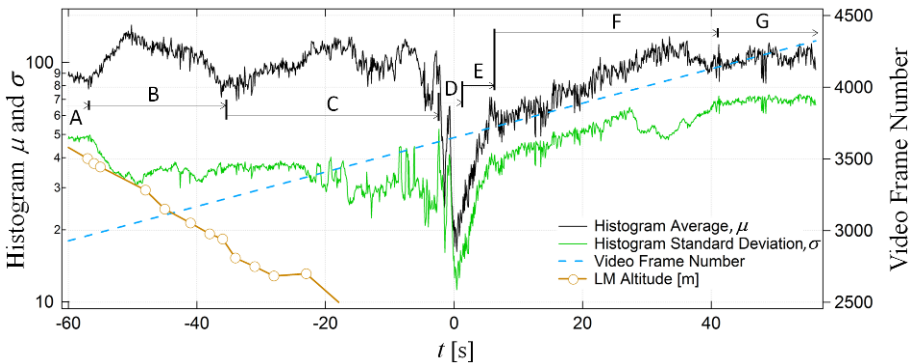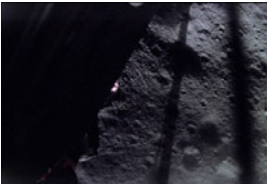

Fig. 3. Histogram parameters for the final 60 s of descent and for 60 s landing and after engine shutoff.

Table 1

Description of regions in Figure 3

	A: Frame 2962. Images in this region can be used as a histogram reference. The variation of the histogram average and standard deviation is minimal during this segment of the video. Frames showing unusually large dark craters need to be excluded from the reference baseline.
	B: Frame 3111. Images in this region of the video are good candidates for the histogram matching method (HMM), described in Appendix C.
	C: Frame 3404. Images in this region are not suitable for the HMM. Manual selection of extinction parameters is done in this region by trial-and-error, comparing the dust treated image visually.
	D: Frame 3647. Images in this region experience a total blackout due to LM shadows on the top of the dust cloud. As the dust cloud settles the shadow quickly disappears. This region is not useable for extracting extinction information due to the effect of the LM shadows.
	E: Frame 3713. As the dust cloud height decreases, the LM shadows fade. A smaller amount of dust persists for at least another 30 s.
	F: Frame 4004. The dust in this region is clearing at a slow rate, indicative of levitation due to effects such as electrostatic repulsion or escaping rocket exhaust gas previously forced into the regolith. The mechanism of the dust levitation is an area of current research.
	G: Frame 4311. The dust has cleared in this region. Variations of the histogram average or standard deviation are due to camera noise and/or noise introduced during the image digitization process.

Metzger 2014a) are discussed in Appendix C. HMM is more than just adjusting contrast and brightness of the images. The dust erosion angle ($\theta \approx 3°$ for Apollo 12) and radius of erosion is used to apply different amounts of contrast-brightness equalization to strips across the image, rotated to align with the horizon (and the dust cloud top). By iteratively adjusting the optical extinction factor α and radius of erosion a_0 and comparing the histogram averages and standard deviations, a best fit a_0 and α are found. It may be feasible to estimate a dust erosion angle θ using HMM, an area of possible future work.

The total mass ejected (total mass displaced) is Eq. 4 integrated over vehicle descent time and over the area where soil is eroded:

$$m_T = 2\pi \int\limits_{-\infty}^{0} \int\limits_{0}^{a_0(t)} \dot{m}(t) \, r dr dt = \pi \int\limits_{-\infty}^{0} a_0^2(t) \, \dot{m}(t) \, dt \; , \tag{6}$$

where $a_0(t)$ is the radius on the surface, referenced to the engine nozzle centerline, where erosion is taking place. The assumption inherent in Eq. 6 is that erosion is uniform over a circle of radius $a_0(t)$ and zero outside of that circle. An estimate of $a_0(t)$ for the Apollo 12 LM is shown in Fig. 4, which is an output of the HMM algorithm. Note that in this and all previous discussions of $S(D)$ and its moments, it has been assumed that the particle size distribution is homogeneous over the extent of measurement, *i.e.*, within a circle of radius $a_0(t)$, and all temporal effects due to engine LM altitude and thrust occur instantaneously over this spatial extent.
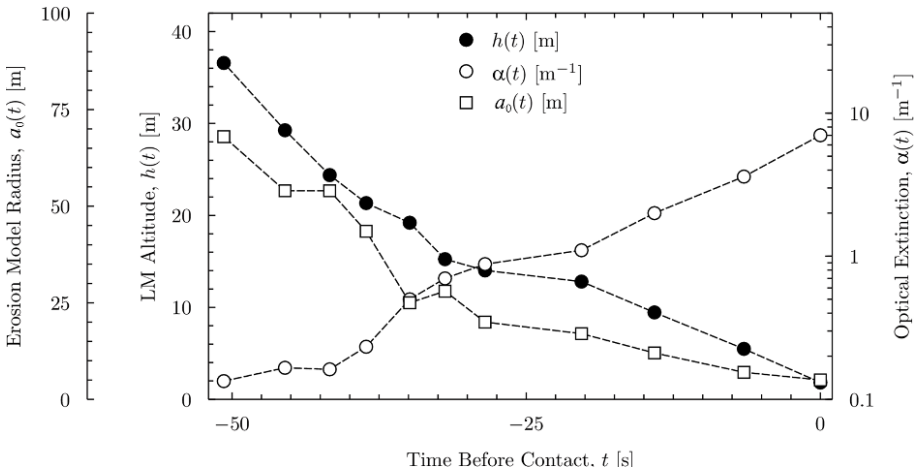

Fig. 4. Apollo 12 optical extinction estimate using histogram matching method. The time intervals correspond to the voice callouts of LM pilot, Alan Bean.

The erosion rate should actually vary with the state of the gas flowing across the soil, including its shear stress, rarefaction, and turbulence; it should also vary with saltation, including the downward flux of larger particles that are too heavy to be carried away by the gas as well as smaller particles that are scattered back down from the entrained cloud via particle collisions. Examination of the sandblasting effects on Surveyor III has shown that the downward flux of scattered particles is significant (Immer *et al*. 2011a) and discrete element computer simulations show the important but largely unexplored role of mid-flight particle scattering in enhancing erosion rate (Berger *et al*. 2013). Influence of the gas upon erosion rate should be greatest in an annular region around the vehicle (Roberts 1963) while the influence of saltation may be greater then another annulus with larger radius since particles travel downrange before striking the surface. Thus the net erosion rate may be somewhat more uniform and spread over a broader region than if gas effects alone are considered. The details of erosion physics, especially in lunar rocket exhaust conditions, are not yet well understood, so a constant erosion over a finite area assumed, as it is the simplest model and therefore a sensible first step.

2.2 Particle velocity function

Single particle trajectory modelling, based on CFD simulations of the Apollo LM engine and the lunar environment (Lane *et al*. 2010), yield a particle velocity function which can be described by the empirical fit shown in Eq. 7, analogous to the hydrometeor terminal velocity formulas used in meteorology:

$$u(h,D) = \begin{cases} 0 & \log D + c_1 \log h > c_0 \\ \dfrac{10^{(b_1 + b_2 \log h)\tanh(b_3 + b_4 \log h)}}{b_0\left(h^{9/20} + D^{1/2}\right)} & \text{otherwise} \end{cases}, \qquad (7)$$

where $b_0 = 0.2964$, $b_1 = -0.225$, $b_2 = 0.1954$, $b_3 = 5$, $b_4 = 4.343$, $c_0 = 2.212$, and $c_1 = 3.53$ (h and D in meters). Equation 7 is plotted in Fig. 5. For every point in $\{h, D\}$ space, a distribution of particle velocities is computed from the particle trajectory code using an equivalent Monte Carlo distribution of initial particle trajectory starting points, height above the surface and horizontal distance from the engine nozzle centerline.

