



# Microseismic Time-Series Prediction of Rockburst Events Based on an Adaptive Multiscale Noise-Resilient Mechanism

Qiyuan Xia<sup>1,2</sup>, Qiang Wu<sup>3</sup>, Baoquan Zhang<sup>4</sup>, Jun Gu<sup>3</sup>, Hai Wu<sup>1,2\*</sup>, Hailong Gan<sup>4</sup>, Chengjun Xu<sup>4</sup>

<sup>1</sup> School of Resource, Environment and Safety Engineering, Hunan University of Science and Technology, 411201 Xiangtan, China

<sup>2</sup> Sanya Institute, Hunan University of Science and Technology, 572000 Sanya, China

<sup>3</sup> Guizhou Ganxing Coal Industry Co., Ltd., 553300 Bijie, China

<sup>4</sup> Jiangxi Provincial Investment Group Co., Ltd., 330029 Nanchang, China

\* Correspondence: Hai Wu (wuhai@hnust.edu.cn)

Received: 12-26-2024

Revised: 02-18-2025

Accepted: 02-24-2025

**Citation:** Q. Y. Xia, Q. Wu, B. Q. Zhang, J. Gu, H. Wu, H. L. Gan, and C. J. Xu, "Microseismic time-series prediction of rockburst events based on an adaptive multiscale noise-resilient mechanism," *GeoStruct. Innov.*, vol. 3, no. 1, pp. 1–12, 2025. <https://doi.org/10.56578/gsi030101>.



© 2025 by the author(s). Licensee Acadlore Publishing Services Limited, Hong Kong. This article can be downloaded for free, and reused and quoted with a citation of the original published version, under the CC BY 4.0 license.

**Abstract:** Rockburst monitoring data acquired during underground coal mining are characterized by strong noise, nonlinearity, and multiscale coupling, which severely limit the predictive performance of existing models. To address these challenges, an innovative deep learning model was proposed. First, a hybrid denoising strategy combining Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) and wavelet transform was applied to enhance the quality of microseismic data. Subsequently, a trend–residual decomposition module was constructed to decouple the complex microseismic data into a trend component, representing long-term stress accumulation, and a residual component, capturing short-term fracture-induced fluctuations. On this basis, an improved adaptive multiscale noise-resilient Long Short-Term Memory (LSTM) unit was designed. A dynamic noise-control mechanism and a multiscale memory strategy were introduced to enable targeted feature extraction from the trend and residual branches, respectively. Furthermore, a multiscale interactive fusion (MSIF) module incorporating a channel attention mechanism was employed to dynamically integrate complementary information from both branches. The proposed framework was validated using field microseismic monitoring data from a northern coal mine. Experimental results demonstrated that the proposed model consistently outperformed five benchmark models across multiple evaluation metrics, achieving a recall of 88.37% with a false alarm rate of only 3.88%. These results confirm the effectiveness and robustness of the proposed approach for rockburst microseismic time-series prediction under noisy and complex conditions.

**Keywords:** Rockburst; Deep learning; Microseismic monitoring; Time-series prediction; Noise-resilient mechanism

## 1 Introduction

With the progressive increase in mining depth, underground coal extraction has gradually shifted from shallow to deep levels. As mining depth increases, in-situ stress acting on roadway surrounding rock intensifies, leading to increasingly complex geomechanical conditions and frequent occurrences of rockburst. Owing to the highly complex physical mechanisms underlying rockburst and the inherent difficulties associated with monitoring, prediction, and early warning, rockburst has become one of the primary factors constraining the safe production of deep underground coal mines.

At present, rockburst prediction approaches can be broadly classified into physics-driven methods and data-driven methods. In the physics-driven domain, multiple theoretical frameworks have been established. Jiang et al. [1] developed a dynamic assessment method by integrating microseismic, acoustic emission, and stress monitoring systems, thereby overcoming the limitations of conventional static evaluation approaches. Lei [2] constructed a rockburst hazard evaluation model through the comprehensive consideration of geological conditions and mining technological factors.

