

Hindawi Publishing Corporation
Shock and Vibration
Volume 2016, Article ID 4051701, 6 pages
<http://dx.doi.org/10.1155/2016/4051701>



Research Article

Algorithm Indicating Moment of P-Wave Arrival Based on Second-Moment Characteristic

**Jakub Sokolowski,¹ Jakub Obuchowski,¹ Radosław Zimroz,¹
Agnieszka Wylomanska,² and Eugeniusz Koziarz³**

¹KGHM CUPRUM Ltd., CBR, Sikorskiego 2-8, 53-659 Wrocław, Poland

²Faculty of Pure and Applied Mathematics, Hugo Steinhaus Center, Wrocław University of Science and Technology, Janiszewskiego 14a, 50-370 Wrocław, Poland

³KGHM Polska Miedź S.A. O/ZG “Rudna”, Henryka Dąbrowskiego 50, 59-100 Polkowice, Poland

Correspondence should be addressed to Jakub Sokolowski; jsokolowski@cuprum.wroc.pl

Received 22 April 2016; Accepted 1 August 2016

Academic Editor: Shimin Liu

Copyright © 2016 Jakub Sokolowski et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The moment of P-wave arrival can provide us with many information about the nature of a seismic event. Without adequate knowledge regarding the onset moment, many properties of the events related to location, polarization of P-wave, and so forth are impossible to receive. In order to save time required to indicate P-wave arrival moment manually, one can benefit from automatic picking algorithms. In this paper two algorithms based on a method finding a regime switch point are applied to seismic event data in order to find P-wave arrival time. The algorithms are based on signals transformed via a basic transform rather than on raw recordings. They involve partitioning the transformed signal into two separate series and fitting logarithm function to the first subset (which corresponds to pure noise and therefore it is considered stationary), exponent or power function to the second subset (which corresponds to nonstationary seismic event), and finding the point at which these functions best fit the statistic in terms of sum of squared errors. Effectiveness of the algorithms is tested on seismic data acquired from O/ZG “Rudna” underground copper ore mine with moments of P-wave arrival initially picked by broadly known STA/LTA algorithm and then corrected by seismic station specialists. The results of proposed algorithms are compared to those obtained using STA/LTA.

1. Introduction

Obtaining accurate information about seismic phenomena induced by mining activity might be a difficult task. The recordings strongly depend on distance between source and measuring device, energy of the event, lithology of the rock mass, device parameters, noise induced by transmission line, and so forth. In order to acquire exact features of the event (like, e.g., 3-dimensional location) recordings from at least four different one-axial sensors are required.

When the seismic event occurs, its energy is transported via different types of seismic waves, which can be primarily classified as body waves (P-wave, S-wave) and surface waves (Rayleigh wave, Love wave, and Stoneley wave). P-waves possess the highest velocity among others; thus they indicate

onset of the event. Therefore, in order to receive detailed information about particular phenomenon, the first step is to indicate its moment of P-wave arrival.

From mathematical point of view the problem is isometric with finding a moment in time series where it loses stationarity (as the background noise is considered to be stationary) or as a problem of finding structural break point.

The moment of P-wave arrival is commonly used in estimation of event location [1], energy [2], and focal mechanism [3]. Determining such a moment manually is time-consuming and requires considerable experience. However, under development of science and technology, many automatic P-wave picking algorithms were proposed. Implementation and use of such methods are a much faster solution but not 100% reliable, as the results frequently differ from indications

given by seismic station specialists. Thus the algorithms are frequently used as an initial pick followed by experts' manual correction.

So far, there are plenty of different algorithms which can be divided into 2 main groups: proceeding in time and proceeding in frequency domain [4]. Broadly known time domain methods include AR-AIC [5, 6], which fits autoregressive model to the data and determines the moment of P-wave arrival in a point where Akaike Information Criterion [7] is minimized, and STA/LTA algorithm [8, 9], which for fixed characteristic function (e.g., square of the signal) computes its average over short and long time window and indicates the onset time when the ratio of averages exceeds predefined value. The moment of P-wave arrival might be also determined with use of neural networks [10, 11], methods based on wavelet transform [12, 13], spectrogram [14, 15], and cross-correlation [16].