The fit given by of Eq. 7 and Fig. 5 represents a maximum value of particle velocities, where particles originate near the outside rim of the rocket nozzle. The area in the upper left of Fig. 5 represents the region of $\{h, D\}$ where particles do not lift from the surface due to an insufficient lifting force. Since $u(h, D)$ is a maximum velocity, the velocity in Eq. 4 at each
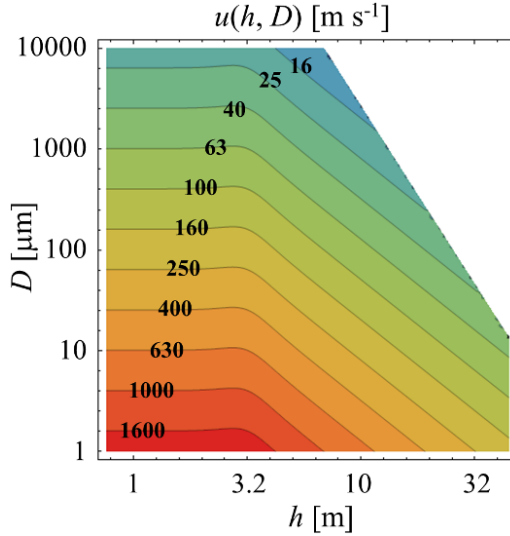
Fig. 5. Maximum particle velocity, $u(h, D)$ [m s^{-1}].

time step is replaced with a reduced value of Eq. 7, $v(D) = \eta \sin\theta\, u(h, D)$, where θ is the plume propelled dust angle, equal to approximately 3 degrees (Immer *et al.* 2011b), where $\eta \leq 1$.

Because of the distribution of particle velocities for a given h and D, the maximum values originate near the engine nozzle. Since the erosion area within $r < a_0$ (where r is the radial distance from the engine nozzle centerline) is much greater for slower velocities, then it is reasonable to expect η to be much smaller than 1. The curve fits in Fig. 6 are of the form $V(r) = v_0 + v_1 \exp(-r/\xi)$, which to first order are assumed to be independent of h and D. The parameter η can be estimated from the weighted area integral of the particle speed profile (see Fig. 6) as a function of r:

$$\eta = \frac{2\pi \int_{R_1}^{R_2} V(r)\, r\, dr}{2\pi\, V(R_1) \int_{R_1}^{R_2} r\, dr}\ , \tag{8}$$

where $R_1 = 0.9$ m and $R_2 = 8.0$ m are limits of the CFD particle trajectory (CFD-PT) simulations. Note that the data points along the horizontal axis are the specific values of r used in the CFD-PT simulations. Performing this integral for the two curve fits in Fig. 6 results in $\eta = 0.147$ for the upper curve and $\eta = 0.130$ for the lower curve. Since the value of η does not change significantly for D or h, a value of $\eta = 1/8$ is used as an approximation throughout the descent and erosion analysis.
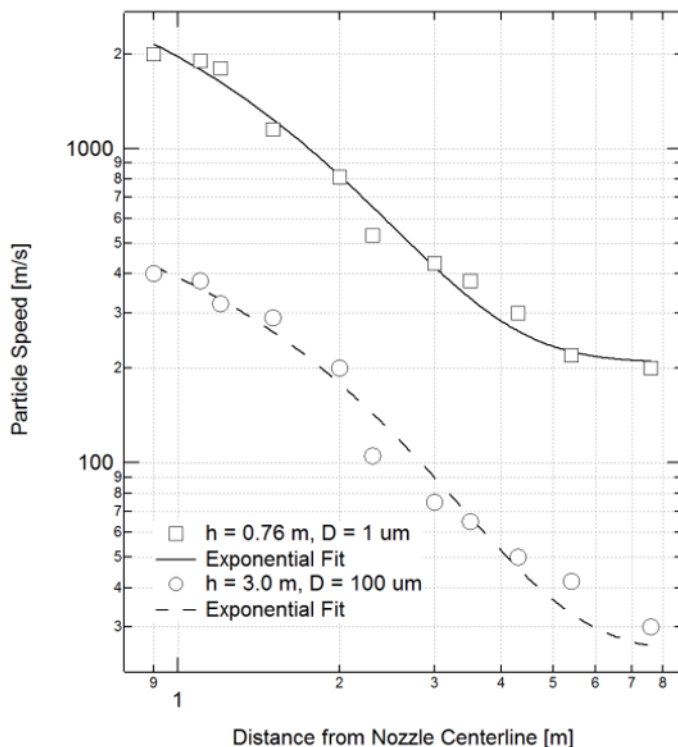
Fig. 6. Particle speed as a function of its radial starting distance from the nozzle center of the Apollo LM descent engine, for two example values of h and D.

2.3 Particle size distribution

The particle size fraction of the lunar soil at the Apollo 12 site is modeled as a combination of two power law functions by fitting Apollo 11 and 17 soil sample data, as well as JSC-1a simulant (Metzger *et al.* 2010):

$$S(D) = \frac{\left(w(D)S_1(D) + S_2(D)\right)^{-1}}{\int_0^\infty \left(w(D)S_1(D) + S_2(D)\right)^{-1} dD} \quad [\text{m}^{-1}] , \qquad (9)$$

where: $S_1(D) = (D/D_1)^{B_1}$, $S_2(D) = (D/D_2)^{B_2}$, and $w(D) = (D_3/D)^{B_3} + 1$. The fitting constants in Eq. 9 are: $D_1 = 4.090 \times 10^{-7}$ m, $B_1 = 1.8$, $D_2 = 9.507 \times 10^{-6}$ m, $B_2 = 5.6$, $D_3 = 2.5 \times 10^{-8}$ m, and $B_3 = 18$. Note that all units are kept in *si* units even though the numbers are more aesthetically pleasing in micrometers. The reason for doing this is to minimize confusion in the integrals involving $S(D)$.

2.4 A justification for Eq. 4

Rather than derive Eq. 4 directly, it can be worked in reverse to yield a familiar result in the meteorological case. If this is shown to be true for hydrometeors, then it follows that a particle distribution composed of granular material and dust should follow similar rules of behavior under similar forces.

Rainfall rate R m s^{-1} is equal to $\dot{m}/\rho$, where ρ is the density of water. With this substitution, and substitution of Eq. 2 for $S(D)$, Eq. 4 becomes:

$$R(t) = \frac{\pi\phi}{6} M_2(t) \frac{\int_0^\infty v(D)\, N(D) D^3\, dD}{\int_0^\infty v(D)\, N(D)\left(D^2/v(D)\right) dD} = \frac{\pi\phi}{6} \int_0^\infty v(D)\, N(D) D^3\, dD. \quad (10)$$

For $\phi = 1$, Eq. 10 is a familiar result for computing rainfall rate in terms of the drop size distribution.

3. INTREPID EROSION ANALYSIS

Equation 4 computes mass erosion at each time step, with the data from Figs. 4, 5, and 7, using $\eta = 1/8$, $\phi = 2$, and $f = 1$. Table 2 summarizes these results. The sum of the eroded soil at each time step in the right column of Table 2 yields the total mass eroded. As can be seen from this table, the majority of the mass weighted erosion takes place in the last 20 s. The total eroded mass, using the parameter values chosen, equals 2594 kg. Table 3 compares the present result with previous work.
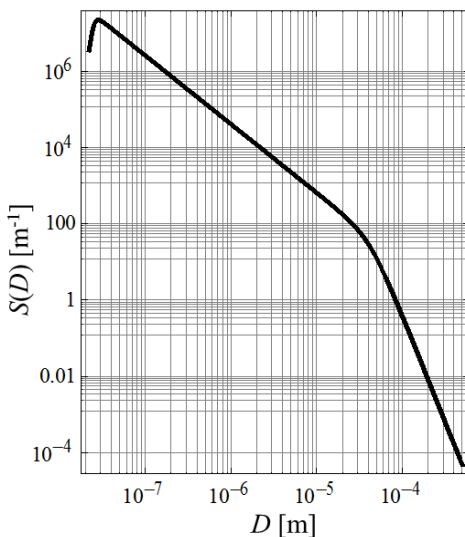

Fig. 7. Lunar regolith particle size fraction $S(D)$, estimated for Apollo 12 site using fit from Eq. 9, based on samples 10084 (Apollo 11) and 70051 (Apollo 17).

