Journal of Rock Mechanics and Geotechnical Engineering 14 (2022) 1406–1420

Contents lists available at ScienceDirect

Journal of Rock Mechanics and Geotechnical Engineering

journal homepage: www.jrmge.cn

Full Length Article

Quantitative assessment of the spatio-temporal correlations of seismic events induced by longwall coal mining

Shuyu Wang [a], Guangyao Si [a,*], Changbin Wang [a], Wu Cai [b], Binglei Li [c,**], Joung Oh [a], Ismet Canbulat [a]

[a] School of Minerals and Energy Resources Engineering, University of New South Wales, Sydney, 2052, Australia
[b] State Key Laboratory of Coal Resources and Safe Mining, School of Mines, China University of Mining and Technology, Xuzhou, 221116, China
[c] Zijin School of Geology and Mining, Fuzhou University, Fuzhou, 350108, China

ARTICLE INFO

Article history:
Received 18 June 2021
Received in revised form
2 March 2022
Accepted 14 April 2022
Available online 28 April 2022

Keywords:
Spatial correlation
Temporal correlation
Autocorrelation function (ACF)
Semivariogram
Scale of fluctuation

ABSTRACT

Rock failure process as a natural response to mining activities is associated with seismic events, which can pose a potential hazard to mine operators, equipment and infrastructures. Mining-induced seismicity has been found to be internally correlated in both time and space domains as a result of rock fracturing during progressive mining activities. Understanding the spatio-temporal (ST) correlation of mining-induced seismic events is an essential step to use seismic data for further analysis, such as rockburst prediction and caving assessment. However, there are no established methods to perform this critical task. Input parameters used for the prediction of seismic hazards, such as the time window of past data and effective prediction distance, are determined based on site-specific experience without statistical or physical reasons to support. Therefore, the accuracy of current seismic prediction methods is largely constrained, which can only be addressed by quantitively assessing the ST correlations of mining-induced seismicity. In this research, the ST correlation of seismic event energy collected from a study mine is quantitatively analysed using various statistical methods, including autocorrelation function (ACF), semivariogram and Moran's I analysis. In addition, based on the integrated ST correlation assessment, seismic events are further classified into seven clusters, so as to assess the correlations within individual clusters. The correlation of seismic events is found to be quantitatively assessable, and their correlations may vary throughout the mineral extraction process.

1. Introduction

Mining-induced seismicity is the response of rock mass to continuous mineral extraction. It represents either the initiation and propagation of new fractures or the slippage of pre-existing weak planes in rock mass induced by stress redistribution during mining. Mining-induced seismicity is usually controlled by the mining depth, mining speed, excavation geometry and geological discontinuities. Also, one or the combination of the above factors would lead to different seismic behaviours (Guha, 2000). Correspondingly, the inversion of seismic data collected at mine sites can reflect the characteristics of the above parameters to some extent (Bosch et al., 2010). For instance, the temporal variation of geomechanical properties of rock material can be inferred by the analysis of seismic data (Zhao et al., 2018). A clustering of seismic events could indicate local rock instability or/and substantial stress changes (Mendecki, 1999). Therefore, seismic monitoring has been used to optimise rock failure-related engineering designs (hydraulic fracturing) (Schultz et al., 2020) and predict potential rock failure and the induced seismic hazards (coal/rock bursts and gas outburst) (Zhao et al., 2018).

Based on the distribution of seismic events, seismic monitoring may contribute to predicting mining-induced seismic hazards. Mining induced seismicity does not distribute uniformly in space or time. In the space domain, most of the explosive types of seismic events caused by mining activities are energetically weak. In contrast, the events with high energy commonly occur in tectonic regions and are presumably caused by the interaction between tectonic stresses and mining-induced stresses (Stec, 2007). While

* Corresponding author.
** Corresponding author.
E-mail addresses: g.si@unsw.edu.au (G. Si), libinglei@fzu.edu.cn (B. Li).
Peer review under responsibility of Institute of Rock and Soil Mechanics, Chinese Academy of Sciences.

https://doi.org/10.1016/j.jrmge.2022.04.002

in the time domain, the seismic events tend to form nests, swarms and sequences (Gibowicz, 2009). Previous research indicates that seismic hazards are related to high-energy events near mine openings (Leśniak and Isakow, 2009; Cai et al., 2019). The occurrence of the rockburst is a result of a high energy storage within brittle-hard rocks under high stress, part of the energy will release as seismic energy and monitored by seismic waves (W. Cai et al., 2021). A direct relationship between seismicity and gas emission rate has been reported by Si et al. (2015), which can be used to provide early warning for uncontrolled gas emissions. Fault slip and seismic activities can be numerically simulated to comprehensively explore seismicity induced by mine extraction (Cao et al., 2018).

The difficulty of using a large amount of seismic data collected from mining operations for prediction purposes lies in the lack of understanding of the internal correlation between seismic events, as mining-induced seismicity is not a random process (Gibowicz, 2009) but has a high correlation with mining activities both spatially and temporally (Arabasz et al., 2005). The number and the released energy of the seismic events can directly indicate microfractures induced by mining activities (Song et al., 2019). Invalid prediction results or misleading data interpretation can be derived if the correlation is not well understood. For instance, during seismic data analysis, questions need to be addressed beforehand, such as how much past data (time window) are required to predict future events and the maximum distance that can be effectively predicted with confidence (grid size). The time window and grid size are essential parameters for investigating spatio-temporal (ST) evolutions of seismic events. An undersized time window may not be enough to reflect the general pattern of seismic events. An oversized time window may include unnecessary noisy data that reduce prediction accuracy (Kijko and Funk, 1996). Acoustic emission laboratory tests have proven an appropriate time window as the stress drop and the acoustic quiescence period within 10–30 min of the microfracturing process (Zhang and Zhou, 2020a; b; Zhou and Zhang, 2021) However, this paper seeks a correlated time window of a broader range that would contain seismic events from a large number of microfractures. Also, a too large grid may significantly reduce the resolution/accuracy of seismic hazard prediction in space (Kisilevich et al., 2010). A too small grid can increase computational time and cause overfitting issues. Therefore, determination of time window and grid size for the ST prediction of seismic hazard, respectively, remains a significant challenge using historical seismic data. In order to determine the appropriate time window and grid size, a correlation assessment on seismic data would be required in both the time and space domains.

The correlation analysis of mining-induced seismicity, including its randomness, stationary and memoryless, would provide an understanding of the past seismic data (Bischoff et al., 2010; González et al., 2016). Additionally, vital work has been conducted to prove the memory and the inter-relation of the seismic events information and discussed the importance of the correlation seismic data (Gibowicz and Lasocki, 2001; Lasocki, 2008; Olszewska et al., 2017). However, this paper mainly focuses on the quantitively assessment of the spatial, temporal and ST correlations of mining-induced seismicity in a systematic manner so far. The knowledge gap has largely affected the accuracy of using seismic data for dynamic hazard prediction in mines, e.g. rockbursts and outbursts (Cai et al., 2020; Cao et al., 2020; Si et al., 2020; Wang et al., 2021a), as well as the interpretation of mining-induced seismicity (X. Cai et al., 2021; Wang et al., 2021b; Duan et al., 2022). To address above research gap, we apply three statistical methods to quantitively assessing the correlation of seismic data in a case study mine:

(1) The autocorrelation function (ACF) has been proven to be a reliable indicator of the self-correlation of seismic parameters (seismic trace, seismic event location, time and magnitude), and it was used to obtain the degree of similarity of seismicity in time or space with itself (Węglarczyk and Lasocki, 2009). This paper will apply ACF to calculating the correlation with a delayed copy of the data itself, and equidistant data are required.
(2) The semivariogram is the basic geostatistical tool for measuring the correlation of a series of spatially regionalized data (Hohn, 1988). This paper applies semivariogram into both time and space seismic energy data to calculating the degree of correlation as a function of distance or time step.
(3) Moran's I (MI) is an index to describe the spatial similarity of a dataset, which is commonly used for cross-comparison and correlation threshold assessment. Tiefelsdorf and Boots (1995) suggested that MI is flexible for investigating the characteristics of the distribution and correlations for distinct spatial data. In this paper, MI is used to measure the correlation of seismic event energy extended in a specific time window.

These quantitative correlation assessment approaches can be applied to any parameters of mining-induced seismicity, including spatial location, onset time, energy, source radius and apparent stress. This paper will focus on radiated energy, which represents the total elastic energy radiated by mining activities and is better reflecting the influence on artificial structures compared to the magnitude and other parameters (Gibowicz and Kijko, 1994; Zhang and Zhou, 2020a; Zhou and Zhang, 2021).

Furthermore, many researchers proposed that seismic events can be divided into clusters due to the spatially distinct rock mass failure processes associated with the temporally dependent mining activities (Gibowicz, 1986; Leśniak and Isakow, 2009; Woodward et al., 2018). The seismic events from different clusters may be independent, whereas events within one cluster are internally correlated (Kijko and Funk, 1996). During a mining process, the overall correlation of the entire seismic dataset may be different from the correlation within individual clusters because the cluster-based data can be recognised as being related to a specific area or time. Thus, it is necessary to re-assess correlation characteristics within each cluster and between clusters after seismic data being clustered.

Seismic monitoring data collected from a Chinese coal mine is used in this study. Firstly, the correlative period and correlative distance of seismic data are calculated by the ACF and semivariogram function, respectively. The MI is used to evaluate the extent of the correlation and temporal variability. Using the results obtained from the above methods, an ST integrated analysis is conducted to examine the seismic correlation in time and space simultaneously. Finally, seismic events are divided into multiple clusters to investigate the local correlation within individual clusters.

2. Quantitative methods of correlation assessment

2.1. ACF

In this paper, the ACF is used to analyse time-series data of seismic energy. This method normally requires the same time interval between data points (an evenly spaced dataset). A gridding process is required to pre-process the unevenly spaced seismic data onset time. Therefore, the raw seismic energy data recorded with uneven time interval are calculated as cumulative daily energy, which has the same time interval. Assuming k is the lag in the time domain, the temporal variability of two seismic data points with a time difference of k can be calculated based on the autocovariance c_k and the autocorrelation ACF_k. The autocovariance c_k is the

covariance of the two seismic data x_i and x_{i+k} at the times i and $i+k$, respectively:

$$c_k = \frac{\sum_{i=1}^{N}(x_i-\mu)(x_{i+k}-\mu)}{N} \tag{1}$$

where N and μ are the number and the mean of the total studied data points, respectively. For an array of seismic data with lag k, its ACF is defined as

$$ACF_k = \frac{c_k}{c_0} = \frac{\sum_{i=1}^{N-k}(x_i-\mu)(x_{i+k}-\mu)}{\sum_{i=1}^{N}(x_i-\mu)^2} \tag{2}$$

where c_0 is the autocovariance when $k=0$, which is the self-covariance of x_i. ACF_k ranges from -1 to 1, and it shows the variation of seismic data correlation along with k. A typical ACF plot is shown in Fig. 1a. ACF_k equals 1 when k is 0, and it shows a downward trend with the increase of k. The seismic data array is regarded as correlated until the ACF_k falls below Bartlett's limit (I_B), which is expressed as (Jaksa et al., 1999):

$$I_B = \pm\frac{1.96}{\sqrt{N}} \tag{3}$$

The range of k before ACF_k reaching Bartlett's limit is called the correlative period. The highest correlative period of seismic data is presented when ACF_k reaches the upper limit of I_B, and the corresponding time lag is called the scale of fluctuation (SOF). For the lag larger than the SOF, it is considered that seismic data presents no correlation. Apart from the correlative period calculated using ACF, SOF can also be used to represent correlative distance, which is calculated using semivariogram (Onyejekwe et al., 2016) (to be introduced in Section 2.2).

