TIME SERIES PREDICTION OF ROCK BURST BASED ON DEEP LEARNING
A Case Study

by

***Hui LI*[a], *Si-Yuan GONG*[a*], *Xiu-Feng ZHANG*[b], *Lin-Ming DOU*[a], *Guo-Ying LI*[b], and *Hao-Ze LI*[a]**

[a] School of Mines, China University of Mining and Technology, Xuzhou, China
[b] Shandong Energy Group Co., Ltd (Coal Industry Management Department), Jinan, China

Original scientific paper
https://doi.org/10.2298/TSCI2502319L

Rockburst is a common mining hazard causing dynamic damage to coal and rock masses, posing significant threats to personnel and equipment safety. Various analytical methods exist to assess impact risks, with microseismic monitoring systems playing a pivotal role due to their stability, dynamism, and continuity. This approach utilizes a dual residual connection and a deeply connected stack architecture to facilitate seasonal-trend predictions and enhance their interpretability in time series prediction tasks using a purely deep learning model. The time-frequency and total energy of microseismic events are predicted using the proposed approach, and a comparative experimental study is conducted on the time window lengths of M = 7 days and M = 4 days. The results indicate that the proposed approach effectively predicts the evolution trend of microseismic event frequency, with minor discrepancies between the predicted results and the actual monitoring values, showing its excellent prediction performance and generalization capability.

Key words: *rockburst, monitoring and warning, time series, deep learning*

Introduction

Rockburst typically refers to the mining-induced dynamic phenomena of damage in coal and rock masses surrounding coal mine roadways or working faces due to the instantaneous release of elastic energy under high stress conditions. This phenomenon is often accompanied by loud noises, the ejection of coal and rock masses into the working space, and air blasts [1]. The microseismic monitoring technique, offering advantages such as stability, 3-D, and continuous temporal and spatial data, can provide information on the location and energy of vibrations, among other metrics. Hence, it has been widely applied in monitoring and providing early warnings of dynamic hazards such as rockburst in coal mines. However, there is limited research on time-series quantitative early warning of rockbursts. These studies often utilize acquired microseismic monitoring information solely for retrospective analyses of impact hazards [2, 3], resulting in poor timeliness and accuracy of impact hazard warnings. Nonetheless, the significance of these two aspects cannot be overlooked. Accurate and quantitative prediction of the time of high energy microseismic occurrences can significantly enhance the effectiveness of warnings and gain valuable time for hazard prevention and response.

[*] Corresponding author, e-mail: gsy_cumt@cumt.edu.cn

In the study of rockburst warnings across the temporal dimension, traditional time series analysis methods, including microseismic time-frequency analysis, time-energy analysis, autoregressive models, moving average models, and autoregressive moving average models [4-6], have yielded significant results in certain areas. For example, Lu *et al.* [2] developed a characteristic function that uses the variance of microseismic energy as a discriminative indicator and predicted trends in microseismic energy release using the auto-regression integrated moving average models and threshold autoregressive models. This approach defined the time series variation trends of microseisms. Pei *et al.* [3] designed a 1-D convolutional neural network that inputs historical microseismic energy levels from a specified interval to predict microseismic energy levels for subsequent time phases. Qin *et al.* [7] decomposed and reconstructed energy-time and frequency-time curves of the original microseismic data using modal decomposition techniques to extract feature information and trained a microseismic time series prediction model employing a deep neural network that integrates CNN-LSTM-GRU.

However, the internal operations of deep learning models often resemble a black box, leading to limited interpretability of the results. This study proposes a purely deep learning-based approach to focus on univariate regression problems in time series, such as predicting microseismic frequency and energy. It performs comparative analyses to determine how unit time series datasets of varying lengths adapt to the model and to optimize the optimal time window span. The model can filter out data noise and extract trend and periodic terms from the most significant time series data.