Table 2

The kth time step corresponds to the cockpit voice recording of altitude

t_k [s]	h_k [m]	M_{2k} [m^{-1}]	a_{0k} [m]	$\dot{m}_k$ [kg s^{-1} m^{-2}]	Δm_k [kg] from Eq. 11
0	1.83	8.91	5.0	1.26	644.4260
−6.5	5.49	4.58	7.0	0.395	462.3470
−14.1	9.45	2.55	12.0	0.133	373.4560
−20.3	12.80	1.40	17.0	0.0542	403.7720
−28.5	14.00	1.12	20.0	0.0392	167.5450
−31.9	15.20	0.891	28.0	0.0282	208.2450
−34.9	19.20	0.637	25.0	0.0140	101.8150
−38.6	21.30	0.297	43.5	0.00521	96.0130
−41.7	24.40	0.206	54.0	0.00260	90.4892
−45.5	29.30	0.211	54.0	0.00165	78.4725
−50.7	36.60	0.171	68.0	0.000734	55.4542

Table 3

Comparison of total mass erosion estimates for Apollo landings

Reference	Apollo mission	Total mass erosion [kg]
Scott (1975)*	Apollo 12	4500 to 6400
Metzger et al. (2008)	Apollo 12	2400
Metzger et al. (2010)	Apollo average	1200
present work: $m_T = \sum_{k=-10}^{0} \Delta m_k$	Apollo 12	2600

*$^{)}$from Metzger et al. (2011) – interpretations of data reported by Scott (1975)

The total regolith transported from its initial resting position by erosion induced by the Apollo LM rocket engine, is estimated by integrating the mass erosion rates from Table 2 over surface area and time. Note that $t = 0$ is the surface "contact time" when the LM is approximately 1.5 m above the surface and the descent engine is turned off. Based on fall time in lunar gravity, the LM continues to descend for up to an additional 1.3 s. The index k corresponding to entries in Table 2 ascend from bottom to top. The total eroded mass can be approximated by linear interpolation of $\dot{m}_k$ and a_{0k} at each kth point:

$$m_T = \pi \int_{-10}^{0} \dot{m}(t)\, a_0^2(t)\; dt \approx \sum_{k=-10}^{0} \Delta m_k$$

$$\approx \frac{\pi}{12} \sum_{k=-10}^{0} \left(\left(3\dot{m}_k + \dot{m}_{k+1}\right) a_{0k}^2 + 2\left(\dot{m}_k + \dot{m}_{k+1}\right) a_{0k} a_{0k+1} + \left(\dot{m}_k + 3\dot{m}_{k+1}\right) a_{0k+1}^2 \right)\left(t_{k+1} - t_k\right).$$

(11)

The computed value of $m_T \approx 2600$ kg, and as shown in Table 3, is well within the range of other estimates of Apollo 12 total mass erosion.

The LM landing profile shown in Fig. 8 can be compared to Table 2 and Fig. 3. The appearance of dust and erosion begins at an altitude of about 40 m, which is approximately 65 m from the landing site. Figure 9 shows a view of the landing site towards the south with Surveyor Crater to the left and possible surface scouring due to plume interaction, just to the left of the engine nozzle. The area of possible soil removal is shown as a discolored region, slightly browner than the greyer regolith surrounding it. Using simple image scaling, it is possible to roughly estimate the crater contour that is highlighted by the discoloration.

Figure 10a shows the results of a crude photogrammetry analysis of the scouring depth. The offset of the deep part of the crater is unusual and may indicate a burst of thrust just before touchdown, or a pre-existing depression in the surface. The lack of a large dug-out directly beneath the engine nozzle could be explained by the combination of a small horizontal velocity, a slight LM tilt, and engine shutoff at 1.5 m altitude (point of "contact"). For comparison, the crater from the Table 2 data is shown in the Fig. 10b. The erosion picture from the optical extinction model is much shallower and greater in extent than the photogrammetry derived crater analysis.
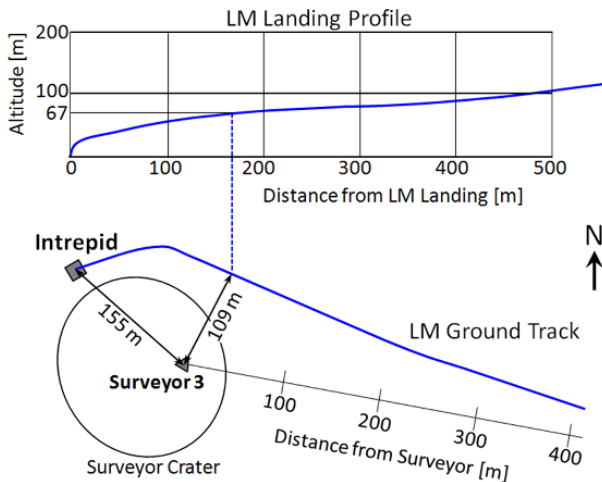

Fig. 8. Landing profile of Apollo 12 LM, Intrepid, showing Surveyor 3 landing site.
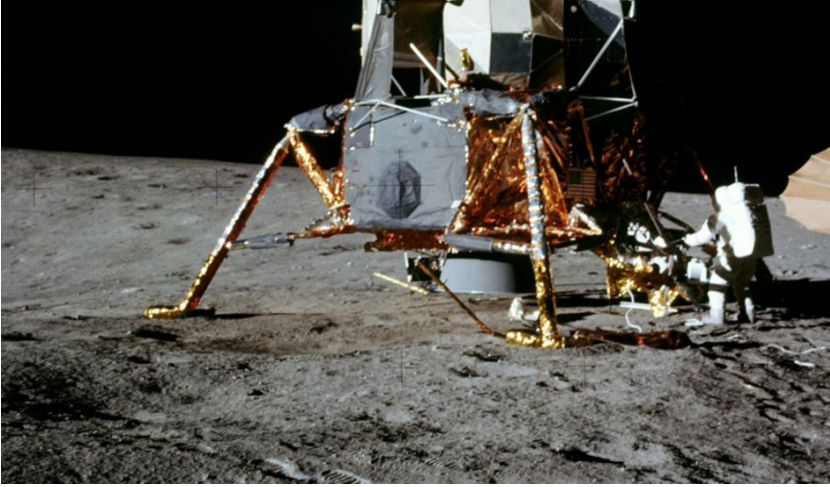
Fig. 9. View towards the south with Surveyor Crater to the left and possible scouring crater to the left of the engine nozzle (note area of discoloration).
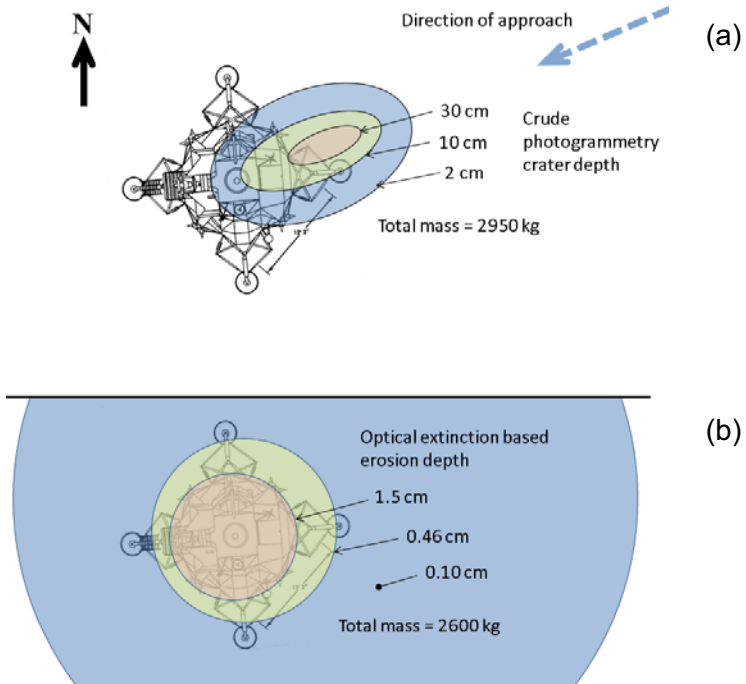

Fig. 10: (a) Crater profile based on crude photogrammetric measurements of the Intrepid landing site; (b) Contour map of erosion based on mass erosion from optical extinction measurements.

The size and depth of the predicted erosion depth (Fig. 10b) is too small to measure photographically (maximum depth is less than 2 cm). It must be pointed out that the cratering diagrammed in Fig. 10a is at best a hazardous guess since the previous elevation of the soil is unknown, and it is extremely difficult to get a quantifiable measurement of depth, even though multiple high resolution photographs are publically available at NASA Apollo mission archives. If the actual soil was eroded in the same fashion as the model suggests, that erosion might only be visible from a bird's eye view which might appear as a change in brightness of the surface in a large radius around the landing sit. This effect, known as "blast zone brightening", has been observed and as yet lacks a concrete explanation (Clegg *et al.* 2014).

Another outcome of this study is that optical extinction due to scattering of light from hydrometeors can be used to estimate rainfall rate (Lane *et al.* 2014b, Atlas 1953), just as microwave radar may be used to measure soil erosion rate. This connection is a consequence of the similarity of size range of the particle distributions of hydrometeors and lunar soil and the fact that the index of refraction which determines the details of electromagnetic scattering is similar. Before this equivalence is taken too far, however, it must be recognized that 10 cm weather radar does not detect fog size particles, which are comparable in size to the smallest lunar dust particles that may contribute to soil erosion. Therefore, to measure dust particles effectively, millimeter wave radar would need to be utilized to correctly complete this analogy.