When dealing with the problem of P-wave arrival moment, one may investigate it as an element of a signal segmentation procedure [17, 18], as the indication of onset moment is basis for segmentation. Common methods are often used in both problems.

Recently, a method of finding a critical point which divides the time series into two stationary parts with different variances has been proposed [19]. The basis for this method is statistical property of the second central statistical moment; that is, the expected value of cumulative sum of squares for stationary time series increases linearly with time. Such property is independent of the underlying probability distribution, as long as the variance is finite. The method has been already utilized in structural break detection method [20]. It was decided to involve this idea for P-wave arrival point estimation. However, the entire seismic event does not possess stationarity property, nor it can be split into two stationary time series. Thus the method requires a modification. In this paper two similar methods are proposed and compared to a widely used STA/LTA algorithm. All of the investigated automatic P-wave picking methods are compared with arrivals indicated by specialists of O/ZG "Rudna" underground copper ore mine seismic station experts due to their extensive experience in analysis of mining-induced seismic events.

The rest of the paper is organized as follows: in Methodology the new method of structural break detection is presented. Moreover, we recall the STA/LTA algorithm (the classical method used to detection of P-wave arrival time). Next, in Section 3 of application to real data, the new methodology is applied to the real seismic signals. Obtained results are compared with the STA/LTA technique. The last section contains conclusions.

2. Methodology

2.1. STA/LTA Algorithm. One of the classical algorithms that are often used in the problem of P-wave arrival moment detection is based on the short-term-average and long-term-average (STA/LTA) trigger method. The underlying idea of this method is to evaluate in a continuous fashion the value

of characteristic function (CF) of a seismic signal in two moving-time windows (one short and one long) in order to detect the seismic event. The characteristic function used for calculation purpose can be defined as energy, absolute amplitude, or envelope function of the microseismic trace. Irrespective of the definition of the characteristic function (CF), the short time window (STA) is supposed to measure the instantaneous amplitude of the seismic signal, whereas the long time window (LTA) provides information about the amplitude of seismic noise. When their ratio exceeds a predefined value τ_A (activation threshold), the following recorded samples are marked to be event-driven until the ratio falls below another predefined value τ_D called the deactivation threshold. In this algorithm, for a raw signal x_1, \dots, x_n , the following statistic SLR_t is being calculated:

$$SLR_t = \frac{(1/s) \sum_{k=t-s}^t CF(x_k)}{(1/l) \sum_{k=t-l}^t CF(x_k)}, \quad (1)$$

where s and l denote short and long time windows lengths (in samples), respectively. Moreover, in the above equation $CF(x)$ is a specific characteristic function defined in terms of signal energy. In the literature different characteristic functions can be found, such as absolute value of the signal or envelope of the microseismic trace. In this paper we consider $CF(x) = x^2$.

In the STA/LTA algorithm the inspection of the SLR_t statistic is performed and on such basis one can detect the moment of P-wave arrival. This moment is the minimum t for which the ratio STA/LTA exceeds the predefined value ρ ; that is,

$$t_P = \min_t SLR_t > \rho. \quad (2)$$

In this paper we compare the classical approach based on the STA/LTA algorithm with the new algorithm based on the cumulative empirical second moment of given raw signal.

2.2. Algorithm Based on the Empirical Second Moment. As it was mentioned, the proposed method is based on the empirical second moment of given raw signal x_1, \dots, x_n . First, we introduce the statistic which is a cumulative second moment of given sample:

$$C_x(t) = \sum_{i=1}^t x_i^2. \quad (3)$$

The $C_x(t)$ statistic was used in [19] as a base of the method applied in the segmentation problem in case when in real data we observe that some characteristics change with respect to time. This statistic was also a main point of the testing procedure whether in the given sample a structural break point exists or not.