In recent years, with the rapid development of artificial intelligence techniques, data-driven approaches have attracted increasing attention. Wojtecki et al. [3] combined the rock mass bursting tendency index (WTG) with

machine learning algorithms and achieved satisfactory prediction performance under hard roof conditions. Guo and Yang [4] and Shi et al. [5] employed Random Forest (RF) and Generalized Regression Neural Network (GRNN) models, respectively, for rockburst risk prediction. Jiang et al. [6], taking the excavation of roadways in steeply inclined and ultra-thick coal seams at the Wudong Coal Mine in Xinjiang as a case study, analyzed rockburst precursory characteristics using microseismic data and multiphysical indicators, and subsequently constructed four machine learning prediction models, providing effective tools for rockburst prediction in steeply inclined coal seam roadways. For time-series data, LSTM networks have been widely adopted due to their strong capability in capturing long-range temporal dependencies [7–10]. Ding et al. [8] proposed a multi-source signal integrated early warning method based on LSTM-Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs). Cao et al. [11] investigated the fusion of explicit physical indicators and implicit deep learning features, while Liu et al. [12] developed a CNN–LSTM hybrid model for rockburst-related prediction tasks.

Despite the substantial progress achieved by the aforementioned studies from both physics-driven and data-driven perspectives, rockburst prediction remains a highly challenging task. The primary difficulty arises from the intrinsic characteristics of microseismic signals, which are typically dominated by strong nonlinearity and low signal-to-noise ratios. When confronted with such complex data, existing prediction models often struggle to effectively distinguish long-term evolutionary trends from short-term abrupt disturbances embedded in the effective signal. Moreover, conventional prediction models are generally highly sensitive to noise.

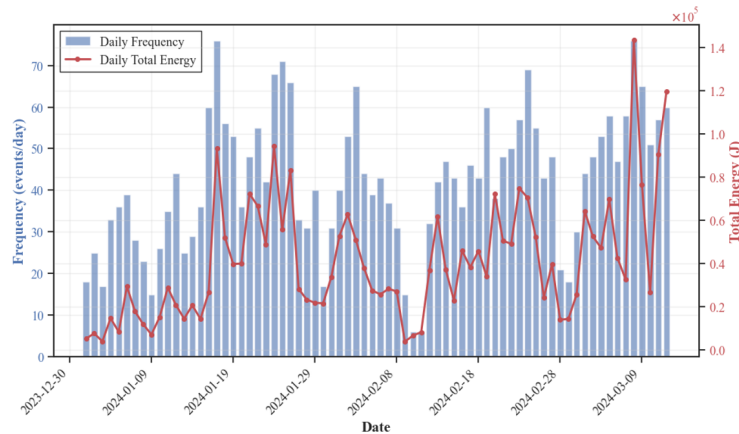
To address these challenges, an innovative deep learning framework, termed Trend–Residual Synergistic Network with Multiscale Collaborative Attention (TRSNet), was proposed. The core concept of TRSNet is inspired by decomposition–reconstruction paradigms in signal processing, in which the original information is decomposed to facilitate multidimensional and multiscale information mining, followed by reconstruction to achieve synergistic fusion across heterogeneous scales, enabling robust modeling.

## 2 Data Source and Data Preprocessing

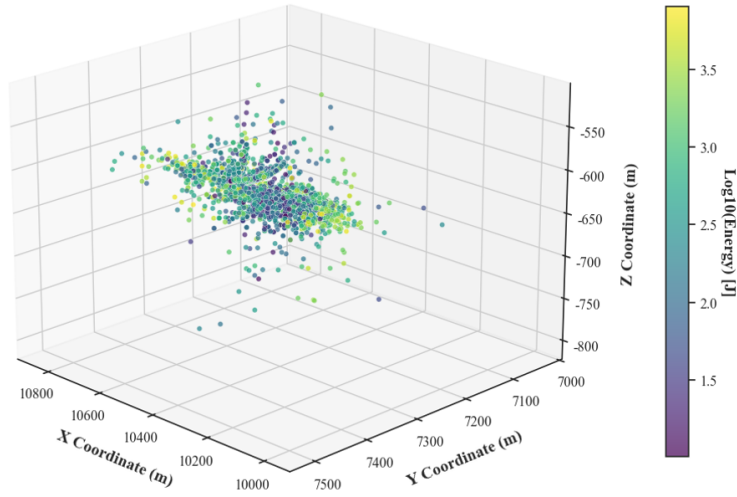
The microseismic monitoring data employed in this study were acquired from a longwall mining face in a northern underground coal mine and were primarily associated with microseismic activity generated during face advancement. The mining face is located within a deep mining district, with an average burial depth exceeding 600 m. A segment of continuous microseismic monitoring data collected during the extraction period was selected as the experimental dataset. The dataset spans three months and contains 3,035 recorded effective microseismic events.