4. EROSION RATE *VERSUS* SHEAR STRESS

Shear stress σ is computed from the CFD output as a post process and is equal to the product of the dynamic viscosity μ and the vertical gradient of the radial component of the plume gas velocity v_r. This can be expressed in terms of gas temperature using Sutherland's formula (Smits and Dussauge 2006):

$$\sigma = \mu \nabla_z v_r = \mu_0 \frac{T_0 + T_C}{\left(T(r,z) + T_C\right)} \cdot \frac{T(r,z)}{T_0} \cdot \frac{\partial v_r}{\partial z} \ , \tag{12}$$

where $\mu_0 = 1.83 \times 10^{-5}\,\text{kg m}^{-1}\,\text{s}^{-1}$, $T_0 = 291.2$ K, $T_C = 120$ K, and $T(r, z)$ is the gas temperature at a distance r from the nozzle centerline and a distance z above the surface. With this definition, shear stress has units N m^{-2}. Figure 11 shows the average shear stress computed by Eq. 12 for four engine heights above the surface (open squares), by averaging the shear stress over the radial distance, similar to the particle velocity averaging of Eq. 8:

$$\bar{\sigma}(h) = \frac{2\pi \int_0^{a_0(h)} \sigma(r,h)\,r\,dr}{\pi a_0^2(h)} \ . \tag{13}$$

The dotted line in Fig. 11 is an exponential fit of these data points, as a function of LM height $h(t)$. The mass erosion rate $\dot{m}(h(t))$ from Table 2, as defined by the model of Eq. 4 (open circles), is also an average value over the area of constant radius $a_0(h(t))$. By fitting both the average shear stress to Eq. 13 and average erosion rate to exponentials as a function of $h(t)$ to Eq. 4, the erosion rate can be expressed as a function of shear stress by eliminating $h(t)$:

$$
\begin{aligned}
\sigma(t) &= \sigma_0 e^{-\Gamma h(t)}, \\
\dot{m}(t) &= m_0 e^{-\Lambda h(t)}, \\
\dot{m}(t) &= m_0 \left(\sigma(t)/\sigma_0\right)^{\Lambda/\Gamma},
\end{aligned}
\tag{14}
$$

where $m_0 = 2.20$, $\sigma_0 = 6.21$, $\Gamma = 0.123$, and $\Lambda = 0.309$. The final relationship, based on this data, is:

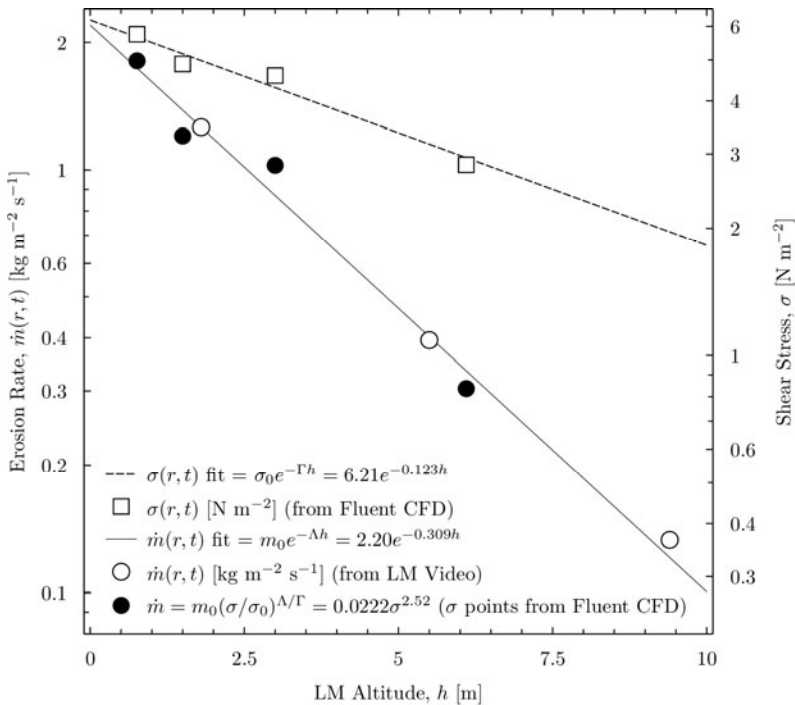$$\dot{m}(t) = 0.0222 \ \sigma^{2.52} \ \mathrm{kg\,s^{-1}\,m^{-2}} \ . \tag{15}$$

Fig. 11. Averaged shear stress and mass erosion rate as a function of $h(t)$.

5.   SUMMARY

A method for estimating lunar soil erosion rate due to plume impingement of the Apollo 12 Lunar Module Intrepid during its descent to the lunar surface has been presented. The observables are optical extinction and particle size distributions of soil samples returned from the lunar surface. The optical extinction is measured between the camera mounted inside of the cockpit window and the lunar surface during landing. CFD analysis of the Apollo LM descent engine, as well as particle trajectory analysis based on the CFD simulations, provides the remainder of the necessary data.

The CFD simulations provide a key piece of information: the velocity profile of particles as a function of starting distance from the engine nozzle, size of the particle, and height of the lander from the surface. Note that there may at times be some confusion as to what is defined as height above the surface: camera, engine nozzle, or landing pads. In this paper, height $h(t)$ when used in a quantifiable analysis, is defined as the height of the engine nozzle opening to the surface, which is generally half a meter or less after landing. At other times height may refer to altitude of the landing pads.

Taking an approach similar to the problem of estimating rainfall rate from weather radar, Eq. 4 was presented as the solution to the problem of estimating soil erosion rate from optical extinction measurements (see Fig. 12). In both cases, the particle velocities must be known, as well as the par-
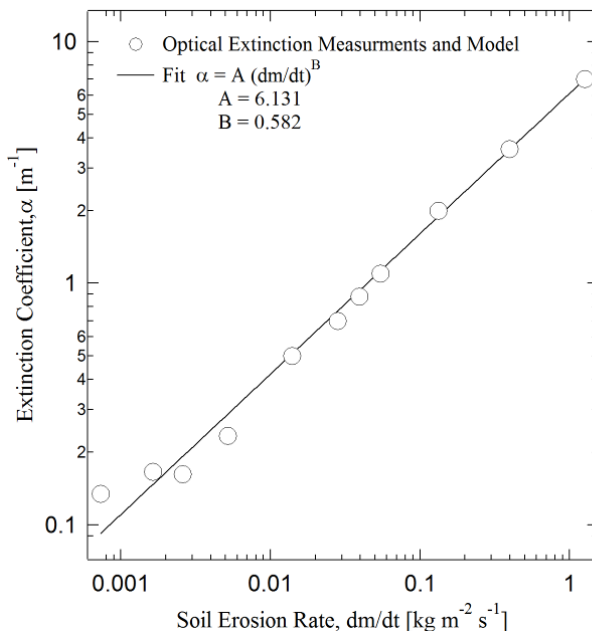

Fig. 12. Optical extinction α *versus* soil mass erosion rate $\dot{m}$, showing power law fit.

ticle size distributions. Quantifying the particle velocity function is in some sense the most difficult part of the problem for both lunar plume bservations and weather radar estimation of rainfall. For this reason, the velocity function is likely the source of greatest error. More advanced coupled gas-particle flow simulations should be able to provide an improved velocity function.

5.1 Estimating mass erosion rate from optical extinction

Equation 3 is used to estimate mass erosion rate from measurements of optical extinction once A and B are known. The A and B parameters should correspond to a particular engine design with a corresponding total vehicle mass M and surface gravity g, which implies a nominal thrust $T = Mg$ for a slow descent or hovering. The A and B also correspond to a specific soil type characterized by a size distribution $S(D)$.

The following summarizes the key points in determining A and B for the first time:

1. For various values of measured extinction factor α, corresponding values of mass erosion rate are computed using Eq. 4. These point pairs can be plotted on a log-log graph with a straight line fit to the scatter plot, providing the A and B parameters, as demonstrated in Fig. 12. A method to determine α from descent videos using histogram matching has been described, as a special case.

2. The velocity $v(D)$ in Eq. 4 is computed from a CFD based empirical function $u(h, D)$, such as the velocity model given by Eq. 7. Then $v(D) = \eta \sin \theta\, u(h, D)$ where θ is the angle of the dust sheet relative to horizontal.

3. η is part of the velocity model and corresponds to the ratio of mean velocity to the maximum velocity. In the example given by Eq. 8, $\eta \approx 1/8$. It is assumed that η is a constant of the engine design and has no dependencies on $S(D)$.

4. The symbol φ in Eq. 4 is a geometrical value relating the dispersion of the dust to view angle along the optical extinction path. In Appendix A, this is shown to be a constant ≈ 2. Based on the arguments given in Appendix A, φ should not vary much from this approximate value of 2 under varying conditions of engine design or soil type.