2.2. Semivariogram function (V_S)

In order to quantitively evaluate the correlation of the unevenly spaced seismic data, semivariogram function is used here. Semivariogram is a graph showing the variation of semivariance with different lags. For a given lag k in the time or space domain, the semivariance V_S of a seismic data array is calculated as introduced by Clark (1979):

$$V_s = \frac{\sum_{i=1}^{N(k)}(x_i - x_{i+k})^2}{2N(k)} \tag{4}$$

where $N(k)$ is the number of data pairs separated by lag k, x_i represents the ith seismic datum, and x_{i+k} represents the paired seismic datum of x_i with a spatial or temporal interval of k. Semivariogram is the curve of the semivariance results at different lags fitted by selected mathematical models. A typical semivariogram is shown in Fig. 1b. The semivariance of the data array increases along with the lag increase until a maximum is reached at a certain lag. The increasing of the semivariances indicates the decline of autocorrelation. Following Three parameters are used to characterise the correlation of a semivariogram:

(1) Nugget , the semivariance when $k=0$;
(2) Sill, the maximum semivariance of the data array; and
(3) Range, i.e. SOF, the critical lag length for the semivariance to reach the sill.

In a semivariogram, lower nugget and sill indicate a higher correlation. A lower SOF suggests the faster attenuation of correlation along with the lag increase. The calculation of SOF varies slightly between different mathematical models. Table 1 lists three fitting models used in this research to calculate SOF.

2.3. MI

MI is an index to describe the spatial similarity of a dataset. For a seismic data array with N seismic events, its MI is defined as (Tiefelsdorf and Boots, 1995):

$$MI = \frac{N}{\sum_i\sum_j w_{ij}}\frac{\sum_j w_{ij}(x_i-\bar{x})(x_j-\bar{x})}{\sum_i(x_i-\bar{x})} \tag{5}$$

where $\bar{x}$ is the mean of the seismic data array; x_i and x_j are the seismic data of events i and j, respectively; and w_{ij} is a matrix of spatial weights. In this equation, spatial weights are calculated based on the inverse distance weighting of k-nearest points (10 points are selected in this paper).

MI can also be calculated graphically using MI scatter plot shown in Fig. 1c. In this figure, the horizontal axis is $(x_i-\bar{x})$ for all events. The vertical axis shows the difference between the average of the above 10 nearest seismic data (x_i) and $\bar{x}$. MI is then calculated as the linearly fitted line slope for all data points across the origin, which is shown as the red line. As the calculation of the slope in this figure is essentially the same as Eq. (5), identical MI results will be achieved using this graphic method or Eq. (5). In Fig. 1c, the four quadrants demonstrate different correlation conditions of seismic

Table 1
Mathematical models available for semivariogram fitting.

Model	Fitting function	SOF, θ
Gaussian	$G(x) = C(1-e^{-k^2/a^2})+C'$	$\theta = \pi^{0.5}a$
Spherical	$G(x) = C\left(\frac{3k}{2a}-\frac{k^3}{2a^3}\right)+C'\ (k<a)$ $G(x) = C+C'\ (k>a)$	$\theta = 3a/4$
Exponential	$G(x) = C(1-e^{-k/a})+C'$	$\theta = 2a$

Note: a, C and C' are the fitting parameters.

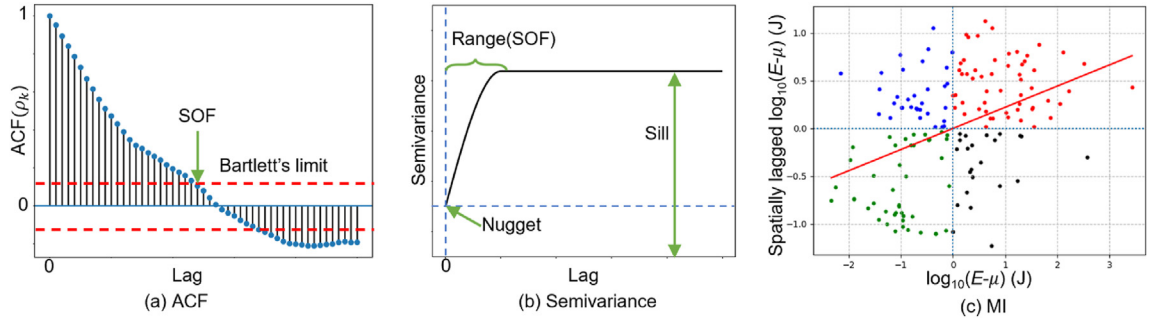

Fig. 1. An example plot of (a) ACF, (b) semivariogram, and (c) MI.

events. The first quadrant indicates that high energy seismic events are clustered with high energy neighbouring events, and the third quadrant indicates that low energy seismic events are clustered with low energy neighbours. Seismic events in the second and fourth quadrants show that low energy events are sitting close to high energy neighbours and high energy events are close to low energy neighbours, respectively.

3. Background of the study mine site

Hujiahe Coal Mine is located in the west of Shaanxi Province, China. The studied panel, longwall (LW) 102 is about 700 m in depth, 1493 m in length and 180 m in width. The target coal seam has a thickness varying from 13 m to 22.5 m, and the maximum dip angle is 9°. A fully mechanised longwall top coal caving method is adopted, with 3.5 m mined by shearer and the remaining coal at the top extracted by gravity caving. The coal seam is sequentially overlaid with a 5.95 m thick sandy mudstone layer, a 23.7 m thick siltstone, a 4.65 m thick mudstone, and a 4.8 m thick siltstone.

The 16-channel "ARAMIS M/E" seismic monitoring system developed by EMAG in Poland was installed in the mine in September 2013. For more information about this monitoring system, please refer to Cai et al. (2019). LW102 started to retreat in May 2014 and completed in July 2015. Although the seismic monitoring system was installed before the start of the panel, due to the calibration delay in the early stage, the system was only available to record reliable seismic activities in the panel from September 2014. Therefore, seismic data from September 2014 to July 2015, a total of 293 d, were used for the ST correlation analysis in this research.

In this paper, the seismic dataset contains 14,024 seismic events, and each seismic event is a five-dimensional array including the three-dimensional spatial location (longitude, latitude and depth), onset time and recorded seismic energy. For longwall coal mining, the majority of geophones are deployed on a horizontal planar surface as constrained by the tabular orebody. This yields a relatively high accuracy for locating seismic events in the horizontal direction (x and y coordinates), which is within 5–10 m in this research given the coverage of geophones. On the other hand, for the vertical direction, due to limit access to install geophones at large vertical variations, the accuracy for the z coordinate is estimated at 25 m. Monitoring results show that most of seismic events were recorded within 40 m above the mining level, which also suggested low vertical variation of seismic events. Due to the lower accuracy and variance compared to the longitude and latitude, the depth of each seismic event is not considered in this research.

4. Exploratory analysis of seismic data

To ensure the completeness of the seismic data array during the study period, the magnitude of completeness (m_c) is applied to determining the lowest energy magnitude of seismic events that the seismic monitoring system can fully detect. Only seismic events with energy magnitudes larger than m_c are regarded as complete and used for further analysis. Fig. 2 shows the probability density function of recorded seismic events over the monitoring period in LW102, where the m_c is found to be at $\log_{10}E = 2.3$. Thus, a total of 8024 seismic events with $\log_{10}E > 2.3$ are selected for the ST correlation analysis.

Fig. 3a shows the spatial distribution of the selected seismic events in LW102. Note that only 7 out of 16 geophones are shown here. The rest of geophones are far away from the study longwall panel and did not receive high quality data for further analysis. A large number of events with higher energy magnitudes are located around Fault 5–6. Also, in Fig. 3b, the contour map of seismic event probability density indicates that intensive seismic activities were reported at the tailgate side of the panel due to the nearby goaf zone. The detailed monthly evolution of seismic data over the monitoring period can be found in Appendix A.

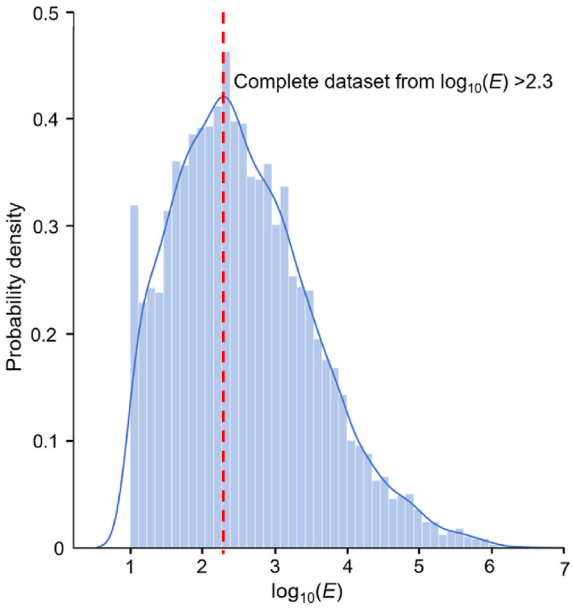

Fig. 2. Probability density plot of $\log_{10}E$ for all recorded seismic events.

5. Assessment of ST correlations of seismic energy during longwall mining

5.1. Temporal correlations

In this section, both ACF and semivariogram are used to investigate the temporal correlation of the seismic data in LW102. ACF is used to assess the temporal correlation of an evenly spaced seismic data array (cumulative daily seismic energy). Semivariogram is used to investigate the temporal correlation of non-evenly spaced seismic data (seismic energy per event). The SOF in the time domain (SOF-time) is determined for both ACF and semivariogram to quantify the correlative period of the seismic data.

5.1.1. ACF

As mentioned in Section 2.1, the ACF analysis is applied to the evenly spaced data with the same interval. The raw seismic events are pre-processed into daily cumulative data. A total of 293 data points representing the cumulative energy within 293 monitoring days are applied to calculate ACF_k by Eqs. (1) and (2), where x is substituted by $\log_{10}E_c$, in which E_c is the cumulative daily energy. The Bartlett's limit is calculated by Eq. (3).

Fig. 4 shows the ACF results of cumulative daily seismic energy in LW102 with different lags. In this figure, the I_B is ±0.1145 shown as the red dashed lines, and the SOF-time is calculated as 6 d, which indicates a relatively strong correlation of cumulative daily energy in 6 d. In other words, the future daily seismic energy in LW102 is likely to be dependent on the recorded seismic energy in the past 6 d. It will be independent with any data beyond that period. The ACF of daily seismic energy drops to 0.44 when the lag is only 1 d, which shows the increase of randomness.