Figure 1 illustrates the temporal evolution of daily microseismic event frequency and daily cumulative released energy during the monitoring period. The overall pattern exhibits a typical alternation between stress accumulation and stress release. The maximum daily released energy reached  $4.46 \times 10^4$  J. Although no strong rockburst events were observed during the study period, low-energy microseismic events occurred frequently. Such non-stationary time-series fluctuations further increase the difficulty of reliable rockburst hazard prediction.

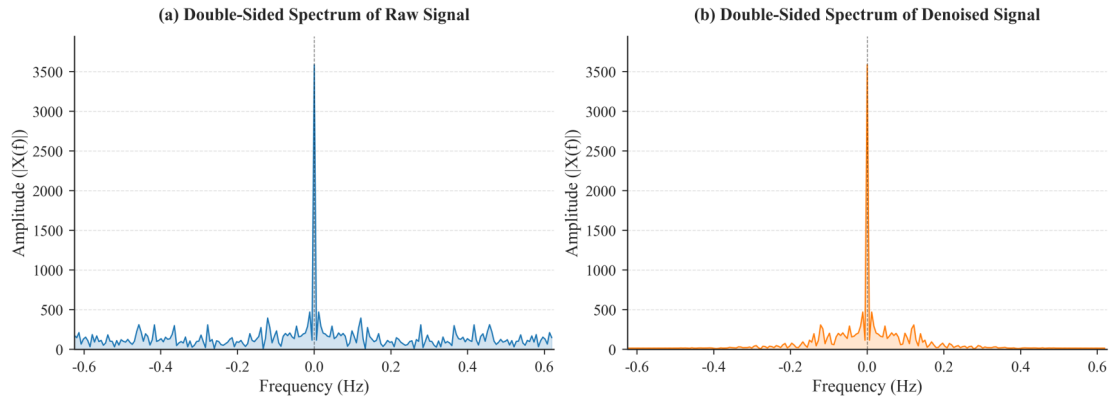
Figure 2 presents the three-dimensional spatial distribution of microseismic events. The events are shown to be non-uniformly distributed in space, displaying a pronounced clustered aggregation pattern along the roadway direction. In addition, the majority of high-energy microseismic events are concentrated within a depth range of approximately 650 m.



**Figure 1.** Temporal evolution of microseismic activity



**Figure 2.** Spatial distribution of microseismic events



**Figure 3.** Comparative spectral analysis of microseismic energy series before and after denoising

## 2.1 Data Noise Analysis and Stationarity Assessment

In the complex operational environment of underground coal mines, data acquisition is inevitably affected by various sources of noise. For prediction models, such noise can obscure latent patterns embedded in the data, thereby impairing model convergence and leading to degraded prediction accuracy. To enhance data quality, a hybrid denoising strategy combining CEEMDAN [13] and the wavelet transform [14] was employed to suppress noise components in the data.

To evaluate the effectiveness of the denoising procedure, Fast Fourier Transform (FFT)-based spectral analysis was conducted on the data before and after denoising, as illustrated in Figure 3. Prior to denoising, the effective dominant frequency components were found to be severely masked by noise. After denoising, noise components in the frequency spectrum were significantly attenuated. To avoid over-denoising-induced signal distortion, a small amount of residual noise was intentionally retained in the processed data. Consequently, no further aggressive denoising was applied; instead, subsequent improvements in prediction performance were pursued through enhanced noise robustness of the proposed prediction model.