5. $s(f)$ in Eq. 4 is an extinction shape factor which, according to the arguments given in Appendix B, can be approximated by a sphere with $s(f) = 1$, with less than 10% error for an ellipsoid when $f < 3$.

5.2 Image analysis and optical extinction

The details of the optical extinction model and data analysis of the Apollo LM video have been deferred to Appendix C in order to avoid obscuring the details and significance of Eq. 4 in estimating soil erosion rates. The optical extinction model described in Appendix C assumes from the start that the erosion rate is uniform over an area defined by radius a_0, which is a function of lander height. The primary influence of optical extinction as measured by a reduction of brightness of the surface and increase in brightness of the dust cloud, is the spatial dust geometry. In this model, erosion is uniform over πa_0^2 but diverges radially from all points on the surface within $r \leq a_0$. No erosion occurs for $r > a_0$. The underlying assumption of HMM is that by matching the average and standard deviation of the histogram of a reference image to the histogram of a processed image (modified by brightness and contrast equalization), the optical extinction factor can be deduced. Even though the results seem promising, the accuracy of the HMM output parameters have not been quantified. This is a possible area of future work.

5.3 Erosion rate as a function of shear stress

A relationship between soil erosion rate and shear stress as computed from specific engine design characteristics is highly desirable. The value of this relationship is that the total plume/erosion effects of engine design on any surface can be predicted. To this end, an empirical relationship was established between shear stress as determined by CFD simulation and erosion rate estimated by optical extinction measurements for the case of the Apollo 12 LM. The extension of these results to other engine designs for landing on any celestial body lacking an atmosphere, such as the Earth's moon or asteroids, can be used within the limits of this analysis. The result of the predicted Apollo 12 $\dot{m}(\sigma)$ relation, as shown by Eq. 15, shows an approximate 5/2 power dependence of erosion rate on shear stress. Since the value of the exponent is one of the two parameters in the relationship, it is subject to sensitivity of the data measurement and analysis, as well as model assumptions. It is in fact not terribly difficult to force a linear relation (as previously believed) by substituting different values of the optical extinction data that are within credible limits of measurement error.

Equation 15 is an empirical relation for mass erosion rate $\dot{m}$ as a function of shear stress σ. Previous work concluded that the relationship should be a linear one of the form:

$$\dot{m}(t) = c^{-1}\left(\sigma - \sigma_c\right)^{\beta} \; , \tag{16}$$

where σ_c is the threshold shear stress associated with a saltation velocity threshold. The inverse proportionality constant c has units of velocity. Equation 16 with $\beta = 1$ is the form predicted by Roberts (1963). Also, experiments at KSC revealed that mass erosion was proportional to the dynamic pressure of the jet leaving the pipe, *i.e.*, ρv^2, times the area of the pipe. That is also equivalent to the total thrust. It is also equal to momentum flux, which agrees with Roberts that erosion is a momentum-driven process, not an energy-driven process. According to Roberts' plume analysis theory, shear stress everywhere on the surface is proportional to thrust of the rocket, indirectly implying that the relationship is linear ($\beta = 1$). Haehnel and Dade (2008) conducted experiments where they directly measured shear stress and erosion rate locally everywhere on the surface. Erosion rate and shear stress were found to be linearly related through a global pair of constants, σ_c and β, with $\beta = 1$.

However, complexities of the lunar case that the above three efforts do not account for include: saltation due to particles scattering out of the cloud back down to the surface, rarefaction effects, and turbulence effects, which are different in rarefied or transitional flow than in continuum flow and have never been adequately studied. Turbulence is not modeled in the existing rarefied/transitional gas flow codes. Therefore, it is not unreasonable that in the lunar environment the actual value of β may be a non-integer, as indicated by the result shown in Eq. 15.

Appendix A

Erosion model geometry

To quantify the effect of erosion flux divergence, shown as grey arrows originating from the surface under the rocket plume in Fig. 13, Eq. 4 ($\varphi = 1$ describes the non-divergent case) can be applied to a small differential of erosion $\dot{m}_{jk}$. The camera image is then affected by the optical extinction occurring over a small distance l_{jk} along ray j due to $\dot{m}_{jk}$:

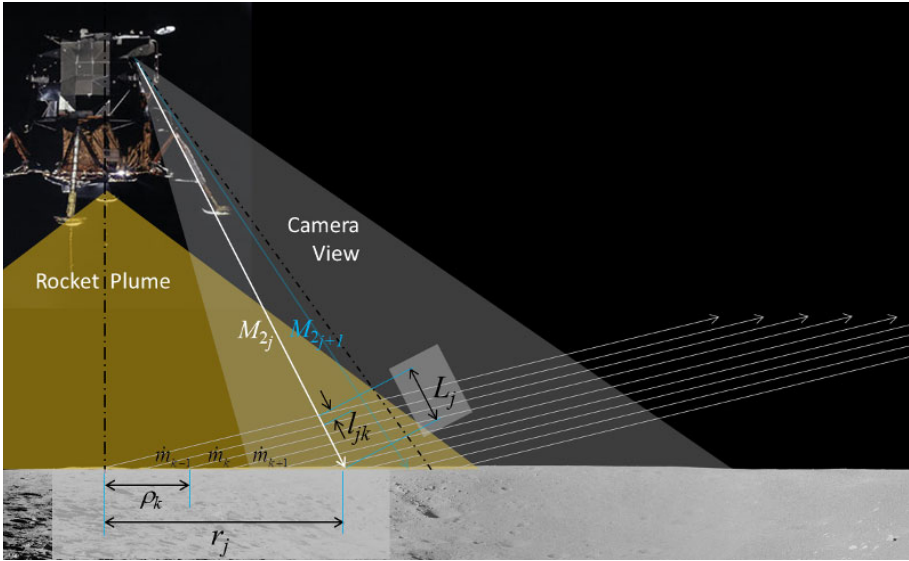$$M_{2_{jk}} = \Phi_k^{-1} \dot{m}_{jk} \ , \tag{A1}$$

Fig. 13. Schematic diagram of erosion model geometry, leading to $\varphi = 2$ in Eq. 4.

where $M_{2_{jk}}$ is the optical extinction described by the second moment of the size distribution at path differential l_{jk} along ray j; Φ_k is the collection of all other terms on the right hand side of Eq. 4, characterized by velocity distribution modeled of Eq. 7.

The "first model assumption" is:

$$\dot{m}_{jk} = \frac{\rho_k}{r_j} \dot{m}_k \ , \qquad (A2)$$

which approximates the flux divergence as constrained to the shallow grey conical surface shown as arrows in Fig. 13. The total optical extinction along ray j is then:

$$M_{2_j} = \sum_k \frac{l_{jk}}{L_j} M_{2_{jk}} \ . \qquad (A3)$$

Substituting Eqs. A1 and A2 into A3, and letting $l_{jk} = L_j/n$:

$$M_{2_j} = \frac{1}{n} \sum_k \Phi_k^{-1} \frac{\rho_k}{r_j} \dot{m}_k \ . \qquad (A4)$$

The "second model assumption" is to set all $\dot{m}_k$ equal, corresponding to constant erosion over radius a_0. The "third model assumption" is to set all Φ_k

equal, which is in the spirit of Eq. 8 where the radial dependence of particle velocity is modeled as a constant. Then Eq. A4 becomes:

$$M_{2_j} = \frac{1}{n}\Phi_0^{-1}\dot{m}_0 \sum_k \frac{\rho_k}{r_j} \ . \tag{A5}$$

The radial distance ρ_k in Eq. A5 can be replaced by $k\,r_j/n$:

$$M_{2_j} = \Phi_0^{-1}\dot{m}_0 \frac{1}{n^2}\sum_k k \ \lim_{n\to\infty} M_{2_j} \to \frac{1}{2}\Phi_0^{-1}\dot{m}_0 \ . \tag{A6}$$

Comparing Eq. A6 to Eq. 4, $\varphi = 2$. Note that this is not likely a fundamental physical principle, but is more than likely a consequence of this simple model and its set of assumptions.