Besides, the value of cumulative daily energy is also affected by the number of events that occurred during the day and face advance distance on that day. Therefore, the daily average energy and energy per meter of face advance are also applied in the ACF analysis. The average daily energy can be calculated as $\log_{10}(E_c/N)$,

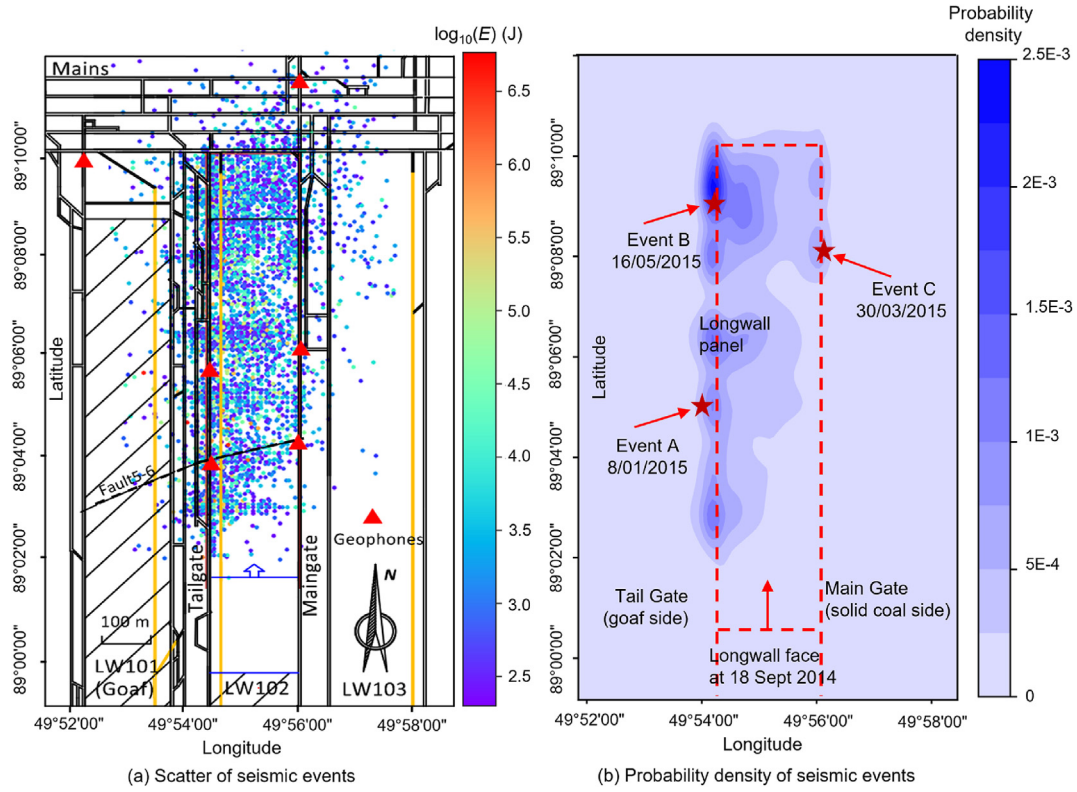

Fig. 3. (a) Spatial distribution of seismic events (colour represents $\log_{10}E$), longwall layout and geophone stations, and (b) The probability density distribution of seismic events in a horizontal plane. The red dashed line shows the longwall panel, and Event A, B and C are three typical event locations, which will be discussed in Section 6.

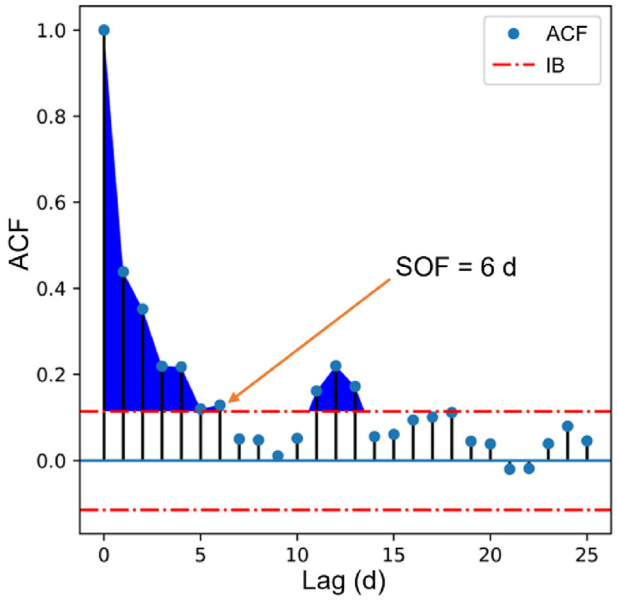

Fig. 4. ACF plot for cumulative daily energy.

where N is the number of seismic events that occurred on that day. The average energy per metre can be calculated as $\log_{10}(E_c/F)$, where F is the face advance distance on that day.

The ACF values of the above two parameters are presented in Fig. 5. The SOF-time of average daily energy decreases to 4 d, and the SOF-time of average energy per metre remains at 6 d compared to Fig. 4. The similar SOF-time calculated from Figs. 4 and 5 indicates that the number of events and the face advance rate only have a marginal effect on the autocorrelation of cumulative energy. One possible reason could be that $\log_{10}(E_c/N)$ is equal to $\log_{10}E_c - \log_{10}N$, and comparing with E_c, the value of N is much smaller in orders of magnitude. The same reason also applies to the face advance rate F. In other words, if one seismic event has a very high energy release in a typical production day, it will significantly increase the average value and decrease the correlation of cumulative energy.

Therefore, the correlation of cumulative energy analysed via ACF provides an approach to quantitively assess the correlation of seismic event energy. The event number and face rate can also be considered and analysed. But the result is sensitive to high energy events. The method in this section is suitable for a zone with similar size of fractures that can trigger seismic events with comparable energy levels.

Apart from the daily cumulative seismic energy, the daily seismic event number can also be analysed via ACF. Fig. 6 shows the ACF results of the daily event number. It indicates that the SOF-time of the daily event number is 26 d, given the I_B at ±0.1145, representing a strong correlation of daily seismic event number within 26 d. Also, the ACF of the daily event number is larger than 0.5 when the lag is lower than 4 d. It indicates that the daily event number is significantly influenced by those event numbers recorded in the past 4 d. Due to the strong correlation in the low lag of seismic event number, seismic event number prediction would yield more reliable results than that of cumulative energy.

5.1.2. Semivariogram function

(1) Overall semivariogram evaluation

Unlike ACF dealing with evenly spaced seismic data (cumulative daily energy), the semivariogram function is applied to assessing the temporal correlation of the unevenly spaced seismic data, such

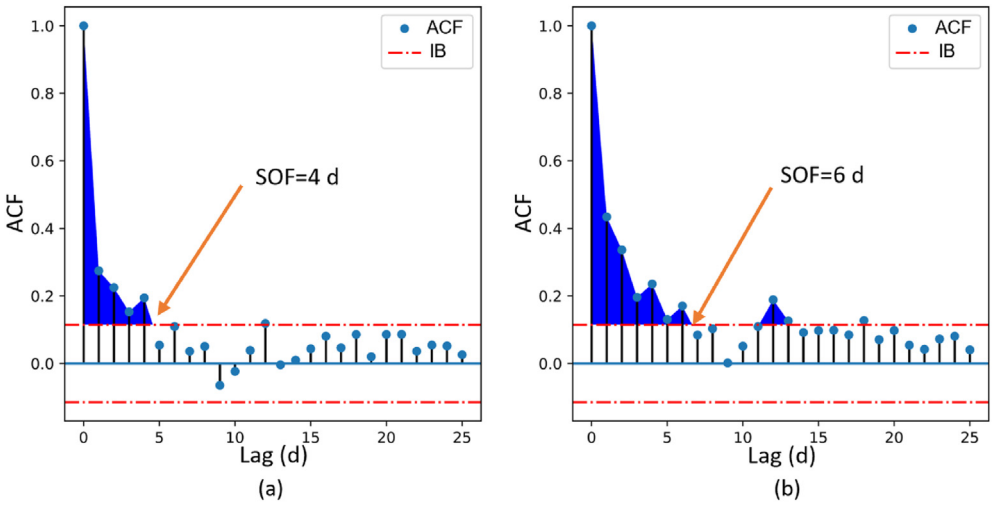

Fig. 5. ACF plot for (a) the average daily energy and (b) the average energy per metre of face advance.

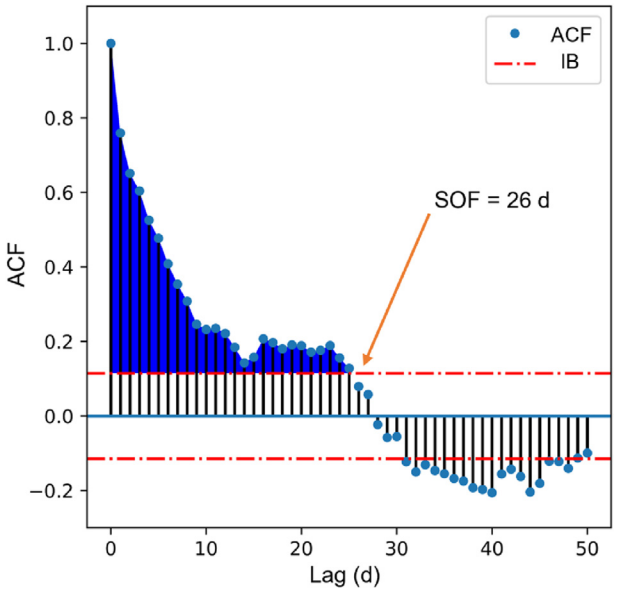

Fig. 6. ACF plot for daily event number.

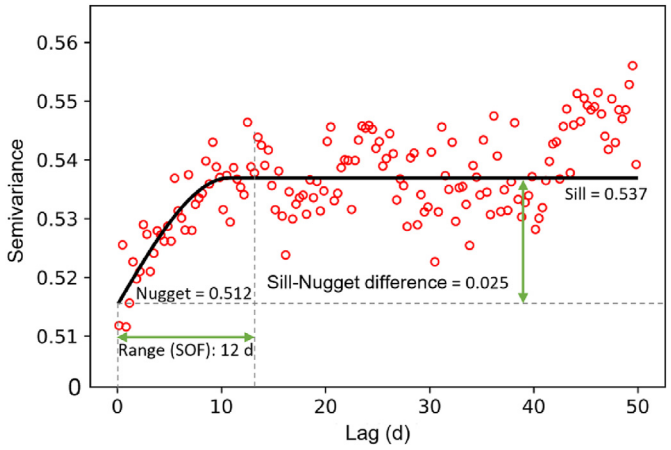

Fig. 7. Semivariogram of seismic energy. The red circles are the semivariance values at each time lag, and they are fitted by the exponential function as presented in the black curve.

Table 2
Comparison of three fitting models for the temporal semivariogram of seismic energy.

Model	*RMS*	R^2
Gaussian	0.00642483	0.362562
Spherical	0.00641485	0.364542
Exponential	0.00635507	0.37633

as the onset time of individual seismic events. The algorithm of this method can be referred to as Section 2.2. The semivariances of seismic energy based on time lag are calculated and presented in Fig. 7 as red circles. To obtain the semivariogram function, a goodness-of-fit test is conducted to select the best-fit model from three mathematical models, i.e. Gaussian, spherical and exponential. The fitting performance of each model is evaluated by root-mean-square (RMS) and coefficient of determination (R^2). These results are summarised in Table 2, indicating that the exponential model best fits the semivariances of the studied seismic data because of the lowest RMS and the highest R^2. It should be noted that the exponential fitting is the best-fit model when analysing the temporal correlation using all seismic energy data. The best-fit model may vary depending on the type and amount of seismic data inputted in this method.

The black line in Fig. 7 shows the semivariogram of the seismic energy of the studied events in LW102. The SOF-time of about 12 d indicates that the seismic energy of individual events is becoming less correlated along with the lag increase within 12 d. The difference between the sill and nugget is 0.025, which is significantly lower than the nugget, which implies a rapid decline of the correlation and high variability of seismic energy in a short time range.

It should be noted that the SOF-time of 12 d correlative period is calculated based on the unevenly spaced seismic energy data. In comparison, the SOF-time of 6 d correlative period obtained by ACF is based on the daily cumulative seismic energy data. Compared with the semivariogram method, ACF requires pre-processing unevenly spaced seismic energy as cumulative daily energy, which may introduce an artificial effect or bias to the correlation analysis. Thus, it is believed that SOF-time calculated by the semivariogram function can better reflect the correlation nature of raw seismic data. An application of the 12 d SOF-time will be presented in Section 5.2.

(2) The maximum correlative period determined by semivariance

The SOF-time calculated by semivariogram is the average correlative period of all seismic energy data in the study period, which means that if the correlative period of partial data is assessed, the result could fluctuate around 12 d. Besides the average correlative period of the overall data, the maximum correlative period is critical to be evaluated. This could be achieved by the semivariances calculation based on all the seismic data with a time difference smaller than the lag k, rather than using the paired seismic data x_{i+k} in Eq. (4).