Materials and methods

Materials

The dataset chosen for this study's deep learning time series prediction model comprises mining-induced microseismic activities recorded from January 1, 2024, to July 31, 2024, primarily collected from the 8302 working faces of the Xinjulong Coal Mine. The advanced real-time acquisition and measurement integration microseismic monitoring system is utilized at the mine, where the deployed microseismic stations consist of uniaxial velocity sensors with a natural frequency of 1 Hz and a sampling rate set to 500 samples per second. Underground sensors are synchronized with surface timing, and six monitoring stations are strategically placed around the study area. As mining progresses in the working face, the positions of these stations are periodically adjusted toward the mining advancement, ensuring that the distance from the stations to the study area does not exceed 1 km. The lay-out of the scheme is illustrated in fig. 1. A uniform velocity model of 4000 m/s is employed for locating microseismic events. Comparisons between several known blasting locations and event location results indicated that the 3-D positioning errors in the *X*-, *Y*-, and *Z*-directions range from 20-40 m, 30-50 m, and

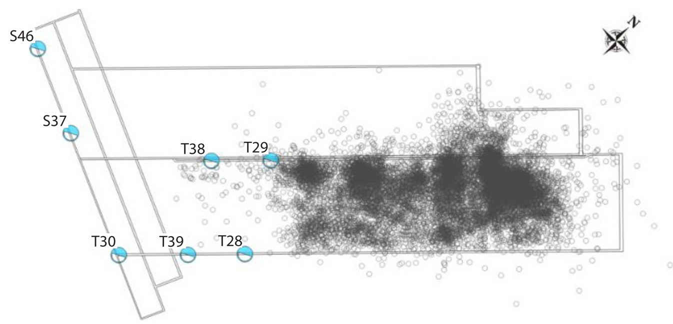

Figure 1. Lay-out of microseismic motoring points and location results at the 8302 working faces of the Xinjulong Coal Mine (January 1, 2024 to July 31, 2024)

40-80 m, respectively. Such accuracy meets the requirements for mine data analysis. A total of 12757 seismic events were recorded during the analysis period, with the majority occurring at an intermediate energy level of 10^3 J, while the counts for $<10^2$ J and $>10^4$ J were relatively minor. The energy distribution of the microseismic data generally follows a normal distribution, as shown in fig. 2.

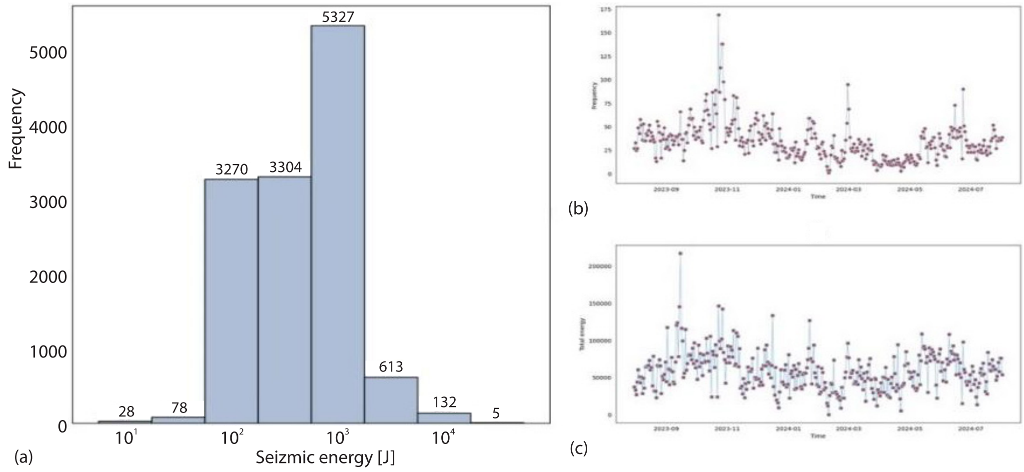

Figure 2. (a) Microseismic energy level distribution at the 8023 working faces of the Xinjulong coal mine, (b) time-frequency variation curve, and (c) time-total energy variation curve

Methods

With the extraction of the working face, monitoring equipment records the microseismic events caused by structural damage to the coal and rock masses near the mining area, forming a continuous time series database. Relevant studies indicated that microseismic events' frequency and energy variations over time exhibit certain regularities [8], supporting the correlation between historical and future microseismic events within specific time intervals. Similar patterns have also been observed in monitoring data from multiple coal mine sites, providing a theoretical and empirical foundation for predicting microseismic events.