Appendix B

Calculation of particle shape factor for spheroid

A rudimentary particle shape model, one level of improvement over a spherical particle, is the spheroid, described by aspect ratio $f = r_b/r_a$. If $f > 1$, the particle is a "prolate spheroid". If $f < 1$, the particle is an "oblate spheroid". And of course when $f = 1$ it is a sphere. The volume weighted diameter is $D = 2r_a f^{1/3}$. Figure 14 shows a prolate spheroid with $f = 2.5$. The surface of
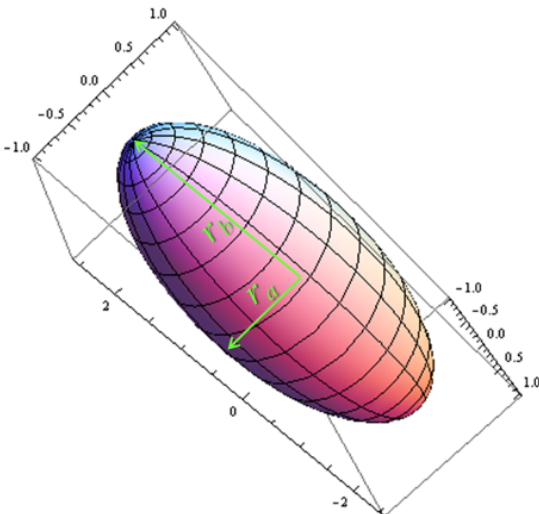

Fig. 14. Prolate spheroid with aspect ratio $f = r_b/r_a = 2.5$.

the spheroid can be represented by a Cartesian vector $\mathbf{P}$, which is a function of parametric angles u and v (similar to spherical coordinate angles θ and ϕ):

$$\mathbf{P} = r_a \cdot \begin{pmatrix} \cos u \sin v \\ \sin u \sin v \\ f \cos v \end{pmatrix} \qquad \begin{array}{l} 0 \le u < 2\pi \\ 0 \le v < \pi \end{array} . \tag{B1}$$

The average area $\bar{A}$ of a randomly oriented spheroid can be found by integrating over all values of the randomly projected major axis r_b, where the projection is a sinusoidal function with limits between r_a and r_b:

$$\bar{A} = r_a^2 \int_0^\pi \left(1 + (f-1)\sin\theta\right) d\theta = \left(\pi + 2(f-1)\right) r_a^2 . \tag{B2}$$

The shape factor $s(f)$ can then be equated to $\bar{A}$, normalized by the volume weighted cross-section of the spheroid:

$$s(f) = \frac{\bar{A}}{\pi \left(r_a f^{1/3}\right)^2} = \frac{\pi + 2(f-1)}{\pi f^{2/3}} . \tag{B3}$$

This result is valid for both the prolate and oblate cases. Figure 15 is a plot of $s(f)$ for f ranging from 0.2 to 5. Note that in the case of $f = 2.5$, $s(f) = 1.06$, which will decrease the erosion rate of Eq. 4 by approximately 6%.
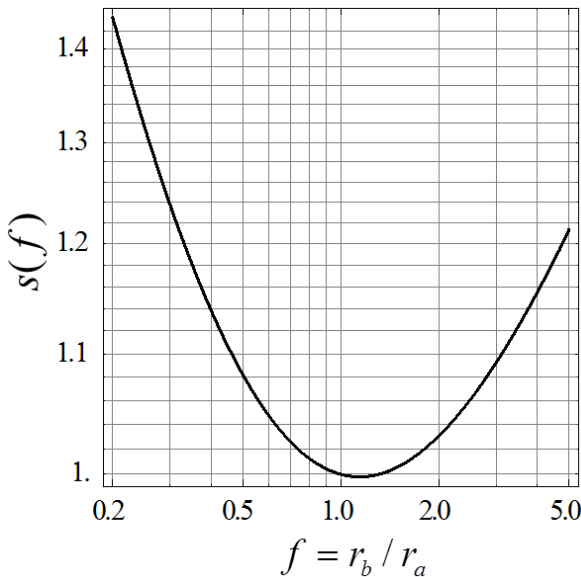

Fig. 15. Shape factor $s(f)$ used in Eq. 4, described by Eq. B3.

Appendix C

Histogram matching method

Characterizing dust plumes on the moon's surface during a rocket landing is imperative to the success of future operations on the moon or any other celestial body with a dusty or soil surface (including cold surfaces covered by frozen gas ice crystals, such as the moons of the outer planets). The most practical method of characterizing the dust clouds is to analyze video or still camera images of the dust illuminated by the sun or on-board light sources (such as lasers). The method described below was used to characterize the dust plumes from the Apollo 12 landing.

In this context, the histogram matching method (HMM) is an image processing technique for determining dust optical density in Apollo landing videos. The software implementation of HMM creates a greyscale image histogram and calculates the histogram mean and standard deviation, which is then used to match dusty and clear images for the purpose of estimating an effective optical density and optical extinction factor α. A dust thickness model, based on the tilt of the camera and increasing height of the dust layer towards the top of the image, is used to account for the distance light travels through the dust.

Previous methods relied on comparing specific features in clear *versus* dusty images, which severely limited ability to analyze video frames. This method compares the statistical nature of a clear image to the statistical nature of a dusty image, assuming that the average scene's description (as characterized by an image histogram) due to surface reflectance and sun angle is invariant throughout the frame sequence. This assumption fails when shadows show up on the scene, which is evident in the last 20 s of the landing descent. In the last 20 s, the error minimization of the histogram matching is by-passed and the matching is done manually by visually comparing images.

The output of the HMM algorithm is a modified image, where "dust" has been added (mode 1) or removed (mode 0). The output image pixel p'_{ij} is computed from the input image pixel p_{ij}:

$$p'_{ij} = \begin{cases} p_{ij}e^{-2\alpha x} + p_0\left(1 - e^{-\alpha x}\right) & \text{mode} = 1 \\ p_0 e^{\alpha x} - \left(p_0 - p_{ij}\right)e^{2\alpha x} & \text{mode} = 0 \end{cases}, \tag{C1}$$

where α is the optical extinction factor of the dust and p_0 is a fitting parameter associated with the dynamic range of the image (ideally $p_0 = 255$ for an

8-bit image). The factor of two in the extinction term is the result of light re-
flecting off of the lunar surface back to the camera. The exponent term with-
out the factor of two corresponds to light scattered back to the camera from
the dust cloud. The distance x in Eq. C1 is the effective optical-dust path
length model along the camera view ray through the dust cloud, correspond-
ing to each ij image pixel in the image. It is equal to the physical path length
x_D of the dust for $r \leq a_0$, where r is the radial distance from the engine noz-
zle centerline and a_0 is a parameter. For $r > a_0$, the effective path length is x_D
scaled by the radial dispersion factor:

$$x = \begin{cases} x_D & r \leq a_0 \\ x_D \left(\dfrac{a_0}{r} \right)^2 & r > a_0 \end{cases} , \qquad (C2)$$

$$x_D = \frac{\tan\theta \cot\varphi \left(l_c + h\tan\varphi \right) \sec\left(\varphi + \tan^{-1}\left(qd/F \right) \right)}{\cot\left(\varphi + \tan^{-1}\left(qd/F \right) \right) + \tan\theta} , \qquad (C3)$$

where F is the focal length of the camera ($F = 10$ mm), d is the pixel width
($d \approx 15$ μm), θ is the dust angle relative to horizontal ($\theta \approx 3°$), φ is the cam-
era angle relative to vertical ($\varphi = 33°$), $l_c = 1.2$ m is the camera offset dis-
tance from the nozzle center line, and h is height of the LM above the
surface. The variable q is the vertical distance in the image in pixel units
from the ij pixel to a horizontal centerline in the rotated camera view:

$$q = \left(j - \tfrac{1}{2}N \right)\cos\zeta - \left(i - \tfrac{1}{2}M \right)\sin\zeta , \qquad (C4)$$