This specific semivariance is called the cumulative semivariance, which represents the evolution of semivariance along with the increasing of input data. Fig. 8 shows the cumulative semivariance of the studied seismic data (marked as red crosses), which increases rapidly at the beginning and then flattens. The first-order derivative of the cumulative semivariance (marked as blue dots) is also calculated. The derivative has a gradual downward trend and tends to reach the elbow point when lag k is at around 40 d. The slope drops to around 0 for the first time at this elbow point. It suggests that 40 d is the maximum correlative period of studied seismic energy data. When lag k is larger than 40 d, most of the first-order derivative is close to 0 due to the slow increase of the cumulative semivariance. It means that little difference can be made if the semivariance considers the data with more than 40 d. Therefore, the correlation analysis of seismic data within 40 d can capture the general correlation characteristics over the period.

(3) The evolution of seismic data correlation in the time domain

Besides the average correlative period and the maximum correlative period, an attempt has been made to analyse the evolution of the temporal correlation of the studied seismic data over the 10 months of the monitoring period, the semivariogram function is used to assess the correlation of seismic data recorded within sequential periods. A moving time window based on the maximum correlative period is defined here, which can sequentially select seismic data within that time window. The moving window is large enough to capture the inherent correlation of seismic data within the selected period. Still, it should not be too large to ensure enough sequential periods to reflect the temporal variation of seismic data correlation.

According to the maximum correlative period calculated above, semivariogram analysis can be applied to every moving time window with 40 d of seismic data. The moving step is set as 1 d. The SOF-time, nugget and sill are sequentially calculated by the best-fit from Gaussian, exponential and spherical models for each time window. The variations of SOF-time selected by the best-fit and its RMS are shown in Fig. 9. It should be noted that the trend of how SOF changes at a different time is more critical than the exact magnitude of SOF because of the variability of fitting quality in each period. As expected, the temporal correlation of seismic energy is not a constant value but varies with time. The trend of the correlative period evolution shows a periodic distribution during the whole process of the coal extraction, and it tends to form seven peaks with different sizes. It should also be noted that before March 2015, the SOF shows a significant variance with a relatively high RMS, indicating the low quality of the semivariogram fitting. After March 2015, a more reliable SOF can be derived due to the lower RMS.

5.2. Spatial correlations

MI and the semivariogram are used to investigate the spatial correlation characteristics of the seismic energy. The degree of spatial dependence between seismic events is quantified. Only the horizontal locations of the studied seismic events are used here due to the location errors in the vertical direction of seismic data recorded in tabulate coal deposits. The SOF in the space domain (SOF-space) is determined based on semivariogram to quantify the correlative area of the seismic data.

5.2.1. MI

MI is applied to assessing the extent of spatial correlation and identifying a strong spatial correlative period based on a moving time window method. As mentioned in the previous section, the correlative period of the studied seismic data is about 12 d, which is the average SOF-time of unevenly spaced seismic energy data calculated in Section 5.1. Therefore, the time window of MI calculation is set as 12 d. The reason for not using the maximum correlative period is that MI does not require the moving time window to be large enough to reflect the general correlation characteristics over a long period. The variation of the spatial correlation is more critical in this section. Also, since the analysis should be based on randomisation assumption and conduct null hypothesis testing, the MI results are further standardised as the Z-value (Eq. (6)). A P-value less than 5% is considered significant (reject the null hypothesis), suggesting that seismic events are spatially autocorrelated on the global scale.

$$Z = \frac{MI - E(MI)}{\sqrt{\mathrm{var}(MI)}} \tag{6}$$

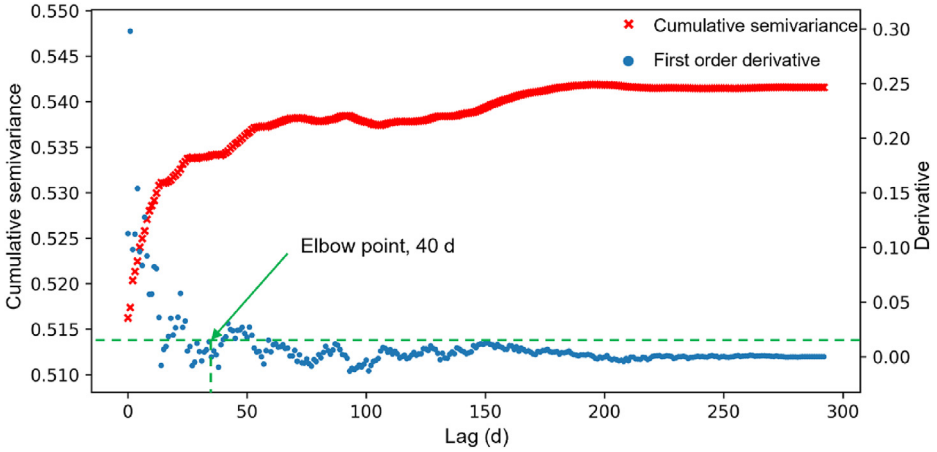

Fig. 8. Cumulative semivariance (marked as red cross) and its first-order derivative (marked as blue dot). The green arrow represents the elbow point. The green dashed line represents its lag of 40 d and cumulative semivariance of 0.514.

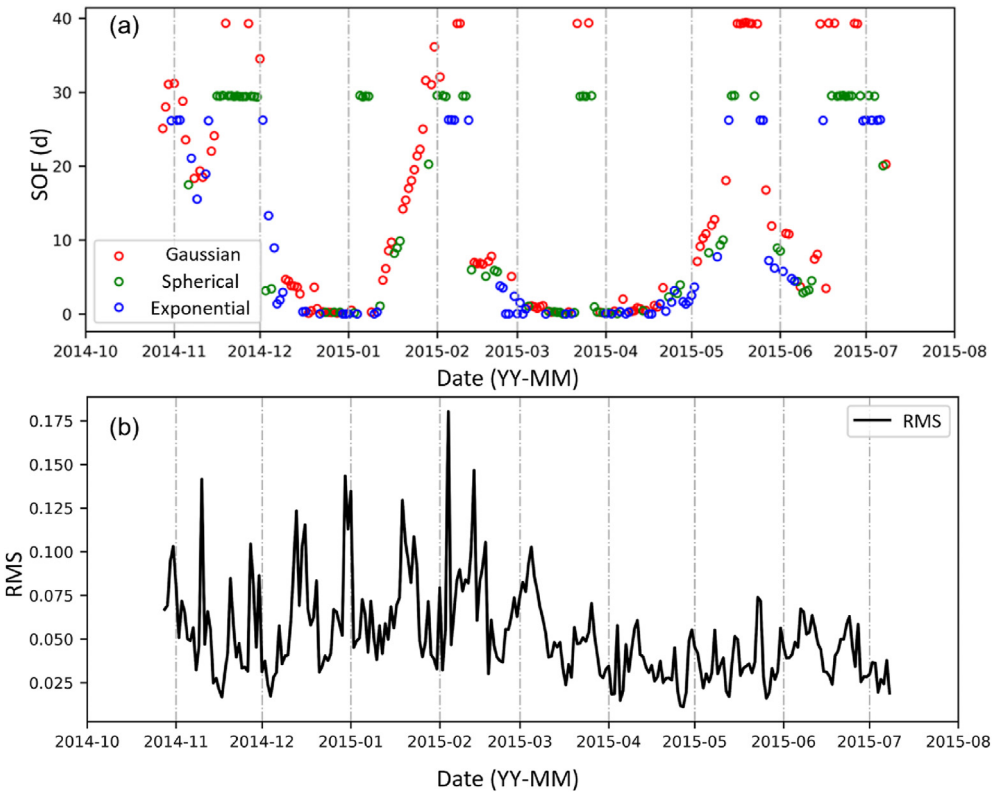

Fig. 9. (a) The evolution of SOF-time in each time window calculated by semivariogram using the best-fit method from Gaussian, spherical and exponential models. (b) The RMS of semivariogram fitting in each time window.

where $E(MI)$ means the expectation of the MI, and var(MI) means the variance of MI.

Fig. 10 shows the plot of MI and Z for seismic events recorded over the monitoring period. The MI value is calculated by both the mathematic equation (Eq. (5)) and the graphic method mentioned in Section 2.3. These two methods show identical MI values. The evolution of MI value indicates various correlation degrees of seismic events in space when LW102 is retreating. Except for one date below the confidence limit, most of the Z-values are higher than the 5% confidence interval limit, indicating that seismic energy are spatially correlated at various degrees, which supports the concept of seismic event prediction in space.

To examine the spatial energy distribution when different MI values are detected, the seismic event distributions at two typical periods (Period I and Period II) with $MI > 0$ and $MI = 0$, respectively, are presented in Fig. 11. Fig. 11a and b shows the scatter plot of MI values calculated using the graphic method for Period I and Period II. Fig. 11c and d illustrates the spatial distributions of the seismic events at these two periods. According to Fig. 11a and c, from 31 October 2014 to 12 November 2014 (Period I), the MI reaches the maximum of the entire monitoring period. The seismic events have a clear trend of high energy events clustered together in space, and so as the low energy events. From 9 February 2015 to 21 February 2015 (Period II) as Fig. 11b and d shows, MI is close to 0, which means no spatial correlation of seismic energy among these events. By visually inspecting the plot of the seismic event distribution in Period II in Fig. 11c and d, the same conclusion can be achieved: the high energy events and low energy events are distributed randomly in this two-dimensional (2D) horizontal space.

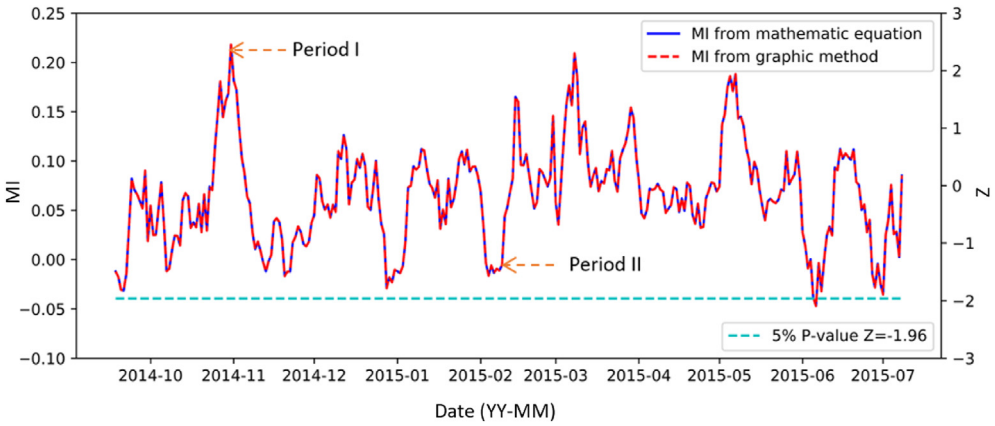

Fig. 10. MI values calculated from two different methods (in the left axis) and Z-value (in the right axis) over the monitoring period. A 5% P-value of Z is drawn as the blue dashed line.