## 2.2 Selection of Prediction Indicators

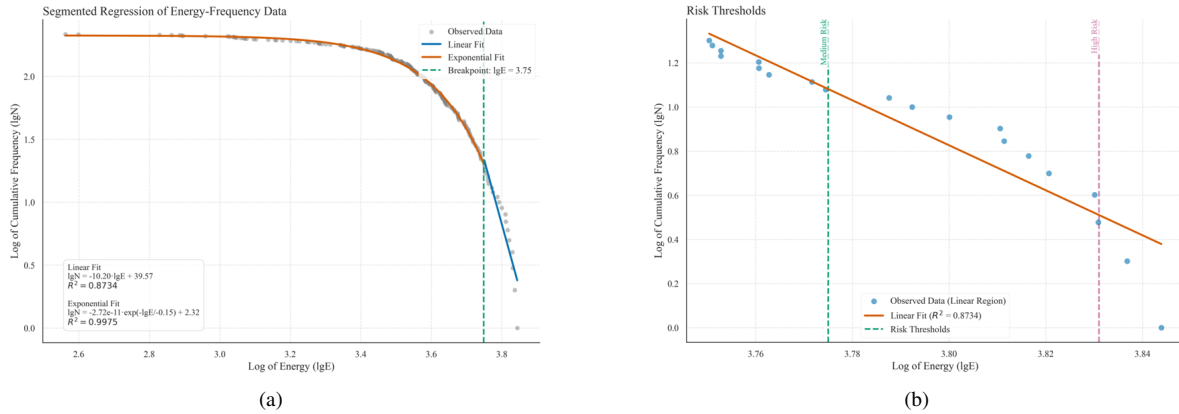
Feature selection constitutes a critical step in improving model learning efficiency and predictive performance, as feature quality directly determines whether effective information can be learned from the data. In this study, a set of statistical features was derived from microseismic monitoring data and employed as input variables for the rockburst prediction model. These features were computed within an 8-hour time window, enabling a comprehensive characterization of energy fluctuation and accumulation processes from both local and global perspectives. The selected prediction indicators are defined as follows:

- **Rolling mean:** The rolling mean characterizes the average level of microseismic energy within the region and reflects the energy release behavior associated with each coal-cutting cycle during face advancement.

- Local energy range: The local energy range captures the amplitude of local energy fluctuations, facilitating the identification of advancement segments exhibiting abnormally intense energy release.
- Local cumulative energy: Local cumulative energy reveals the influence of local energy accumulation on system stability.
- Local average energy: Local average energy mitigates the influence of extreme values on local feature representation, thereby enhancing the stability of the energy distribution.
- Rolling standard deviation: The rolling standard deviation identifies high-variability regions, serving as an indicator of potential anomalies in energy release behavior.
- Global cumulative energy: Global cumulative energy describes the long-term energy accumulation trend throughout the face advancement process from a global perspective, revealing the overall dynamics of energy release in the system.

### 2.3 Rockburst Risk Classification Thresholds

To rationally define microseismic energy intervals corresponding to different levels of rockburst risk, the classification strategy was established based on the Gutenberg–Richter (G–R) relationship [15] and the method proposed by Xie et al. [16]. According to the G–R relationship, an energy–frequency curve of microseismic events can be constructed for the mine. This curve is typically divided into three distinct segments: an exponential distribution segment, a linear distribution segment, and a discrete distribution segment. The transition point between the exponential and linear segments was identified as the early warning threshold for low rockburst risk. By further analyzing the two transition points between the linear distribution segment and the discrete distribution segment, the corresponding energy levels were determined as the early warning thresholds for medium and high rockburst risk, respectively. Figure 4 presents the energy–frequency curve with the three key transition energy levels clearly marked. The resulting classification of rockburst risk levels based on microseismic energy thresholds is summarized in Table 1.