where ζ is the camera rotation angle about the camera axis ($\zeta \approx -33°$), i is
the horizontal pixel index, j is the vertical pixel index, N is the total number
of horizontal pixels, and M is the total number of vertical pixels.
Figure 16a shows frame F3077 ($h = 34$ m) of the cockpit video camera. Fig-
ure 16b displays a map of the same field of view for this frame, showing the
effective optical-dust path length model, Eq. C2, which is based on the tilt of
the camera and increasing depth of the dust layer towards the top of the im-
age. The horizontal and vertical axes of the plot are in pixel units. The con-
tours are graded in increments of 0.3 m, starting with the minimum $x = 0.3$ m
at the bottom (purple) to a maximum $x = 3.0$ m at the top (red).
    Figure 17 is a similar image set, occurring 38.8 s later at an LM altitude
of $h = 11$ m. The contours are graded in increments of 0.017 m, starting with
the minimum $x = 0.1$ m at the bottom (purple) to a maximum $x = 0.25$ m
near the center (red). Note that the video frame numbers F3077 (Fig. 16) and
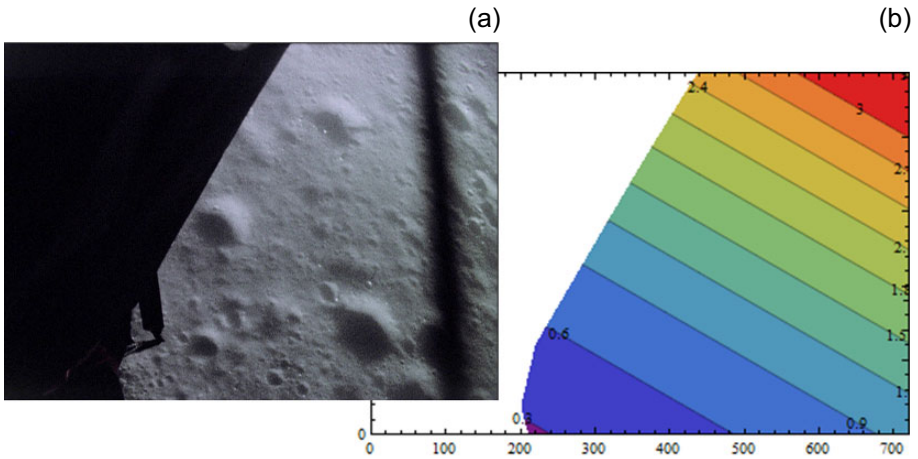F3543 (Fig. 17) correspond to a constant frame rate of 12 fps.

(a)   (b)

Fig. 16. Dust depth model: (a) video camera frame (F3077) with LM altitude $h = 34$ m, (b) effective camera dust length x with radius $a_0 = 46$ m.

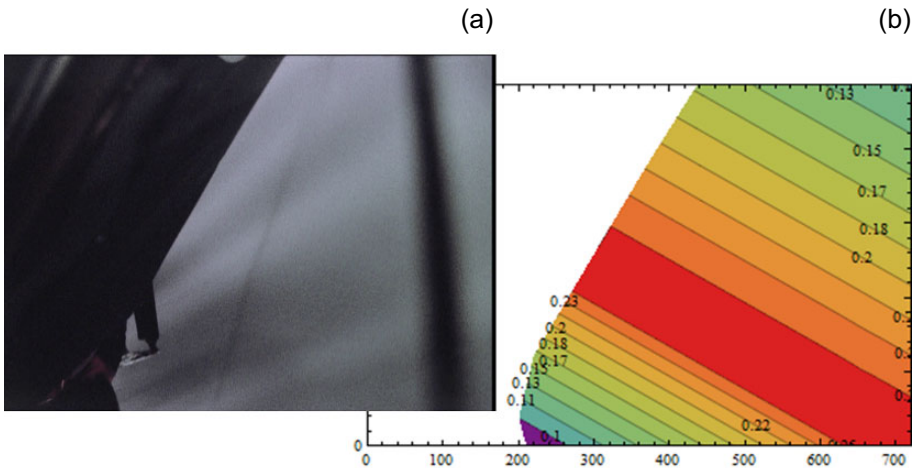(a)   (b)

Fig. 17. Dust depth model: (a) video camera frame (F3543) with LM altitude $h = 11$ m, (b) effective dust depth with radius $a_0 = 6.5$ m.

The HMM algorithm processes two input images, pixel by pixel. The first input image is represented by pixel p_{ij} as shown by Eq. C1. A reference image is represented by q_{ij}. For mode = 0, the HHM algorithm applies the transformation described by Eq. C1 to the input image p_{ij} (frame with dust), creating an output image p'_{ij} (artificially removed dust), as shown in Fig. 18. The reference image q_{ij} (no dust) is then compared to p'_{ij} and by matching the average and standard deviation of the their histograms, the parameters α, p_0, and a_0 are found. Figure 19 shows a similar example for mode = 1.
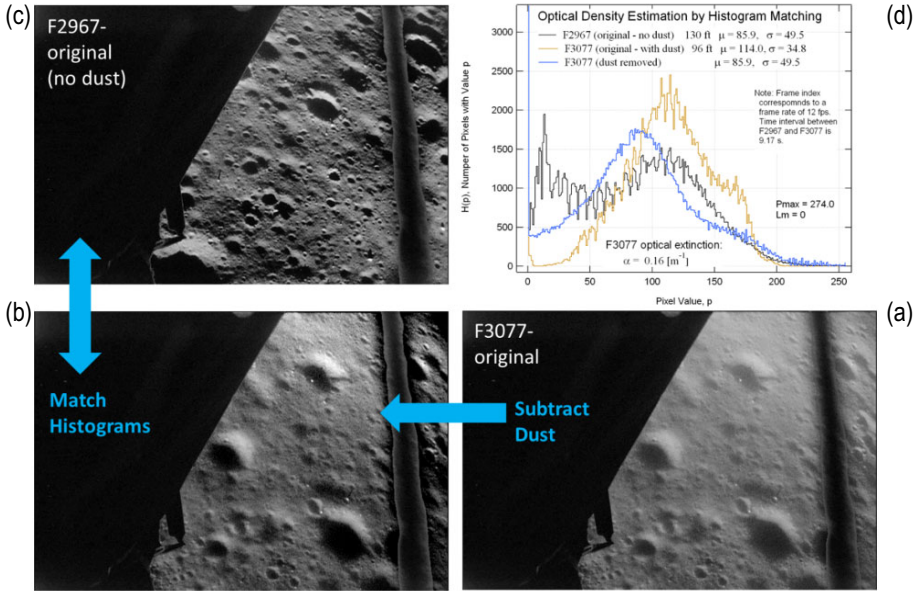
Fig. 18. HMM algorithm example for mode = 0: (a) image p_{ij}, (b) output image p'_{ij}, (c) reference image q_{ij}, and (d) histograms for the three images.
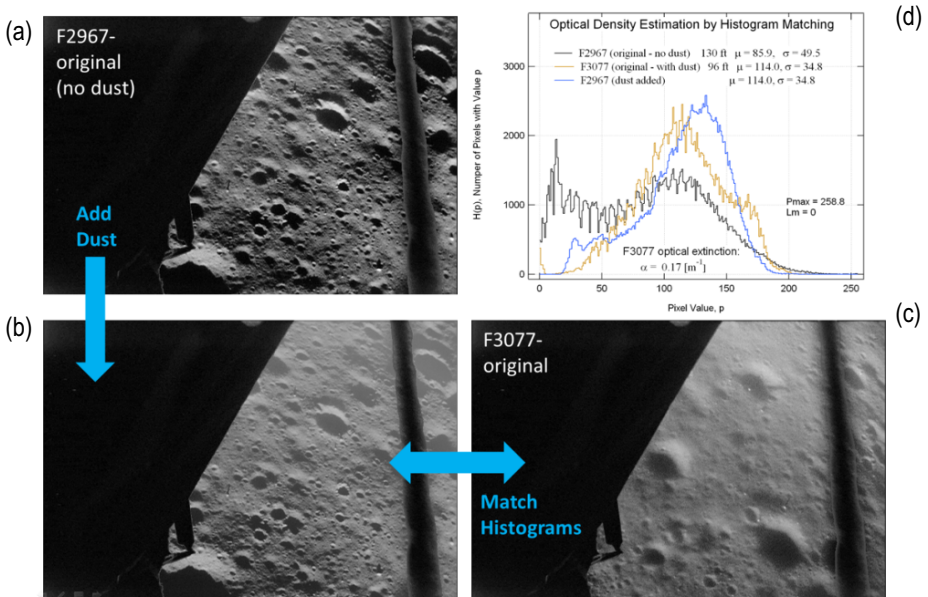

Fig. 19. HMM algorithm example for mode = 1: (a) input image p_{ij}, (b) output image p'_{ij}, (c) reference image q_{ij}, and (d) histograms.