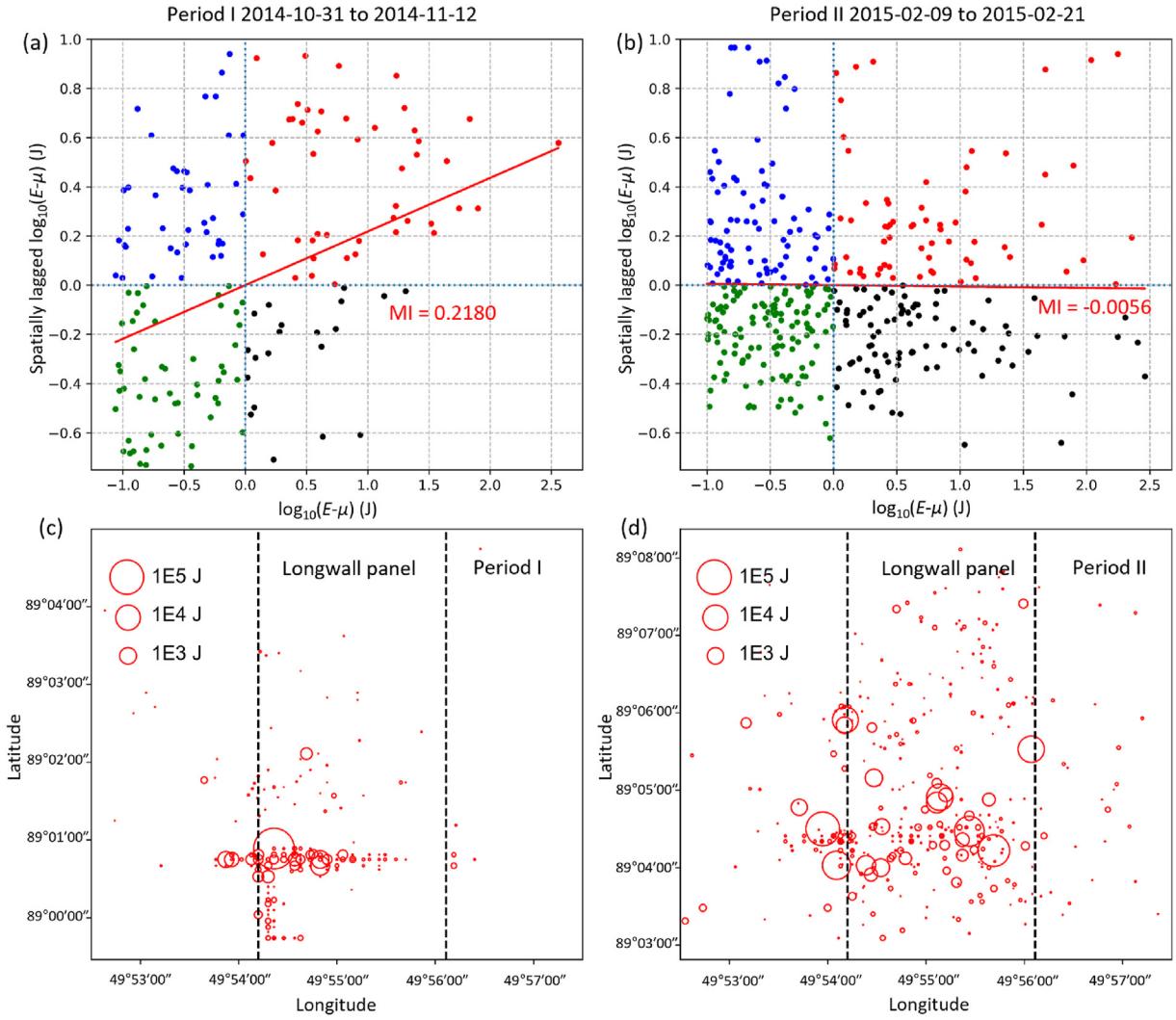

Fig. 11. Typical MI calculated using the graphic method for (a) Period I and (b) Period II. The red line is the linear fitting curve, and the MI value is measured by the slope of this line. Raw spatial distributions of seismic events for (c) Period I and (d) Period II, and size of the circle indicates the seismic energy.

5.2.2. 2D spatial semivariogram

In order to obtain the correlative distance (radius), the method used here to measure the correlation from point to point is the 2D spatial semivariogram, which can quantify the degree of spatial dependence between samples in a specific orientation and assess the degree of attribute's continuity. The algorithm of the semivariogram can be referred to as Section 2.2.

In Eq. (4), x_i indicates the ith seismic event energy and lag k represents the radius of the searching circle in the calculation of spatial semivariogram. Table 3 shows the evaluation results of three fitting methods and the best-fit method is the exponential function. Fig. 12 presents the spatial semivariogram plot for the LW102 working face, in which the calculated semivariance is shown in red circles and the fitted exponential model in black curve. The correlative distance is defined as the SOF-space, which is 23 m in this case. The released energy of a seismic event has a gradually decreasing correlation along with the increase of the distance from its hypocentre within 23 m radius, and no correlation presents beyond this radius. The difference between the sill and nugget is 0.127, which is relatively higher than the temporal semivariogram but still much smaller than the nugget and sill. It implies a rapid decline of the correlation and high variability of seismic energy in a short distance.

Table 3
Comparison of three fitting models for the spatial semivariogram of seismic energy.

Model	*RMS*	R^2
Gaussian	0.018912	0.615597
Spherical	0.017715	0.662701
Exponential	0.017561	0.668547

6. ST correlation assessment using reference seismic events

To simultaneously investigate the ST correlation around a reference location and time registered by the occurrence of a reference seismic event in LW102, the distance and time differences between the reference event and other seismic events need to be considered together. The calculation of distance only considers longitude and latitude coordinates for the same reason mentioned in Section 3. Three seismic events were selected as the reference events in this research. These selected seismic events all have relatively high energy and are located close to the longwall panel because the high energy events in the coal extraction process always represent a large formation of discontinuities and massive energy release. It is worth noting that the method proposed in this section could be used in any reference events with a specific onset time and location. The selected Event A has the seismic energy of 2.6 MJ, which is located around Fault 5–6. Event B is located near

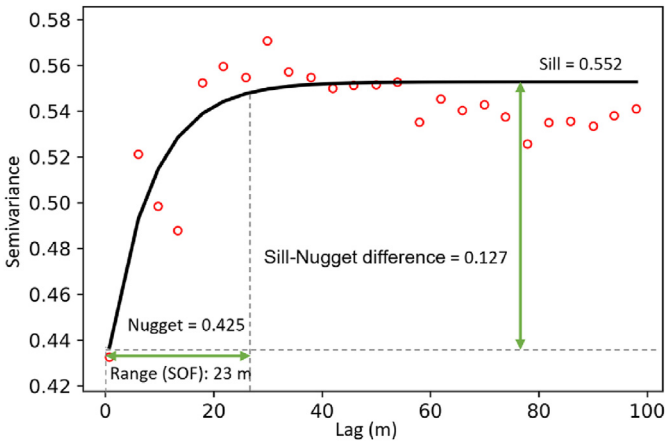

Fig. 12. Spatial variogram plot for seismic energy over the monitoring period. The red circles are the semivariance. They are fitted by the exponential function as presented in the black curve.

the tailgate (goaf side) with seismic energy of 1.4 MJ. Event C is a seismic event with an energy of 280 kJ located near the maingate (solid coal side). The locations of these three events, as well as their occurring time, are highlighted in Fig. 3.

Based on Events A, B and C, the time difference and distance between all other events and the reference events are calculated, which are shown as scatter plots in Fig. 13a. The *y*-axis shows the Euclidean distance difference, while the *x*-axis shows the time difference. Based on the data in Fig. 13a, to calculate the semivariogram for both time and distance simultaneously, a unity-based normalisation needs to be applied to these two parameters. The maximum distance is taken as 800 m, and the maximum time difference is taken as 300 d. The corresponding semivariogram is calculated and presented in Fig. 13b, following the method introduced in Section 2.2. Herein, the property x_i represents the event energy, and lag k is a unity-based normalisation of the time difference and distance. The exponential function is used because it is proved in Section 5.2.2 as the best-fit method when dealing with all studied seismic data. The lag k in the semivariogram plots in Fig. 13b contains the information conveyed in both the time and space domains.

The distance-time difference plots in Fig. 13a show that a linear relationship between distance and time differences can be observed for the reference events. The semivariogram plots in Fig. 13b all show an upward trend before levelling off. The parameters of the three semivariogram plots are also similar, with nugget at around 0.4 and sill at 0.54. The similarity of the nugget and sill for all three reference events indicates that the extent of ST correlation close to or far from the reference events is similar and less affected by the specific location or onset time of the reference events.

Based on the semivariogram plot, the SOF of time difference and distance can be obtained by inverting the process of the unity-based normalisation. The SOF-time for the three reference events A, B, and C are 2 d, 4 d and 2 d, respectively, and the SOF-space are 4.9 m, 9.2 m and 3.9 m, respectively. Compared to the SOF-time of 12 d in Section 5.1.2 and SOF-space of 23 m in Section 5.2.2, the SOF obtained by ST correlation analysis is much smaller in both the time and space domains. This is because it is less likely to have seismic events with strong ST correlation at the same time.

Compared with the SOF-time and the SOF-space in Section 5, The SOF in ST correlation gives relatively stable correlation results when assessing the correlation characteristics around the three reference events, representing a local correlation rather than an overall value. Thus, more holistic results were derived by the SOF in ST correlation, which considers the correlation both in time and space. The nugget and sill are very similar, and the SOF only has a marginal difference. This conclusion seems to be the opposite of the conclusion that the correlation is variational in the time and space domains, as discussed before. But in fact, the ST correlation assessments are designed to detect the overall correlation trend. In contrast, ST correlation chooses three reference points and assesses the correlation by taking the three points as a basis. The correlation near the reference points (nugget) and the correlation far from the reference points (sill) are purely based on the location and onset time of three reference events. The three points all have relatively high energy, which indicates that they might be induced by the slipping of pre-existing fractures. Besides, the three locations all have a similar event density, energy distribution, and even seismicity source mechanism.

7. ST correlation of seismic events in individual clusters

7.1. ST cluster based on face advances

The ST correlation analysis in Section 5 demonstrates the overall correlation of all seismic events in LW102. However, as seismic events from similar sources are more likely to be clustered in time and space, the correlation result of seismic events in one cluster may be interfered by events in other clusters. A seismic event in a cluster commonly presents a significant correlation with other seismic events in that cluster but shows independence to the evens that belong to other clusters (Kijko and Sciocatti, 1995). To remove the interference between different seismic event clusters and explore the correlation within individual clusters, a ST-based clustering method is used.

According to the data of LW102 face position at each production date, the distances of individual seismic events to the longwall face at the time of being recorded can be calculated (hereafter referred to as face-event distance). Fig. 14 shows the boxplot of face-event distances on each production date over the monitoring period in LW102. In this figure, each box shows the face-event distance distribution within the day. The coloured box ranges from the 25th percentile and 75th percentile, and the transverse line within the box indicates the median of the face-event distance.

Fig. 14 shows that seismic events generally first presented at around 200–300 m ahead of the face, and as the progressive advance of the face, the face-event distance reduced to about 50 m. As shown in Fig. 14, each cluster of seismic events was initiated far away from the face, indicating that the regional abutment stress change caused by coal extraction can disturb far-field stress. Within each cluster, the repeated decline trend of face-event distance shows that the abutment stress triggered more seismic events near the longwall opening as the face continued retreating, until a new cluster started with the presence of predominantly far away seismic events. This has the similar result with Kozłowska (2013), as a result of the destressing blasting and static stress transfer caused by the seismic activity (Orlecka-Sikora et al., 2012). This trend repeated seven times over the monitoring period, representing the cyclical change of mining-induced stress, which forms the basis of clustering. This is also synchronised with the periodical weighting of the site, showing seismic events are highly correlated with the mining activities. Therefore, seven clusters of seismic events are determined based on the cyclical tendency of the median face-event distance in Fig. 14. The identified clusters are listed in Table 4, and the spatial distribution of seismic events in seven clusters can be seen in Fig. 15. As recorded in Table 4, the face advance distance and the number of seismic events that occurred within each cluster generally follow a positive linear relationship (except Cluster #1 due to the slow start-up of the monitoring scheme), which indicates a direct correlation between longwall coal extraction and the amount of induced seismicity.