Based on the aforementioned analyses, a functional relationship between historical and current microseismic data along the time axis can be obtained:

$$Ci = f\left(Ci-1, Ci-2, \cdots, Ci-\Delta t\right) \tag{1}$$

where C_i is the attributes of current microseismic data and Δt is the period used for historical data analysis.

Accordingly, a dataset that consists of a fixed window size, M, sliding forward by N lengths of L can be constructed [9], where M is the fixed window length, N is the number of samples in the dataset ($N = \Delta t - M$), and L is the sliding step length of the fixed window (all time units are in days). Generally, L is set to 1 to increase the number of samples in the dataset. The fixed window size serves as a hyperparameter for the dataset model. This study constructs two datasets with $M = 4$ and $M = 7$, comparing their performances in subsequent deep-learning models. Figure 3 illustrates the dataset construction using $M = 7$.

The N-BESTS method is proposed in 2020 [10]. This model is a deep neural network that employs dual residual connections and a deep, fully connected stack, achieving season-

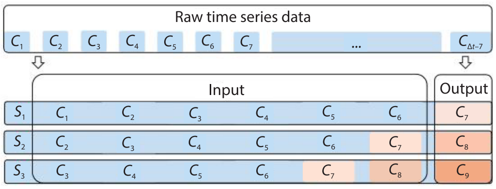

Figure 3. Example of constructing time series dataset

al-trend predictions and their interpretability in time series prediction tasks through a purely deep learning approach. Figure 4 presents the overall structure of the model.

The time series dataset is input into the box to output the amplification coefficients for retrospective history phases and future prediction phases through a 4-layer fully connected (FC) network:

$$h_{l,1} = FC_{l,1}\left(x_l\right),\ h_{l,2} = FC_{l,2}\left(h_{l,1}\right),\ h_{l,3} = FC_{l,3}\left(h_{l,2}\right),\ h_{l,4} = FC_{l,4}\left(h_{l,3}\right)$$
$$\theta_l^b = \text{Linear}_l^b\left(h_{l,4}\right),\ \theta_l^f = \text{Linear}_l^f\left(h_{l,4}\right) \tag{2}$$

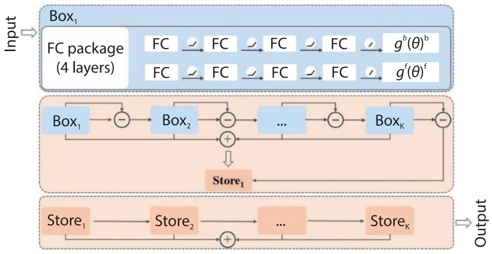

Figure 4. Architecture of N-BESTS deep neural network

Each layer inside the box functions as RELU + FC. For instance, in the first layer of the FC box, $j_{l,1} = \text{RELU}(W_{l,1} + b_{l,1})$, where W is the weight, and b is the bias. The subsequent FC layers perform simple linear mappings $\theta_l^f = W_l^f h_{l,4}$. The data are inserted into the module composed of the box, undergo processing by a structure similar to the residual network, and then move to the Store module to execute the fully connected neural network architecture, ultimately outputting the predicted values of the time series. This experiment sets the predicted time series length to 1 day.