Glossary of symbols

Symbol	Description	Standard units	*si* units
$N(D)$	lofted particle size distribution	$\mu m^{-1}\,m^{-3}$	$m^{-1}\,m^{-3}$
D	particle diameter	μm	m
$S(D)$	normalized soil size distribution, empirical fit	μm^{-1}	m^{-1}
$S_1(D)$	component of $S(D)$ fit	–	–
$S_2(D)$	component of $S(D)$ fit	–	–
$w(D)$	component of $S(D)$ fit	–	–
$D_1 \ldots D_3$, $B_1 \ldots B_3$	fitting constants in particle size fraction model $S(D)$	μm dimensionless	m dimensionless
$v(D)$	vertical component of particle velocity	$m\,s^{-1}$	$m\,s^{-1}$
R	rainfall rate	$mm\,h^{-1}$	$m\,s^{-1}$
Z	radar reflectivity	$mm^6\,m^{-3}$	m^3
a and b	parameters of Z-R, Eq. 1	–	–
A and B	parameters of α-$\dot{m}$, Eq. 3	–	–
$\dot{m}$	lunar soil erosion rate	$kg\,s^{-1}\,m^{-2}$	$kg\,s^{-1}\,m^{-2}$
$\Gamma(D)$	Gamma Function of D	–	–
λ	wavelength of light	nm	m
ρ_L	bulk density of lunar soil	$g\,cm^{-3}$	$kg\,m^{-3}$
M_x	xth moment of size distribution	$mm^x\,m^{-3}$	m^{x-3}
φ	geometry factor in Eq. 4	dimensionless	dimensionless
$s(f)$	particle shape factor	dimensionless	dimensionless
α	optical extinction factor	$\mu m^2\,m^{-3}$	m^{-1}
x	particle size parameter	dimensionless	dimensionless
r_a, r_b	short and long radius of particle ellipsoid, respectively	μm	m
f	particle shape factor, r_a/r_b	dimensionless	dimensionless
Q_e	scattering efficiency factor for extinction	dimensionless	dimensionless
n	refractive index	dimensionless	dimensionless
$a_0(t)$	ideal radius of surface erosion as a function of time t	m	m
$h(t)$	nozzle opening distance from surface as a function of time t	m	m

Glossary of symbols   (continuation)

Symbol	Description	Standard units	Si units
m_T	total mass ejected (total mass displaced)	MT = 1000 kg	kg
$u(h, D)$	CFD based particle maximum speed model	m s^{-1}	m s^{-1}
$b_0 \ldots b_4$, c_0, c_1	empirical fitting constants in CFD particle speed model	–	–
η	mean CFD particle speed compared to maximum	dimensionless	dimensionless
$V(r)$	particle speed fit, r is the trajectory starting point (radial distance from nozzle center)	m s^{-1}	m s^{-1}
R_1, R_2	min. and max. CFD domain distance for determining η	m	m
v_0, v_1, ξ	fitting constants in $V(r)$ fit	m s^{-1}, m s^{-1}, m	m s^{-1}, m s^{-1}, m
σ	CFD derived shear stress	N m^{-2}	N m^{-2}
μ	plume gas dynamic viscosity	kg m^{-1} s^{-1}	kg m^{-1} s^{-1}
$T(r, z)$	CFD gas temperature, r from nozzle center, z above surface	K	K
v_r	radial component of the plume gas velocity	m s^{-1}	m s^{-1}
T_0, T_C, μ_0	constants in Sutherland's formula for shear stress	–	–
$\bar{\sigma}(h)$	CFD shear stress at surface, averaged over radial distance, *versus* engine height	N m^{-2}	N m^{-2}
σ_0, Γ	shear stress fitting parameters, Eq. 13	N m^{-2}, m^{-1}	N m^{-2}, m^{-1}
m_0, Λ	erosion rate fitting parameters, Eq. 13	kg , m^{-1}	kg , m^{-1}
c, σ_c, β	parameters in theoretical shear stress model	m s^{-1}, N m^{-2}	m s^{-1}, N m^{-2}, dimensionless

References

Atlas, D. (1953), Optical extinction by rainfall, *J. Meteor.* **10**, 6, 486-488, DOI: 10.1175/1520-0469(1953)010<0486:OEBR>2.0.CO;2.

Berg, M.J., C.M. Sorensen, and A. Chakrabarti (2011), A new explanation of the extinction paradox, *J. Quant. Spectrosc. Rad. Trans.* **112**, 7, 1170-1181, DOI: 10.1016/j.jqsrt.2010.08.024.

Berger, K.J., A. Anand, P.T. Metzger, and C.M. Hrenya (2013), Role of collisions in erosion of regolith during a lunar landing, *Phys. Rev. E* **87**, 2, 022205, DOI: 10.1103/PhysRevE.87.022205.

Clegg, R.N., B.L. Jolliff, M.S. Robinson, B.W. Hapke, and J.B. Plescia (2014), Effects of rocket exhaust on lunar soil reflectance properties, *Icarus* **227**, 1, 176-194, DOI: 10.1016/j.icarus.2013.09.013.

Haehnel, R., and W.B. Dade (2008), Physics of particle entrainment under the influence of an impinging jet. **In:** *Proc. 26th Army Science Conference, U.S. Army, Orlando, USA*.

Immer, C., P. Metzger, P.E. Hintze, A. Nick, and R. Horan (2011a), Apollo 12 Lunar Module exhaust plume impingement on Lunar Surveyor III, *Icarus* **211**, 2, 1089-1102, DOI: 10.1016/j.icarus.2010.11.013.

Immer, C., J. Lane, P. Metzger, and S. Clements (2011b), Apollo video photogrammetry estimation of plume impingement effects, *Icarus* **214**, 1, 46-52, DOI: 10.1016/j.icarus.2011.04.018.

Lane, J.E., and P.T. Metzger (2014a), Image analysis based estimates of regolith erosion due to plume impingement effects. **In:** *Proc. 14th ASCE Int. Conf. Engineering, Science, Construction and Operations in Challenging Environments "Earth and Space 2014", 27-29 October 2014, St. Louis, USA*.

Lane, J.E., P.T. Metzger, and J.W. Carlson (2010), Lunar dust particles blown by lander engine exhaust in rarefied and compressible flow, **In:** *Proc. 12th ASCE Int. Conf. Engineering, Science, Construction and Operations in Challenging Environments "Earth and Space 2010", 14-17 March 2010, Honolulu, Hawaii, USA*, 134-142, DOI: 10.1061/41096(366)16.

Lane, J.E., T. Kasparis, P.T. Metzger, and W.L. Jones (2014b), In situ disdrometer calibration using multiple DSD moments, *Acta Geophys.* **62**, 6, 1450-1477, DOI: 10.2478/s11600-014-0237-2.

Metzger, P.T., J.E. Lane, and C.D. Immer (2008), Modification of Roberts' theory for rocket exhaust plumes eroding lunar soil. **In:** *Proc. 11th ASCE Int. Conf. Engineering, Science, Construction and Operations in Challenging Environments "Earth and Space 2008", 3-5 March 2008, Long Beach, USA*, 1-8, DOI: 10.1061/40988(323)4.

Metzger, P.T., J.E. Lane, C.D. Immer, J.N. Gamsky, W. Hauslein, X. Li, R.C. Latta III, and C.M. Donahue (2010), Scaling of erosion rate in subsonic jet experiments and Apollo lunar module landings. **In:** *Proc. 12th ASCE Int. Conf. Engineering, Science, Construction and Operations in Challenging Environments "Earth and Space 2010", 14-17 March 2010, Honolulu, Hawaii, USA*, 191-207, DOI: 10.1061/41096(366)21.

Metzger, P.T., J. Smith, and J.E. Lane (2011), Phenomenology of soil erosion due to rocket exhaust on the Moon and the Mauna Kea lunar test site, *J. Geophys. Res.* **116**, E6, E06005, DOI: 10.1029/2010JE003745.

Morris, A.B., D.B. Goldstein, P.L. Varghese, and L.M. Trafton (2011), Plume impingement on a dusty lunar surface, *AIP Conf. Proc.* **1333**, 1187-1192, DOI: 10.1063/1.3562805.

Roberts, L. (1963), The action of a hypersonic jet on a dust layer. **In:** *31st Ann. Meeting, Institute of Aerospace Sciences, New York, USA*, IAS paper No. 63-50.

Rosenfeld, D., D.B. Wolff, and D. Atlas (1993), General probability-matched relations between radar reflectivity and rain rate, *J. Appl. Meteorol.* **32**, 1, 50-72, DOI: 10.1175/1520-0450(1993)032<0050:GPMRBR>2.0.CO;2.

Scott, R.F. (1975), Apollo program soil mechanics experiment, Final report, California Institute of Technology, Pasadena USA.

Shipley, S.T., E.W. Eloranta, and J.A. Weinman (1974), Measurement of rainfall rates by lidar, *J. Appl. Meteorol.* **13**, 7, 800-807, DOI: 10.1175/1520-0450(1974)013<0800:MORRBL>2.0.CO;2.

Smits, A.J., and J.P. Dussauge (2006), *Turbulent Shear Layers in Supersonic Flow*, 2nd ed., Springer Science+Business Media, New York, 424 pp.

van de Hulst, H.C. (1957), *Light Scattering by Small Particles*, John Wiley & Sons, New York.

Wexler, R., and D. Atlas (1963), Radar reflectivity and attenuation of rain, *J. Appl. Meteor.* **2**, 2, 276-280, DOI: 10.1175/1520-0450(1963)002<0276:RRAAOR>2.0.CO;2.