**Figure 4.** Energy-frequency curve

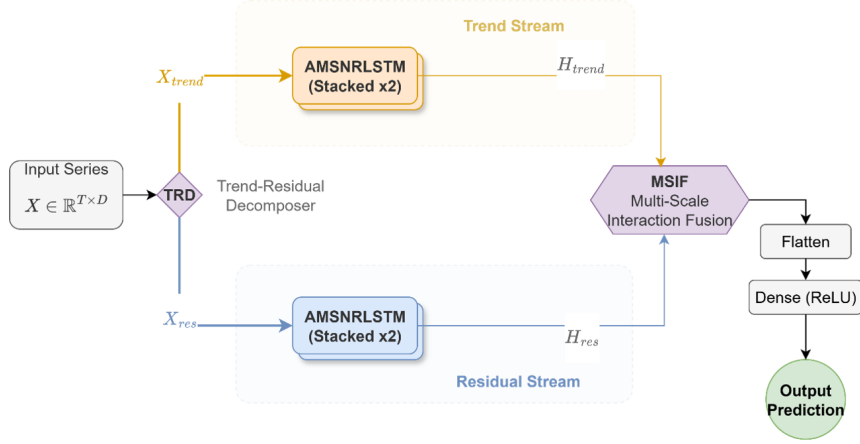
**Table 1.** Classification of rockburst risk levels based on energy thresholds

Energy Level (J)	Risk Level
$\leq 5605.29$	No risk
5605.29-5956.62	Low risk
5956.62-6776.42	Medium risk
$\geq 6776.42$	High risk

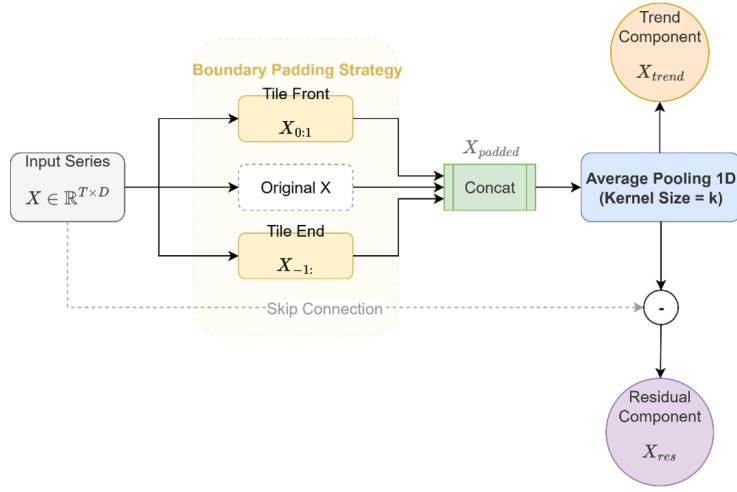
### 3 TRSNet Model Construction

The TRSNet model is formulated as an end-to-end deep learning framework, and its overall architecture is illustrated in Figure 5. The model architecture is designed in a clear and modular manner and consists of three principal stages: decomposition, dual-branch modeling with interaction, and reconstruction.





**Figure 5.** Architecture of the TRSNet model



**Figure 6.** Architecture of the trend-residual decomposition module

### 3.1 Trend-Residual Decomposition Module

Microseismic signals, as highly nonlinear data, typically comprise patterns at multiple scales, including a long-term low-frequency trend associated with gradual stress accumulation and short-term high-frequency fluctuations induced by localized rock fracturing events [17]. In terms of data characteristics, the long-term stress accumulation stage is reflected by slowly varying sequences with pronounced trends, whereas the short-term stress release stage is characterized by intense fluctuations without a clearly defined trend [18]. Owing to the coexistence of these fundamentally different dynamics, direct modeling of the original microseismic sequence often prevents learning models from effectively capturing both patterns simultaneously, thereby resulting in degraded predictive performance.

To address this challenge, a trend-residual decomposition module was introduced, drawing inspiration from decomposition-reconstruction paradigms in signal processing. The module is designed to decouple the complex input sequence into two complementary components: A trend component, representing long-term and smooth temporal evolution, and a residual component, capturing short-term and highly fluctuating dynamics [19]. Subsequent model components were then equipped with independent parameter sets to model these two components separately, thereby implementing a divide-and-conquer modeling strategy. The overall architecture of the trend-residual decomposition module is illustrated in Figure 6.