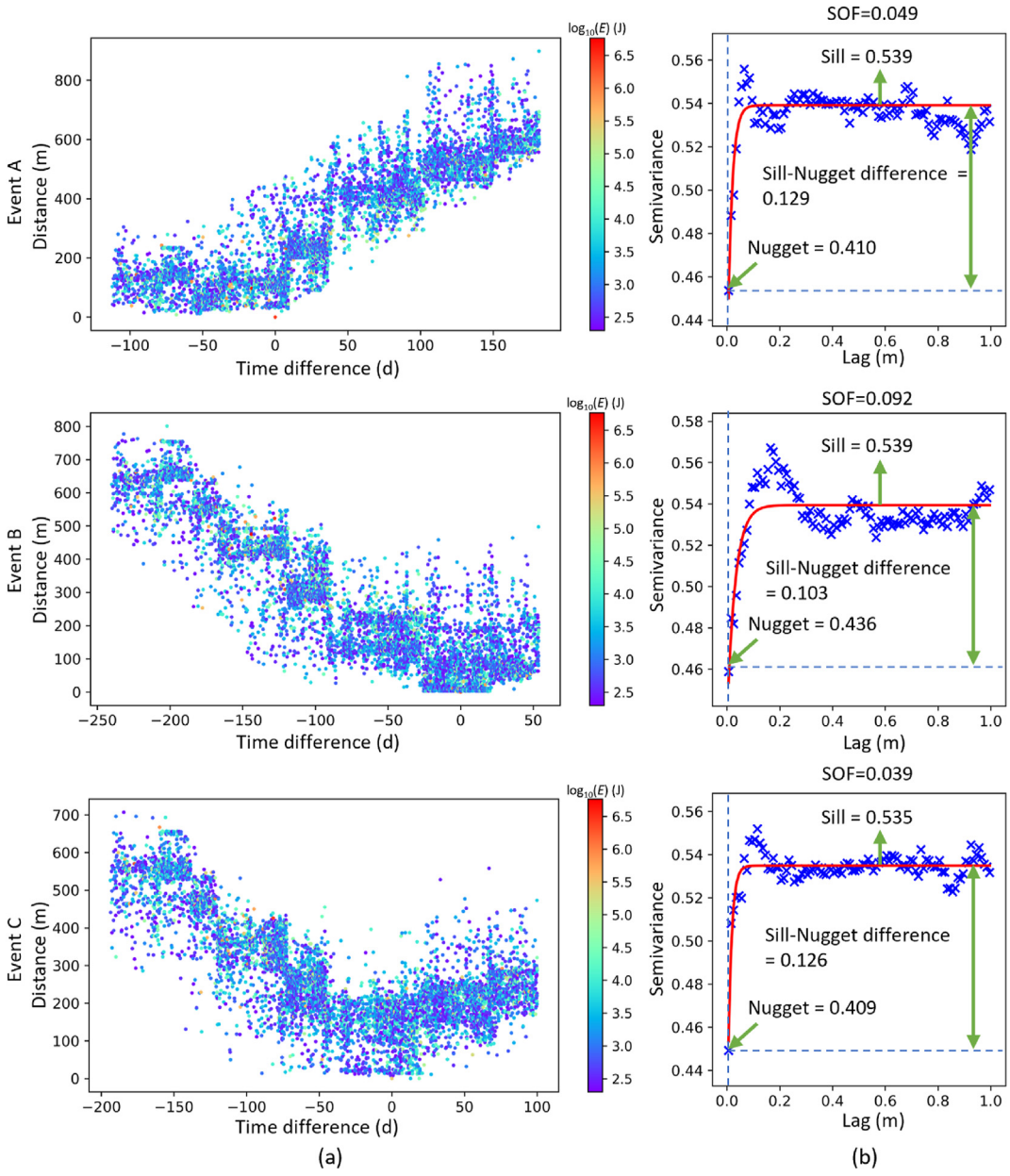

Fig. 13. (a) 2D distribution of time difference and distance between reference events and all other seismic events, and (b) Semivariograms for the three reference events.

7.2. Temporal correlation assessment of seismic events within individual clusters

Based on the clustering result in Section 6.2, temporal correlations within individual clusters can be explored using both ACF and the semivariogram function, similar to the procedure introduced in Section 5.1. Fig. 16 shows the ACF of the identified seven seismic clusters. According to this figure, for most of the seismic clusters, the SOF-time of cumulative daily energy is 2–4 d. Furthermore, the cumulative daily energy shows a lower SOF-time when investigating the ACF for the clustered data. The possible reason is that in the plot of ACF such as Fig. 1a, the ACF value for total seismic data at a specific lag k is approximately the average of the ACF values for individual clusters. For the ACF of each cluster, the ACF value with smaller k tends to be lower. The SOF-time only depends on the first point when the ACF value is lower than Bartlett's limit. Therefore, for a cluster, if the ACF value at a small k is occasionally lower than Bartlett's limit, its SOF-time would be low.

Fig. 17a shows the semivariogram result of the identified seismic clusters. Exponential fitting is used as determined in Section 5.1.2. According to Fig. 17a, seismic clusters show different fitting curves and parameters, indicating various correlations between clusters. The SOF-time, nugget and sill of the seismic event energy for each cluster are summarised in Fig. 17b. Compared to the SOF-time using cumulative daily energy, a higher SOF-time is presented when using the non-evenly spaced seismic energy data. It suggests that converting the non-evenly spaced seismic energy data to the cumulative daily energy data may weaken its temporal correlation.

As discussed in Section 2.2, in order to compare the correlation between clusters, the SOF-time is not the only assessment measure; the nugget and sill can also reflect vital information. A relatively strong correlation cluster should have a large difference between sill and nugget with an appropriate SOF-time. In addition, the small nugget indicates a strong correlation of the events within a short period, and the small sill indicates a strong correlation of the events with a large time difference.

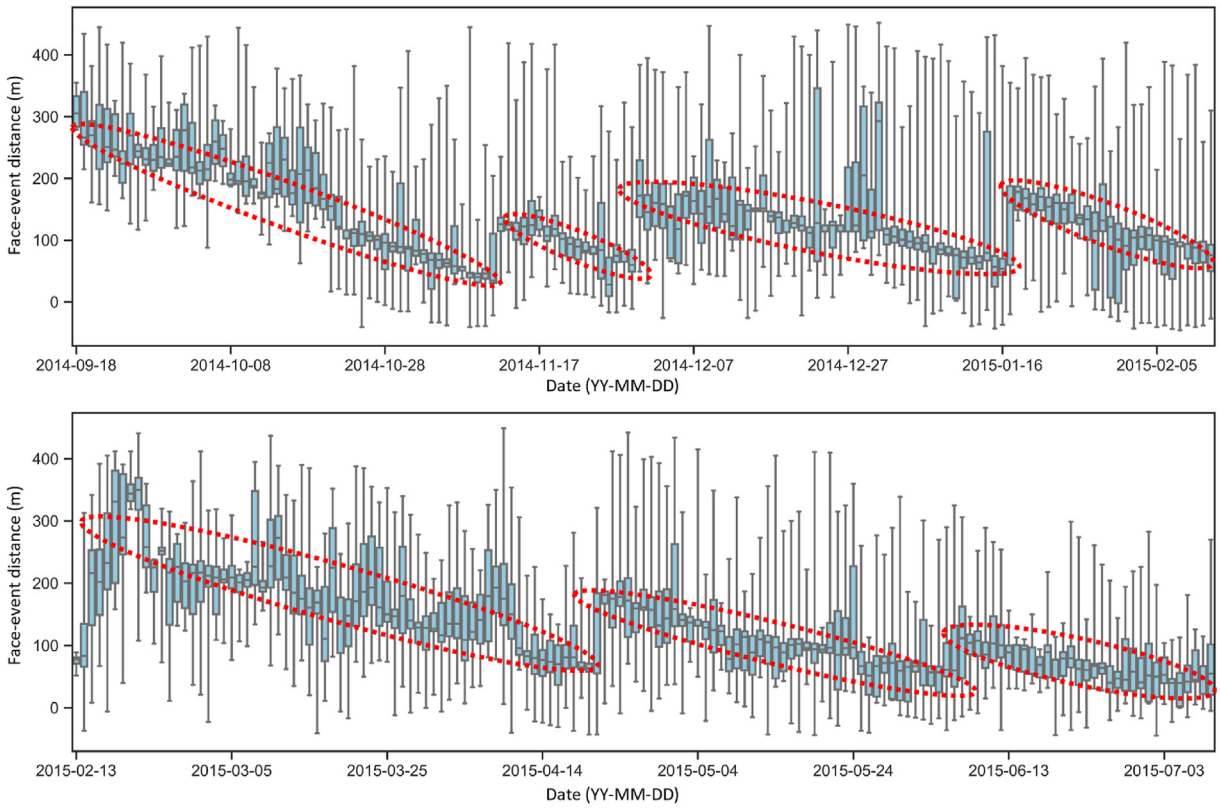

Fig. 14. Boxplot of face-event distance distribution in each day versus the date. Red circle indicates the range of clusters.

Table 4
The details of seven clusters of seismic events.

Cluster No.	Start date (YY-MM-DD)	End time	Face advance distance (m)	Number of seismic events with $\log_{10}E > 2.3$
1	2014-9-18	2014-11-12	220.9	952
2	2014-11-13	2014-11-29	65.15	402
3	2014-11-30	2015-1-17	145.35	1395
4	2015-1-18	2015-2-17	114.87	1436
5	2015-2-18	2015-4-21	177.18	1568
6	2015-4-22	2015-6-6	133.05	1716
7	2015-6-7	2015-7-8	88.4	751

7.3. Spatial correlation assessment of seismic events within clusters

Apart from the temporal correlation, the spatial correlation can also be investigated based on the identified clusters. The spatial correlation of different seismic clusters can be represented by the evolution of MI over the monitoring period. Fig. 18a shows the MI result of seismic data in LW102 during the monitoring period, separated by different clusters. For most clusters, such as Clusters 1, 3, 5, 6 and 7, there will be one or more peaks of MI located in the middle of each cluster period, and MI values at the start and end of the cluster are lower than the peak value. It illustrates the concentration and transfer of the high-density seismic activity area from one cluster to another as a response to progressive coal extraction in the longwall face. The reason could be that the seismic events tend to assemble in the centre of clusters. Still, with the advance of the longwall face, the seismic events transfer from one centre to another, which positively increases the randomness of the event location and decreases the MI value. Furthermore, the peak value, the range of MI, and the evolutionary process show different patterns among the identified seven seismic clusters, which is mainly because of the varying spatial correlation of seismic events during the panel retreating.

To investigate the MI characteristics within one cluster, three typical MI at the start, the highest MI and the end of Cluster 5 are used for analysis. Usually, the transformation of seismic events between clusters leads to a relatively low MI value. However, at the start of Cluster 5, MI may present a higher value if the events are concentrated in more than one centre. Also, in Fig. 18b and d, the semivariogram at the start shows a higher nugget and a lower SOF-space, indicating a low spatial correlation and a low radius of the correlative area. In contrast, in Fig. 18c, the peak MI point presented a lower nugget and higher SOF-space due to a large and concentrated seismic events area. Due to limited seismic data available within each week, discrete semivariance points and poor semivariogram fitting were encountered. This may introduce an error in the correlative radius estimation. Therefore, a certain amount of data that can be used in semivariogram analysis should be required.

The cluster-based correlation analysis did not include the ST correlation since the ST correlation in this paper is based on reference points instead of the total time period. The result of ST method is mainly affected by the reference point selected, which will provide a similar result whether applying ST correlation in cluster-based analysis.

8. Discussion and conclusions

In this study, quantitative approaches were applied for temporal, spatial and ST correlation analysis of a set of seismic data in the longwall mining process.