The mean absolute percentage error (MAPE) [11] is chosen as the model evaluation metric, which can disregard the impact of the model scale. The selected periods for this experiment all fall within the coal mining production period, and the microseismic data does not exhibit zero cases in the time series. The calculation formula of MAPE is:

$$MAPE = \frac{100}{N}\sum_{i=1}^{N}\frac{\left|C_{i+\Delta t} - \hat{C}_{i+\Delta t}\right|}{\left|C_{i+\Delta t}\right|} \tag{3}$$

where N is the total number of samples, $C_{i+\Delta t}$ – the actual value, and $\hat{C}_{i+\Delta t}$ – the predicted value.

Results and discussion

The model data is divided into training and testing sets at a ratio of 8:2 [12]. Figure 5 shows the results of the time-frequency and time-total energy predictions for the Xinjulong coal mine 8302 working face dataset. It can be seen that the prediction performance for time-frequency is superior to that for time-total energy, and the hyperparameter fixed window length $M = 7$ outperforms M = 4. The coal mine's on-site situation indicates that a thick and hard roof can be above the 8302 working face, with a hanging roof present during the working face extraction process. As the working face advances, the rock strata above the mining area gradually fracture, experiencing a relatively long period of continuous fracturing. The deep learning model more effectively learns the periodic patterns of frequency, while the analysis of total energy is impacted by on-site pressure relief measures, such as blasting and large-diameter drilling, which disrupt periodic energy release patterns from roof fractures. In addition, the total energy exhibits significant fluctuations in absolute value, resulting in suboptimal prediction performance in this case.

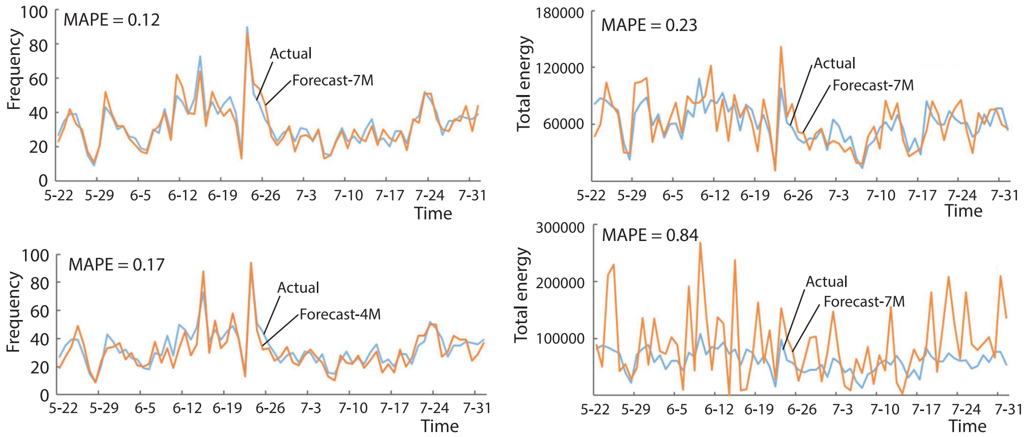

Figure 5. Prediction results of time-frequency and time-total energy by N-BESTS deep neural network under conditions of $M = 7$ and $M = 4$

Conclusions

This study constructs a microseismic time series dataset for coal mines based on a fixed window with sliding steps and applies the purely deep learning model N-BESTS for predictions. Its unique dual residual neural network structure effectively predicts the evolution trends of microseismic events, with slight discrepancies between prediction results and actual monitoring values. Under the setting of $M = 7$, the model's test set results are $MAPE = 0.12$ for time-frequency and $MAPE = 0.23$ for time-total energy, demonstrating good predictive performance.

The current sample size of the dataset is still relatively small for deep learning models, potentially leading to incomplete feature extraction. Future research could apply data augmentation techniques such as spatiotemporal data interpolation [13], generative adversarial networks [14], and diffusion models [15], to expand datasets and enhance microseismic monitoring applications in coal mines.