In this paper we extend the methodology presented in [19] and propose to analyze the following statistic:

$$L_x(t) = \log C_x(t). \quad (4)$$

This choice is motivated by seismic recordings characteristics and discussion is carried out in further sections. As one can

expect L_x can tend to $-\infty$ if at least the first reading is equal to 0. In order to avoid this problem we modify the raw signal and in the further analysis instead of x_1, \dots, x_n we substitute the first reading x_1 with the first nonzero reading. This technical issue is related to a single sample at the very beginning of the recording; thus it does not influence the results. We denote corrected series as x'_1, \dots, x'_n .

Until the moment of P-wave arrival, the seismic recordings x_t consist of ambient noise which is considered stationary [21]; obviously they can be described by independent identically distributed Gaussian random variables. Moreover, we assume that the theoretical second moment of the distribution is finite. It can be shown that for data before the moment of P-wave arrival we have the following:

$$EL_{x'}(n) \sim \log(n). \quad (5)$$

Our methodology is therefore based on this observation. In the procedure, in contrast to [19], we fit the logarithm function $f_1 = a + b \log(x)$ to first k points of $L_{x'}(t)$ statistic. After the moment of P-wave arrival the character of the $L_{x'}(t)$ statistic changes. It is not exactly known what kind of function we can observe after the moment of P-wave arrival; however it was noted that in general the statistic is concave with respect to t . Here we decided to test two different concave functions: exponential $f_2 = c \exp dx + \text{const}$ and power $f_3 = gx^n + \text{const}$. These functions are fitted with time shift; that is, $f(t) = L_{x'}(t - k)$, $t = k + 1, \dots, n$. In order to reduce computational time we subtract $L_{x'}(n)$ or $L_{x'}(k + 1)$ and then fit the exponential or power functions, respectively. Fitted functions coefficients are obtained by using of Levenberg-Marquardt algorithm (LMA) [22, 23] which is an iterative algorithm used to solve nonlinear least squares problems. It combines features of Gauss-Newton method and the method of gradient descent [24]. The LMA algorithm requires at least 3 points to fit considered functions. The next step is to calculate the squared errors between $L_{x'}(t)$ and fitted functions. The estimated point of P-wave arrival is k for which the error is minimized.

Entire detection algorithm can be described as follows:

- (1) Set $k = 3$.
- (2) If $k = n - 2$ go to (7).
- (3) Fit $f_1 = a + b \log x$ to $\{L_{x'}(1), \dots, L_{x'}(k)\}$, $f_2 = c \exp(d(x - k))$ to $\{L_{x'}(k + 1), \dots, L_{x'}(n)\} - L_{x'}(n)$, and $f_3 = e(x - k)^n$ to $\{L_{x'}(k + 1), \dots, L_{x'}(n)\} - L_{x'}(k + 1)$.
- (4) Calculate $e_i^1 = \sum_{j=1}^k (f_1(j) - L_{x'}(j))^2 + \sum_{j=k+1}^n (f_2(j - k) - (L_{x'}(j) - L_{x'}(n)))^2$. Calculate $e_i^2 = \sum_{j=1}^k (f_1(j) - L_{x'}(j))^2 + \sum_{j=k+1}^n (f_3(j - k) - (L_{x'}(j) - L_{x'}(n)))^2$.
- (5) Set $k = k + 1$.
- (6) Go to (2).
- (7) $l_1 = \arg \min (e_k^1)$, $l_2 = \arg \min (e_k^2)$.

l_1 is "exponent" estimator, and l_2 the "power" one.

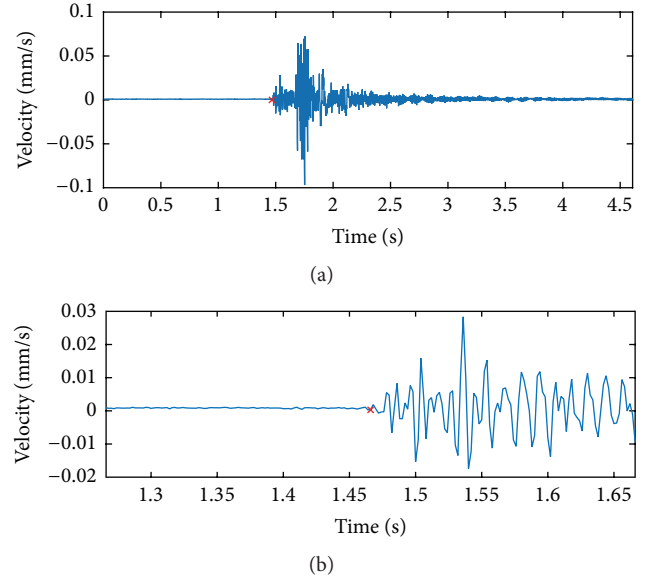


FIGURE 1: Exemplary seismic event with P-wave arrival determined by specialists (a) and its zoom (b).