ACF was used to evaluate the correlation of evenly spaced seismic data and in combination with semivariogram, whereby the temporal correlation of unevenly spaced seismic energy was also assessed. The SOF-time is applied to representing the period that a notable correlation of seismic data shows within. The SOF-time is calculated as 6 d for cumulative daily energy and 12 d for unevenly spaced seismic energy data, representing a potential reference period that seismic events within this period can contribute to further

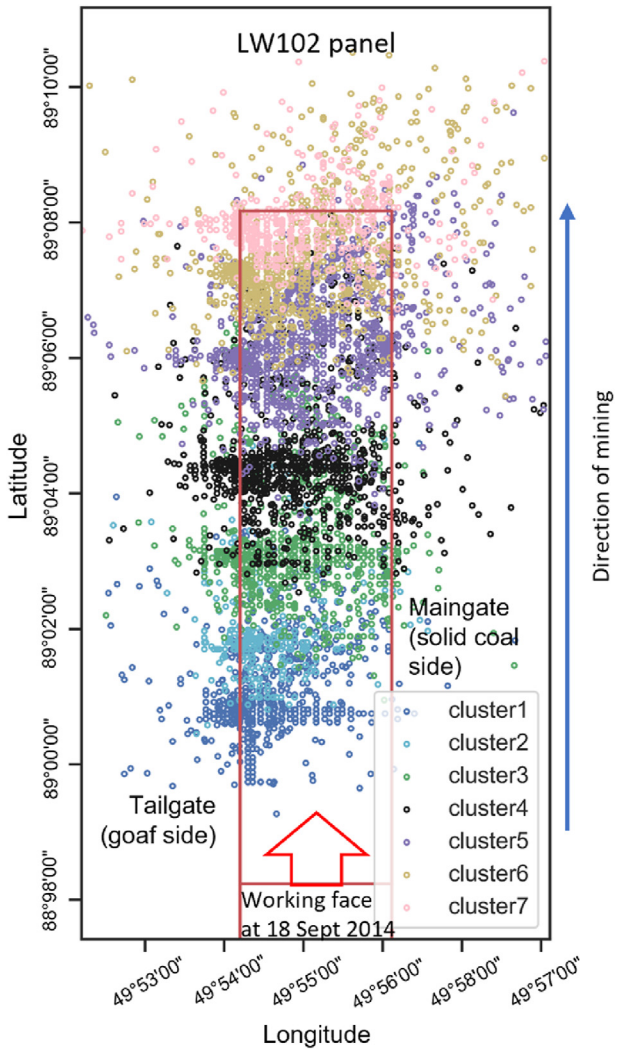

Fig. 15. Spatial distribution of seismic events in different clusters in LW102.

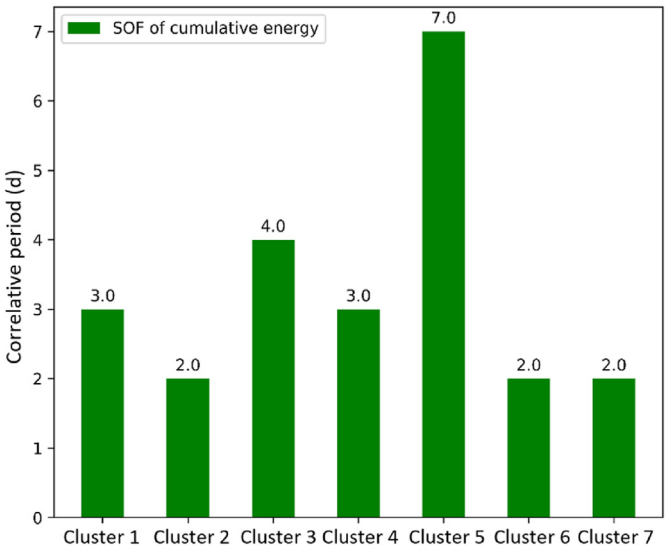

Fig. 16. Bar chart of SOF-time result of each cluster calculated by ACF using the cumulative daily energy.

evaluation and prediction. A semivariances assessment detects the maximum correlative period as 40 d, and the temporal correlation within this period can represent a universal correlation of a period larger than 40 d. With the maximum correlative period as a moving time window, and on account of the long-term mining operation and the variability of the temporal correlation at different mining stages, the evolution of the temporal correlation is determined.

The spatial correlation of the seismic data was estimated using MI. Based on the *Z*-value, most of the monitoring periods present a strong spatial correlation. To determine the radius of the correlative area (SOF-space), the spatial semivariogram assessment was applied. The seismic data show a strong spatial correlation within 23 m area, which can be explained as the seismic response to mining abutment stress or a set of localised discontinuities. The correlative period and distance scale can be used as the critical input parameters for seismic/rockburst hazard prediction, seismic attributes inversion, and mining-induced fracture characterisation.

The ST correlation has been assessed by investigating the distance and time differences with respect to three reference points. The quantitative assessment shows similarity on all three points and can be explained by the fracture behaviour during coal extraction. The proposed method introduced in Sections 5 and 6 improved the understanding of correlation for various purposes and multiple data types. It provides a rational approach to quantitatively assess the seismic data correlations in longwall mining.

To assess the ST correlation between different clusters, clear clustering characteristics have been observed by investigating distance distribution to working face versus time distributions. Within each cluster, the evaluation of correlation can have variable patterns. More parameters such as nugget and sills need to be applied when assessing the correlation between or within clusters. The investigation of correlations within clusters provides an understanding of the correlation within a specific period of mine activities or an area of rock mass discontinuities. The SOF-time and SOF-space of each cluster offer references to select a more accurate time window and grid size for other seismic data-driven prediction tasks. The ACF value, MI value, the nugget and sill in the semivariogram all contribute to evaluating the reliability of the SOF.

The correlation analysis methods (ACF, semivariogram and MI) introduced in this research are universal and can be easily transferred to other mine sites. However, the correlation results in time and space can be site-specific. This is because the variety of geological settings and mining conditions can easily change seismic responses during mining, as well as the associated correlation results. Nevertheless, given the proposed correlation analysis methods are easy to apply to various mines with processed seismic data as input, site-specific correlation results can be quickly obtained to select a proper time-space window for seismic hazard assessment. In addition, it is also possible to explore the correlation within seismic source mechanisms, e.g. fracture pattern or mining rate. This may yield more universal correlation results that are not banded to a specific site, and the seismic spatial/temporal correlation can be estimated once a fracture pattern/mining rate is known. However, this is out of the scope of the current paper and will be investigated in the future.

Nevertheless, the correlation analysis methods proposed in this research are a bold attempt to scientifically guide the selection of time or space windows for seismic hazard forecasts. This answers the following fundamental questions regarding using past seismic data to predict the future: (1) future seismic risk assessment or hazard prediction is possible only if past seismic data are correlated, and the stronger correlation can result in the better prediction results; (2) the ST correlation in seismic data can be quantitively measured, and then fed to hazard prediction algorithms to select a proper window of past seismic data as input; and (3) the correlation window is not a constant number and may vary over the monitoring period, which increases the difficult of seismic hazard prediction. This research also made several assumptions that need to be discussed: (1) the seismic location and

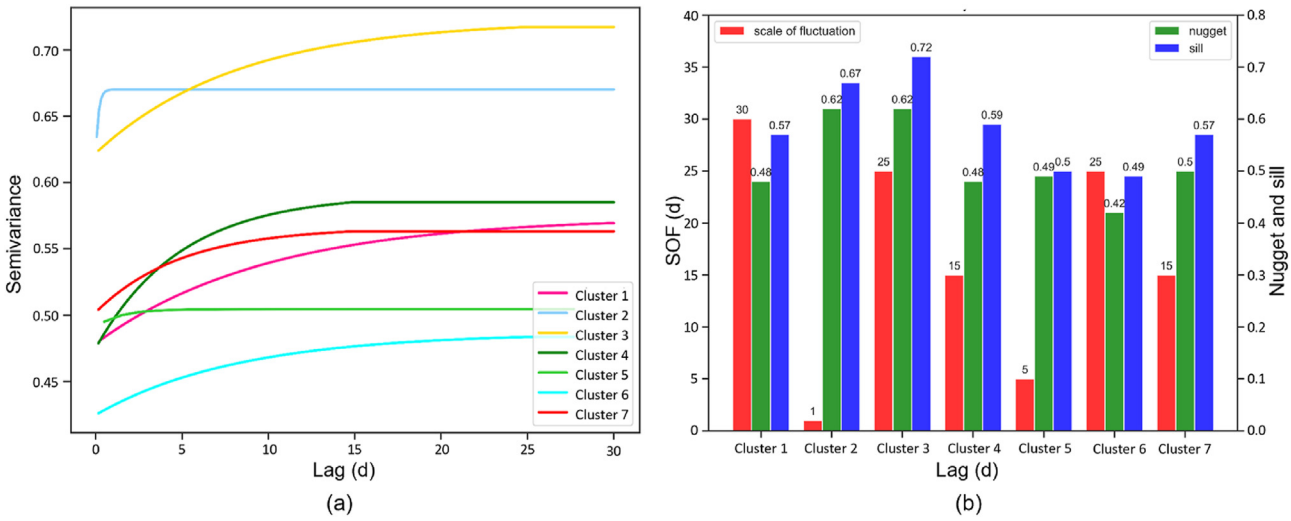

Fig. 17. (a) Semivariogram plot for each cluster with maximum lag of 30 d, and (b) Bar chart of SOF-time (red), nugget (green), and sill (blue) for each cluster.

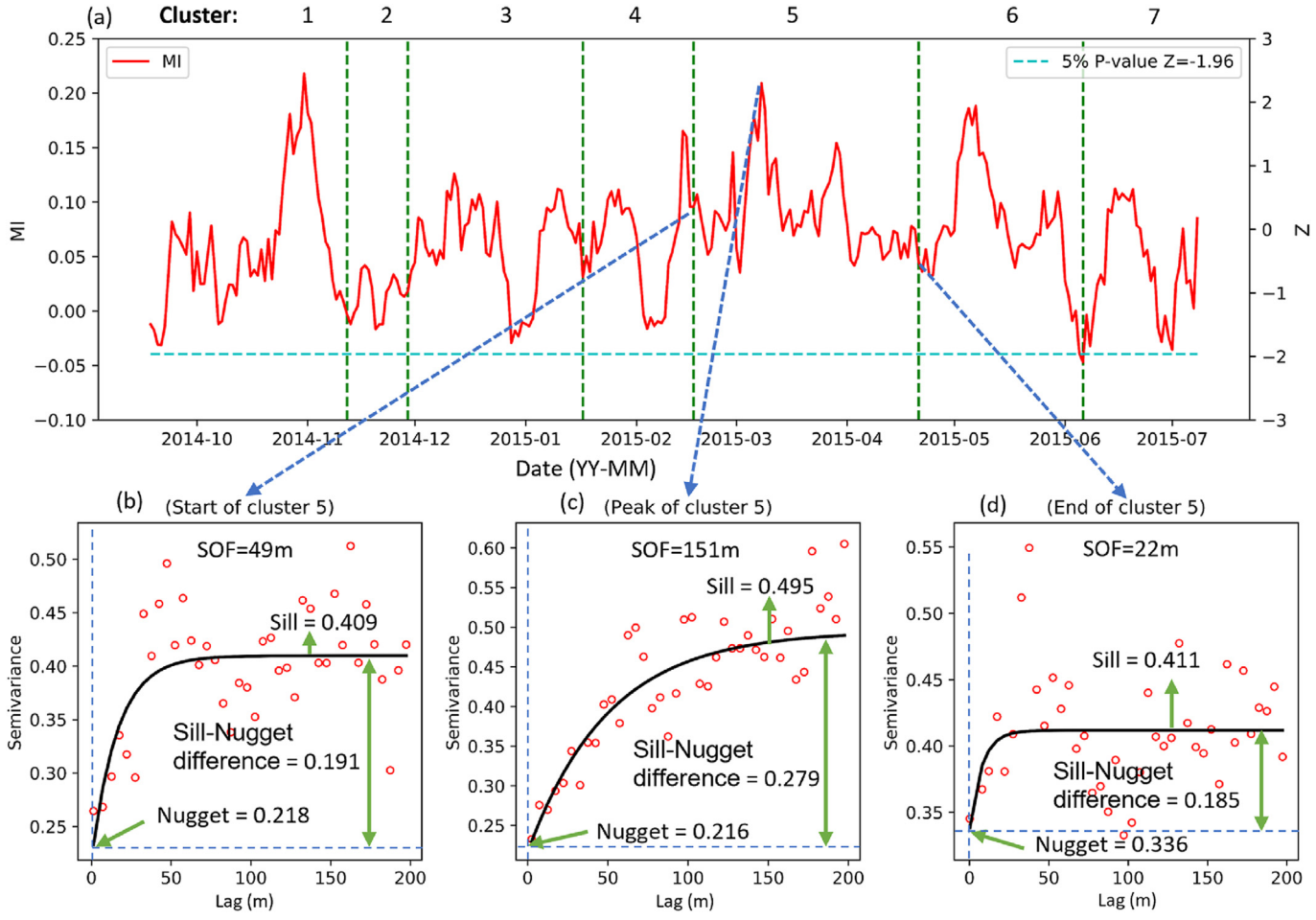

Fig. 18. (a) MI separated by clusters of seismic events and semivariogram for the (b) start, (c) peak and (d) end of Cluster 5.