Acknowledgment

This research was supported by the following projects: the Taishan Industrial Experts Program (No. tscx202408130), the National Natural Science Foundation of China (No. 524742270 and 52304105), the Natural Science Foundation of Jiangsu Province (No. BK20231499), the Shandong Energy Group (No. SNKJ2022BJ03-R28).

Nomenclature

Ci – attributes of current microseismic data, [–]
$C_{i+\Delta t}$ – actual value, [–]
N – total number of samples, [–]
Δt – period used for historical data analysis, [s]

Acronyms

FC – fully connected layer.
MAPE – mean absolute percentage error

References

[1] Dou L. M., *et al.*, T*heory and Technology of Rock Burst Prevention*, China University of Mining and Technology Press, Xuzhou, China, 2001
[2] Lu, J. G., *et al.*, Microseismic Predicting Coal Bump by Time Series Method, *Journal of China Coal Society*, *35* (2010), 2, pp. 2002-2005

[3] Pei, Y. Y, *et al*., Time Series Prediction of Microseismic Energy Level Based on Feature Extraction of 1-D Convolutional Neural Network for Feature Extraction, *Journal of Engineering Science*, *43* (2021), 3, pp. 1003-1009
[4] Wang, J. Z., *et al*., The ARMA Model Identification Using Particle Swarm Optimization Algorithm*, International Conference on Computer Science and Information Technology*, *2008* (2008), 4, pp. 223-227
[5] Bogl, M., W., *et al.*, Visual Analytics for Model Selection in Time Series Analysis, *IEEE Transactions on Visualization and Computer Graphics*, *19* (2021), 2, pp. 2237-2246
[6] Buendia, R. J. M., *et al.*, A Disease Outbreak Detection System Using Autoregressive Moving Average in Time Series Analysis, *Proceedings*, 6th International Conference on Information, Intelligence, Systems and Applications (IISA), Corfu, Greece, 2015, Vol. 1, pp. 1-5
[7] Qin, C. K., *et al.*, Time Series Prediction Method of Microearthquakes in Coal Mine Based on Mode Decomposition and Deep Learning, *Journal of China Coal Society*, *49* (2019), 2, pp. 3781-3797
[8] Cao, A., *et al*., Physical Index and Data Fusion-Driven Method for Coal Burst Prediction in Time Sequence, *Journal of China Coal Society*, *48* (2023), 2, pp. 3659-3673
[9] Xu, Y., *et al*., Deep Transfer Learning for P-Wave Arrival Identification and Automatic Seismic Source Location in Underground Mines, *International Journal of Rock Mechanics and Mining Sciences*, *182* (2024), 3, ID105888
[10] Oreshkin, B. N., *et al.*, The N-BEATS: Neural Basis Expansion Analysis for Interpretable Time Series Forecasting, ICLR2020, https://doi.org/10.48550/arXiv.1905.10437, 2019
[11] Swanson, D. A., *et al.*, A Note on the Measurement of Accuracy for Subnational Demographic Estimates, *Demography*, *37* (2000), 4, pp. 193-201
[12] Vu, Tin, *et al*., Using Deep Learning for Big Spatial Data Partitioning, *ACM Transactions on Spatial Algorithms and Systems (TSAS)*, *7* (2020), 6, pp. 1-37
[13] Amato, F., *et al*., A Novel Framework for Spatio-temporal Prediction of Environmental Data Using Deep Learning, *Scientific Reports*, *10* (2020), 3, ID22243
[14] Krichen, M., Generative Adversarial Networks, *Proceedings*, 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), Delhi, India, 2023, Vol. 4, pp. 1-7
[15] Yang, L., *et al*., Diffusion Models: A Comprehensive Survey of Methods and Applications, *ACM Computing Surveys*, *56* (2022), 5, pp. 1-39

Paper submitted: October 3, 2024
Paper revised: November 15, 2024
Paper accepted: November 26, 2024