3. Application to Real Seismic Data

In this paper the proposed algorithm was applied to a 188 single-event recordings from O/ZG "Rudna" underground copper ore mine. The signals were gathered by seismic system ELOGOR-C which is used to rock mass observation. The system consists of 2 sets of 32 seismometers Willmore MK-III type; each collects velocity data in the frequency band 0.5–150 Hz which is adequate frequency band containing mining-induced events. Such band is enough for localization, seismic energy estimation, and focal mechanism indication by analysis of first motion direction, which is the basic purpose of the monitoring system. The microseismic events in higher frequency are registered in this mine by a different system. The data is transmitted to seismic station using analog transmission (frequency modulation) and sampled with sampling frequency 500 Hz. Due to characteristics of the deposit, the seismic system network is relatively flat and a few additional sensors are located in shafts. Analyzed signals are dated from August 1, 2015, to August 19, 2015. The events length extent from 4.6 s to 33 s. Moments of P-wave arrival was indicated preliminarily using the STA/LTA algorithm and then manually corrected by seismic station experts.

In Figure 1 an exemplary seismic event is presented with moment of P-wave arrival marked by red cross. In Figure 1(b) zoom on the arrival time is shown. It is easy to spot stationarity of the background noise before the arrival of P-wave (red cross).

Application of $C_x(t)$ statistic can be seen in Figure 2.

As it was noted in [19], when applying $C_x(t)$ statistic to a stationary process with variance σ^2 , its expected value would be $EC_x(t) = t\sigma^2$. Seismic recordings before the moment of P-wave arrival (denote p) fulfill the stationarity assumptions.

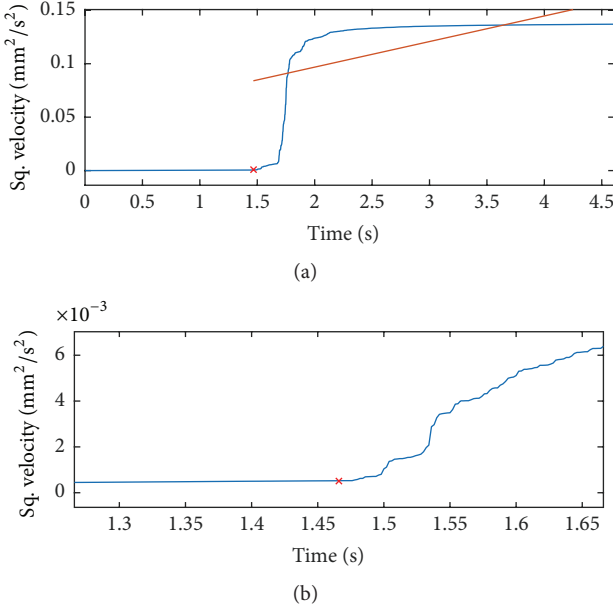


FIGURE 2: $C_{x'}(t)$ statistic applied to exemplary seismic signal presented in Figure 1 (a) and its zoom (b).

However, strict utilization of algorithm proposed in [19] cannot work properly, as after the P-wave arrives the series is not stationary. R^2 statistics were computed for linear fits from arrival point indicated by seismic station specialists to the end of recording (see example in Figure 2). The mean value of these R -statistics for the entire set of seismic records is 0.349 which is unacceptable. Thus the P-wave arrival indicated by using inappropriately fitted function might be false. Application of logarithm function to $C_x(t)$ statistics might highlight the P-wave arrival, since the structural change is sudden in case of $L_{x'}(t)$ statistic contrary to $C_x(t)$.