energy calculation are relatively accurate; (2) as the seismic data with energy magnitude lower than 2.3 were not completed recorded, they were not used in the correlation study; and (3) the seismic events are only induced by either the mine extraction of LW102 or the failure of the geological structures. These assumptions may limit the application of the proposed correlation analysis methods. For example, the first assumption requires relative high accuracy of seismic event location with a dense monitoring array of geophones. Only seismic events that are well-enclosed by the geophone array can be used for correlation analysis. In the second assumption, the selection of the completed energy magnitude (>2.3 in this research) is difficult since it will filter a large amount of low energy seismic data, which may have an inherent

correlation with the rest of the seismic events. This filtering operation ensures the quality of input data, but its impact on correlation analysis needs to be further understood. The last assumption rules out seismic events that may be incurred by mining equipment movement or blasting operations at the site. These events do not indicate any natural correlation of mining-induced seismicity, thus better to be characterised and removed in our future work.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

List of symbols

a	Range of influence
c_0	Autocovariance in lag 0
C, C'	Fitting coefficients for Gaussian function
c_k	Autocovariance
E	Energy of seismic events
E_c	Cumulative daily energy
$G(x)$	Fitting functions
MI	Moran's I value
I_B	Bartlett's limit
k	Lag
N	Total number of seismic events
V_s	Semivariance
μ	Mean of the total studied data points
w_{ij}	An element of a matrix of spatial weights
x_i/x_j	The ith/jth seismic datum
Z	Standardised Moran's I
θ	Scale of fluctuation

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jrmge.2022.04.002.

References

Arabasz, W.J., Nava, S.J., McCarter, M.K., et al., 2005. Coal-mining seismicity and ground-shaking hazard: a case study in the Trail mountain area, Emery County, Utah. Bull. Seismol. Soc. Am. 95, 18–30.

Bischoff, M., Cete, A., Fritschen, R., Meier, T., 2010. Coal mining induced seismicity in the Ruhr area, Germany. Pure Appl. Geophys. 167, 63–75.

Bosch, M., Mukerji, T., Gonzalez, E.F., 2010. Seismic inversion for reservoir properties combining statistical rock physics and geostatistics: a review. Geophysics 75, 165–176.

Cai, W., Bai, X., Si, G., Cao, W., Gong, S., Dou, L., 2020. A monitoring investigation into rock burst mechanism based on the coupled theory of static and dynamic stresses. Rock Mech. Rock Eng. 53, 5451–5471.

Cai, W., Dou, L., Si, G., et al., 2019. A new seismic-based strain energy methodology for coal burst forecasting in underground coal mines. Int. J. Rock Mech. Min. Sci. 123, 104086.

Cai, W., Dou, L., Si, G., Hu, Y., 2021. Fault-induced coal burst mechanism under mining-induced static and dynamic stresses. Engineering 7, 687–700.

Cai, X., Cheng, C., Zhou, Z., Konietzky, H., Song, Z., Wang, S., 2021. Rock mass watering for rock-burst prevention: some thoughts on the mechanisms deduced from laboratory results. Bull. Eng. Geol. Environ. 80, 8725–8743.

Cao, W., Shi, J.Q., Durucan, S., Si, G., Korre, A., 2020. Gas-driven rapid fracture propagation under unloading conditions in coal and gas outbursts. Int. J. Rock Mech. Min. Sci. 130, 104325.

Cao, W., Shi, J.Q., Si, G., Durucan, S., Korre, A., 2018. Numerical modelling of microseismicity associated with longwall coal mining. Int. J. Coal Geol. 193, 30–45.

Clark, I., 1979. Practical Geostatistics. Elsevier.

Duan, Y., Luo, X., Si, G., Canbulat, I., 2022. Seismic source location using the shortest path method based on boundary discretisation scheme for microseismic monitoring in underground mines. Int. J. Rock Mech. Min. Sci. 149, 104982.

Gibowicz, S.J., 2009. Seismicity induced by mining: recent research. Adv. Geophys. 51, 1–53.

Gibowicz, S.J., 1986. Physics of fracturing and seismic energy release: a review. Pure Appl. Geophys. 124, 611–658.

Gibowicz, S.J., Kijko, A., 1994. An Introduction to Mining Seismology. Elsevier.

Gibowicz, S.J., Lasocki, S., 2001. Seismicity induced by mining: Ten years later. Adv. Geophys. 44, 39–181.

González, J.A., Rodríguez-Cortés, F.J., Cronie, O., Mateu, J., 2016. Spatio-temporal point process statistics: a review. Spat. Stat. 18, 505–544.

Guha, S.K., 2000. Induced Earthquakes. Kluver Academic Publishers.

Hohn, M.E., 1988. Geostatistics and Petroleum Geology. Springer.

Jaksa, M.B., Kaggwa, W.S., Brooker, P.I., 1999. Experimental evaluation of the scale of fluctuation of a stiff clay. In: Proceedings of the 8th International Conference on Applications of Statistics and Probability in Civil Engineering, pp. 415–422.

Kijko, A., Funk, C.W., 1996. Space-time interaction amongst clusters of mining induced seismicity. Pure Appl. Geophys. 147, 277–288.

Kisilevich, S., Mansmann, F., Nanni, M., Rinzivillo, S., 2010. Spatio-temporal clustering: a survey. In: Data Mining and Knowledge Discovery Handbook. Springer, pp. 1–22.

Kozłowska, M., 2013. Analysis of spatial distribution of mining tremors occurring in Rudna copper mine (Poland). Acta Geophys. 61, 1156–1169.

Lasocki, S., 2008. Some unique statistical properties of the seismic process in mines. In: Proceedings of the 1st Southern Hemisphere International Rock Mechanics Symposium. Australian Centre for Geomechanics, Perth, pp. 667–678.

Leśniak, A., Isakow, Z., 2009. Space-time clustering of seismic events and hazard assessment in the Zabrze-Bielszowice coal mine, Poland. Int. J. Rock Mech. Min. Sci. 46, 918–928.

Mendecki, A.J., Van Aswegen, G., Mountfort, P., 1999. A guide to routine seismic monitoring in mines. In: A Handbook on Rock Engineering Practice for Tabular Hard Rock Mines. Creda Communications, Cape Town.

Olszewska, D., Lasocki, S., Leptokaropoulos, K., 2017. Non-stationarity and internal correlations of the occurrence process of mining-induced seismic events. Acta Geophys. 65, 507–515.

Onyejekwe, S., Kang, X., Ge, L., 2016. Evaluation of the scale of fluctuation of geotechnical parameters by autocorrelation function and semivariogram function. Eng. Geol. 214, 43–49.

Orlecka-Sikora, B., Lasocki, S., Lizurek, G., Rudziński, Ł., 2012. Response of seismic activity in mines to the stress changes due to mining induced strong seismic events. Int. J. Rock Mech. Min. Sci. 53, 151–158.

Schultz, R., Skoumal, R.J., Brudzinski, M.R., Eaton, D., Baptie, B., Ellsworth, W., 2020. Hydraulic fracturing-induced seismicity. Rev. Geophys. 58, 1–43.

Si, G., Cai, W., Wang, S., Li, X., 2020. Prediction of relatively high-energy seismic events using spatial–temporal parametrisation of mining-induced seismicity. Rock Mech. Rock Eng. 53, 5111–5132.

Si, G., Durucan, S., Jamnikar, S., et al., 2015. Seismic monitoring and analysis of excessive gas emissions in heterogeneous coal seams. Int. J. Coal Geol. 149, 41–54.

Song, Z., Konietzky, H., Herbst, M., 2019. Bonded-particle model-based simulation of artificial rock subjected to cyclic loading. Acta Geotechnol. 14, 955–971.

Stec, K., 2007. Characteristics of seismic activity of the upper Silesian coal basin in Poland. Geophys. J. Int. 168, 757–768.

Tiefelsdorf, M., Boots, B., 1995. The exact distribution of Moran's I. Environ. Plann. 27, 985–999.

Wang, C., Si, G., Zhang, C., Cao, A., Canbulat, I., 2021a. Location error based seismic cluster analysis and its application to burst damage assessment in underground coal mines. Int. J. Rock Mech. Min. Sci. 143, 104784.

Wang, C., Si, G., Zhang, C., Cao, A., Canbulat, I., 2021b. A statistical method to assess the data integrity and reliability of seismic monitoring systems in underground mines. Rock Mech. Rock Eng. 54, 5885–5901.

Węglarczyk, S., Lasocki, S., 2009. Studies of short and long memory in mining-induced seismic processes. Acta Geophys. 57, 696–715.

Woodward, K., Wesseloo, J., Potvin, Y., 2018. A spatially focused clustering methodology for mining seismicity. Eng. Geol. 232, 104–113.

Zhang, J.Z., Zhou, X.P., 2020a. Forecasting catastrophic rupture in brittle rocks using precursory AE time series. J. Geophys. Res. Solid Earth 125, e2019JB019276.

Zhang, J.Z., Zhou, X.P., 2020b. AE event rate characteristics of flawed granite: from damage stress to ultimate failure. Geophys. J. Int. 222, 795–814.

Zhao, Y., Yang, T., Bohnhoff, M., et al., 2018. Study of the rock mass failure process and mechanisms during the transformation from open-pit to underground mining based on microseismic monitoring. Rock Mech. Rock Eng. 51, 1473–1493.

Zhou, X.P., Zhang, J.Z., 2021. Damage progression and acoustic emission in brittle failure of granite and sandstone. Int. J. Rock Mech. Min. Sci. 143, 104789.

Shuyu Wang obtained his BSc degree in Geophysics from University of Science and Technology of China in 2015 and his MSc degree in Petroleum Engineering from University of New South Wales in 2018. Since 2018, he is a PhD candidate in Mining Engineering from University of New South Wales with full scholarship. His research interests include: (1) Statistical analysis of micro-seismic data including clustering and correlation assessment; (2) Processing of micro-seismic data along longwall mining and its aspect to induced fracture activities and rock mechanics during mine extraction to ensure mine safety; (3) Theoretical and numerical modeling of mining induced fracture activities in longwall mining; and (4) Generation of mining induced DFN model with its numerical simulation and matches with micro-seismic data.