In Figure 3 the values of $L_{x'}(t)$ statistic are contained. With use of logarithm, the break point (Figure 3, marked with red cross) might be noticed much easier than it could be indicated from $C_{x'}(t)$ (Figure 2).

It is worth to notice that the $L_{x'}(t)$ statistic can be divided into 2 concave series, and the division point is located in the moment of P-wave arrival (marked with red cross).

Figure 3 presents the quality of fit. The average R^2 statistic for power function fit (fitted on the interval from the onset moment indicated by seismic station experts till the end of recording) is equal 0.976 and 0.974 for exponential function. This indicates that these functions appropriately approximate the $L_{x'}$ statistics.

3.1. Algorithm Results with Exponential Function Fitted. In Figure 4 results of the algorithm (exponential function fitted to the second part of statistic) are contained. Performed data analysis provides that 54.3% of algorithm picks do not differ by more than 10 samples (which corresponds to 0.02 s) from moments indicated by seismic station experts. 79.9% of differences do not exceed 50 samples (0.1 s). The largest difference is 177 samples (0.354 s).

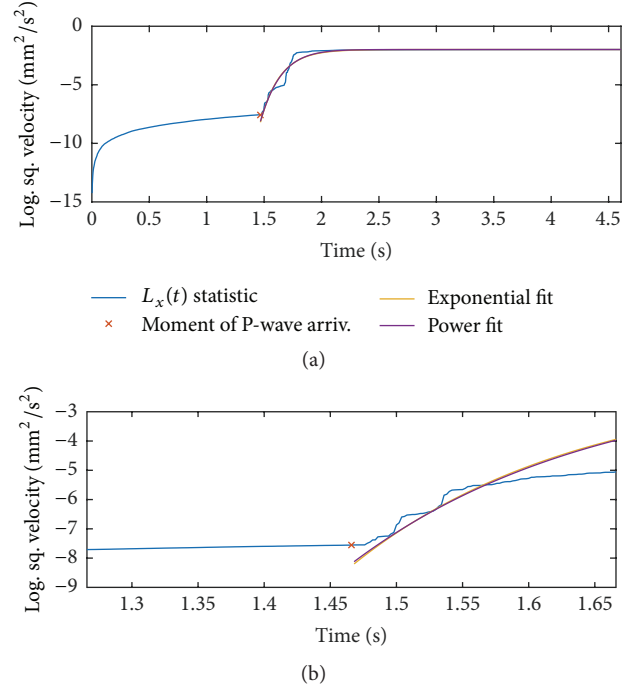


FIGURE 3: $L_{x'}(t)$ statistic applied to exemplary seismic signal presented in Figure 1 (a) and its zoom (b). Fit parameter $R^2 = 0.9557$ and 0.9591 for exponential and power function fitting, respectively.

3.2. Algorithm Results with Power Function Fitted. The results presented in Figures 4 and 5 provide that exponential and power fitting lead to similar results. Within the power function fitted 58.5% of differences do not exceed 10 samples (0.02 s) and 83% do not exceed 50 samples (0.1 s). The largest difference is 255 samples (0.51 s).

3.3. Results Based on STA/LTA. In order to examine performance of the proposed algorithms we compare them to P-wave picks obtained by the STA/LTA method with the most optimal parameters. The method requires predefined threshold ρ . The P-wave arrival is triggered when the STA/LTA ratio exceeds ρ . Also, the length of short and long time windows needs to be predefined. This is a drawback of this method, as optimal values of these quantities can change for different working conditions of sensors.

The algorithm was tested with different ρ values from 1 to 10 (step 0.05). Simultaneously, different lengths of short and long time windows were tested (short window from 10 to 200, step 10 samples, and long window from 10 to 400 samples more than the short one, step 10 samples). The most accurate estimated arrival moments were obtained with $\rho = 2.2$ and $s = 20$ samples and $l = 320$ samples.

In Figure 6 one can observe that results are significantly worse than that these provided by the novel algorithms based on second statistical moment. The analysis provides that 47.8% and 71.7% of picks do not exceed 10 and 50 samples, respectively. Additionally, STA/LTA algorithm missed 3 P-wave arrivals; algorithm proceeded through entire signal and without any moment triggered. The largest difference

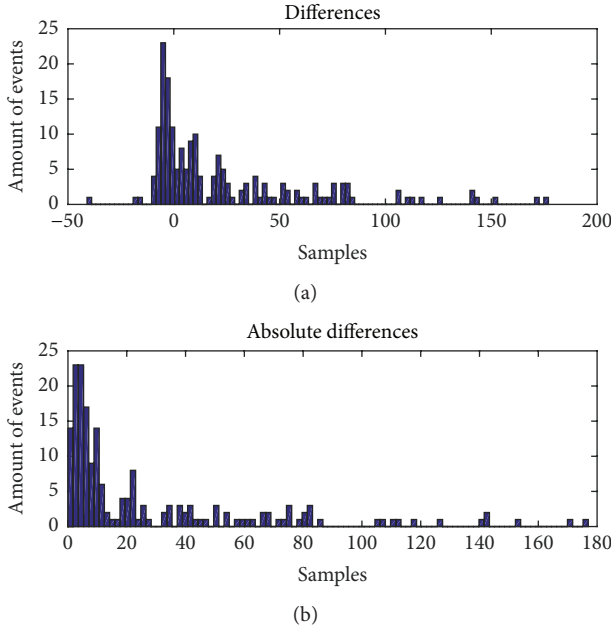


FIGURE 4: Histogram of differences between algorithm picks and those given by specialists: exponential function fitted (a) and histogram of absolute differences (b).

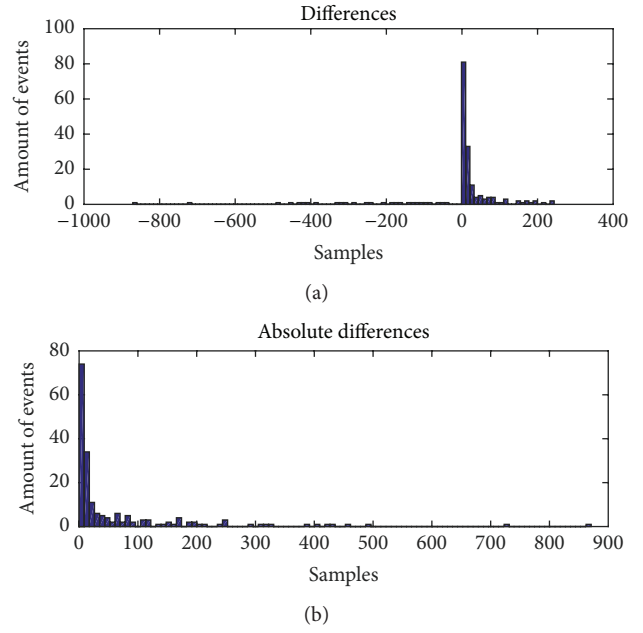


FIGURE 6: Histogram of absolute differences between STA/LTA picks and those given by experts (a) and histogram of absolute differences (b).

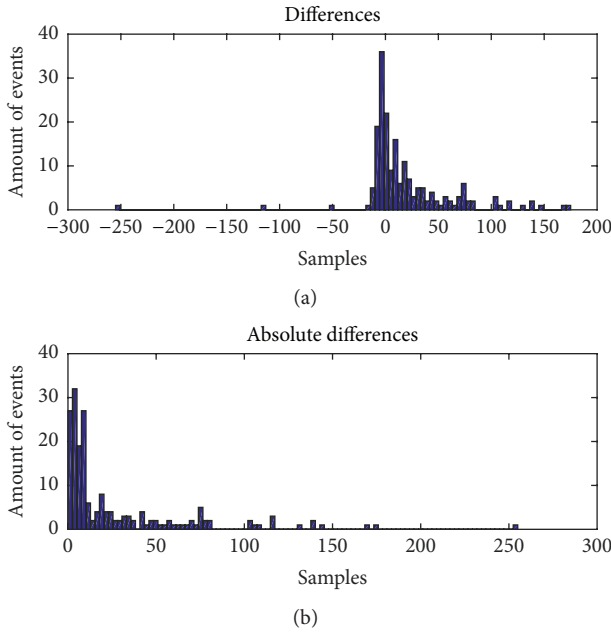


FIGURE 5: Histogram of differences between algorithm picks and those given by specialists: power function fitted (a) and histogram of absolute differences (b).

between algorithm results and seismic station specialists is 490 samples (0.98 s). Moreover, significantly more events are indicated before the actual moment of P-wave arrival. This shows that STA/LTA is prone to outliers.

In Table 1 basic statistics are included in order to compare three investigated methods. As it can be noticed the proposed algorithms outperform STA/LTA algorithm in all aspects,

TABLE 1: Comparison between algorithms based on second statistical moment and STA/LTA method.

	Algorithm (exp)	Algorithm (power)	STA/LTA
Number of correct (exact) picks	5	7	2
Mean of absolute differences	27.2935	27.2287	51.4505
Std. dev. of differences	39.2442	43.8035	101.3341
Std. dev. of abs. differences	36.6021	39.2816	87.2602
Mode of abs. differences	3	3	3

expect the mode of absolute differences which is equal to 3 for all methods. Fitting the power function provides the best results in terms of correct picks and mean of absolute differences but has worse standard deviations than the exponential fitting.

4. Conclusions

In this paper the regime switching detection method was adopted in order to find P-wave arrival. The algorithm was tested on seismic signal recordings from O/ZG "Rudna" underground copper ore mine. The results included in this paper show that the proposed algorithms are capable of indicating P-wave arrival moments as the estimated points were close with points manually indicated by mine station experts. The results were also compared to those provided by

widely used STA/LTA algorithm. Results of both fitting algorithms give better results than the results obtained using the STA/LTA method in terms of descriptive statistics. Moreover, the algorithm does not require any parameters nor calibration to proceed. Comparison of two proposed functions (power and exponential) does not provide clear answer which one is better. The power function leads to higher number of correct picks and lower mean of absolute differences but higher standard deviations of differences between estimated and actual picks.

Despite the fact that algorithms are considerably more accurate than STA/LTA, they still possess some drawbacks. The most noticeable concern is the computational complexity of the basic implementation, since for each signal many fittings (namely, twice the signal length) have to be performed. Additionally, STA/LTA can be computed online and provide information about event occurrence immediately. In case of the novel algorithms the response time might be minimized by considering only a short part of the signal registered after the P-wave arrival. The minimum number of samples acquired after the P-wave arrival that provides accurate pick is therefore of high importance. Moreover, algorithms for exponential or power functions fitting that update the fitted coefficients might significantly increase the computational speed. In future work the proposed methods might be upgraded in order to test whether the seismic event occurred or not. It is also interesting to assess the impact of better P-wave picking on localization accuracy. The method can also be tested for nonanthropogenic teleseismic events.

Competing Interests

The authors declare that there are no competing interests regarding the publication of this article.

References

- [1] W. Spence, "Relative epicenter determination using P-wave arrival-time differences," *Bulletin of the Seismological Society of America*, vol. 70, no. 1, pp. 171–183, 1980.
- [2] J. Boatwright and J. B. Fletcher, "The partition of radiated energy between P and S waves," *Bulletin of the Seismological Society of America*, vol. 74, no. 2, pp. 361–376, 1984.
- [3] G. Kwiatek and Y. Ben-Zion, "Assessment of P and S wave energy radiated from very small shear-tensile seismic events in a deep South African mine," *Journal of Geophysical Research: Solid Earth*, vol. 118, no. 7, pp. 3630–3641, 2013.
- [4] M. Withers, R. Aster, C. Young et al., "A comparison of select trigger algorithms for automated global seismic phase and event detection," *Bulletin of the Seismological Society of America*, vol. 88, no. 1, pp. 95–106, 1998.
- [5] R. Sleeman and T. van Eck, "Robust automatic P-phase picking: an on-line implementation in the analysis of broadband seismogram recordings," *Physics of the Earth and Planetary Interiors*, vol. 113, no. 1–4, pp. 265–275, 1999.
- [6] M. Leonard and B. L. N. Kennett, "Multi-component autoregressive techniques for the analysis of seismograms," *Physics of the Earth and Planetary Interiors*, vol. 113, no. 1–4, pp. 247–263, 1999.
- [7] H. Akaike, "Information theory and an extension of the maximum likelihood principle," in *Selected Papers of Hirotugu Akaike*, Springer Series in Statistics, pp. 199–213, Springer New York, New York, NY, USA, 1998.
- [8] R. V. Allen, "Automatic earthquake recognition and timing from single traces," *Bulletin of the Seismological Society of America*, vol. 68, no. 5, pp. 1521–1532, 1978.
- [9] M. Leonard, "Comparison of manual and automatic onset time picking," *Bulletin of the Seismological Society of America*, vol. 90, no. 6, pp. 1384–1390, 2000.
- [10] J. Wang and T.-L. Teng, "Artificial neural network-based seismic detector," *Bulletin of the Seismological Society of America*, vol. 85, no. 1, pp. 308–319, 1995.
- [11] S. Gentili and A. Michelini, "Automatic picking of P and S phases using a neural tree," *Journal of Seismology*, vol. 10, no. 1, pp. 39–63, 2006.
- [12] K. S. Anant and F. U. Dowla, "Wavelet transform methods for phase identification in three-component seismograms," *Bulletin of the Seismological Society of America*, vol. 87, no. 6, pp. 1598–1612, 1997.
- [13] H. Zhang, C. Thurber, and C. Rowe, "Automatic P-wave arrival detection and picking with multiscale wavelet analysis for single-component recordings," *Bulletin of the Seismological Society of America*, vol. 93, no. 5, pp. 1904–1912, 2003.
- [14] A. G. Hafez, T. A. Khan, and T. Kohda, "Earthquake onset detection using spectro-ratio on multi-threshold time-frequency sub-band," *Digital Signal Processing*, vol. 19, no. 1, pp. 118–126, 2009.
- [15] G. Xiantai, L. Zhimin, Q. Na, and J. Weidong, "Adaptive picking of microseismic event arrival using a power spectrum envelope," *Computers & Geosciences*, vol. 37, no. 2, pp. 158–164, 2011.
- [16] E. H. Ait Laasri, E.-S. Akhouayri, D. Agliz, and A. Atmani, "Automatic detection and picking of P-wave arrival in locally stationary noise using cross-correlation," *Digital Signal Processing*, vol. 26, no. 1, pp. 87–100, 2014.
- [17] D. Kucharczyk, A. Wyłomańska, J. Obuchowski, R. Zimroz, and M. Madziarz, "Stochastic modelling as a tool for seismic signals segmentation," *Shock and Vibration*, vol. 2016, Article ID 8453426, 13 pages, 2016.
- [18] R. Zimroz, M. Madziarz, G. Żak, A. Wyłomańska, and J. Obuchowski, "Seismic signal segmentation procedure using time-frequency decomposition and statistical modelling," *Journal of Vibroengineering*, vol. 17, no. 6, pp. 3111–3120, 2015.
- [19] J. Gajda, G. Sikora, and A. Wyłomańska, "Regime variance testing—a quantile approach," *Acta Physica Polonica B*, vol. 44, no. 5, pp. 1015–1035, 2013.
- [20] D. Kucharczyk, A. Wyłomańska, and R. Zimroz, "Structural break detection method based on the Adaptive Regression Splines technique," <http://arxiv.org/abs/1605.08667v1>.
- [21] S. J. Gibowicz and A. Kijko, *An Introduction to Mining Seismology*, vol. 55, Elsevier, New York, NY, USA, 2013.
- [22] K. Levenberg, "A method for the solution of certain non-linear problems in least squares," *Quarterly of Applied Mathematics*, vol. 2, no. 2, pp. 164–168, 1944.
- [23] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *Journal of the Society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963.
- [24] A. Cauchy, "Méthode générale pour la résolution des systemes d'équations simultanées," *Comptes Rendus de l'Académie des Sciences Paris*, vol. 25, pp. 536–538, 1847.