### 3.2 Improved Adaptive Multiscale Noise-Resilient LSTM

Although LSTM networks [20] have demonstrated strong performance in modeling time-series data, explicit mechanisms for noise suppression are generally absent [21], and a single memory unit is often insufficient to accommodate multiscale dynamic data [22].

To overcome these limitations, an Adaptive Multiscale Noise-Resilient LSTM cell (AMSNRLSTMCell) was developed. The proposed AMSNRLSTMCell was constructed upon the standard LSTM gating architecture and was

further extended through targeted structural enhancements. The core design principle of the AMSNRLSTMCell is twofold. First, dynamic noise suppression is introduced prior to memory-state updates, enabling the noise level of the input signal to be actively assessed and attenuated. Second, multiple memory units with distinct temporal decay scales are maintained in parallel to dynamically fuse multiscale memory representations. The overall architecture of the AMSNRLSTMCell is illustrated in Figure 7.

To mitigate the pronounced noise sensitivity of standard LSTM architectures, the AMSNRLSTMCell maintains a dynamically updated noise coefficient, denoted as  $\sigma_t$ , which is adaptively learned in a data-driven manner. This coefficient is computed by the model based on the correlation between the historical hidden state (query) and the current input representation (key). When the current input contains a substantial amount of information that is weakly correlated with the historical state, the value of  $\sigma_t$  approaches zero, ensuring that ineffective or noisy inputs do not affect the memory units.

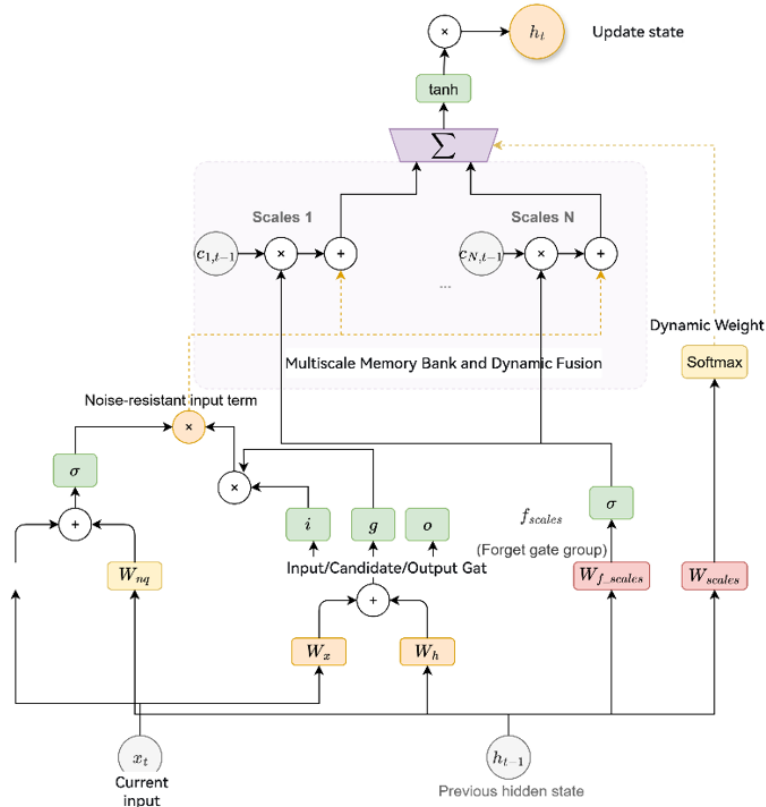
To capture dependencies at multiple scales, the AMSNRLSTMCell maintains  $N_s$  independent memory states in parallel. Each memory state is equipped with an independent forget gate, denoted as  $f_t^s$ , thereby enabling distinct memory decay rates across different states. After memory updates at all scales are completed, the multiscale memory representations are integrated through fusion to produce a unified final cell state, denoted as  $C_t^{final}$ .

### 3.3 MSIF Module

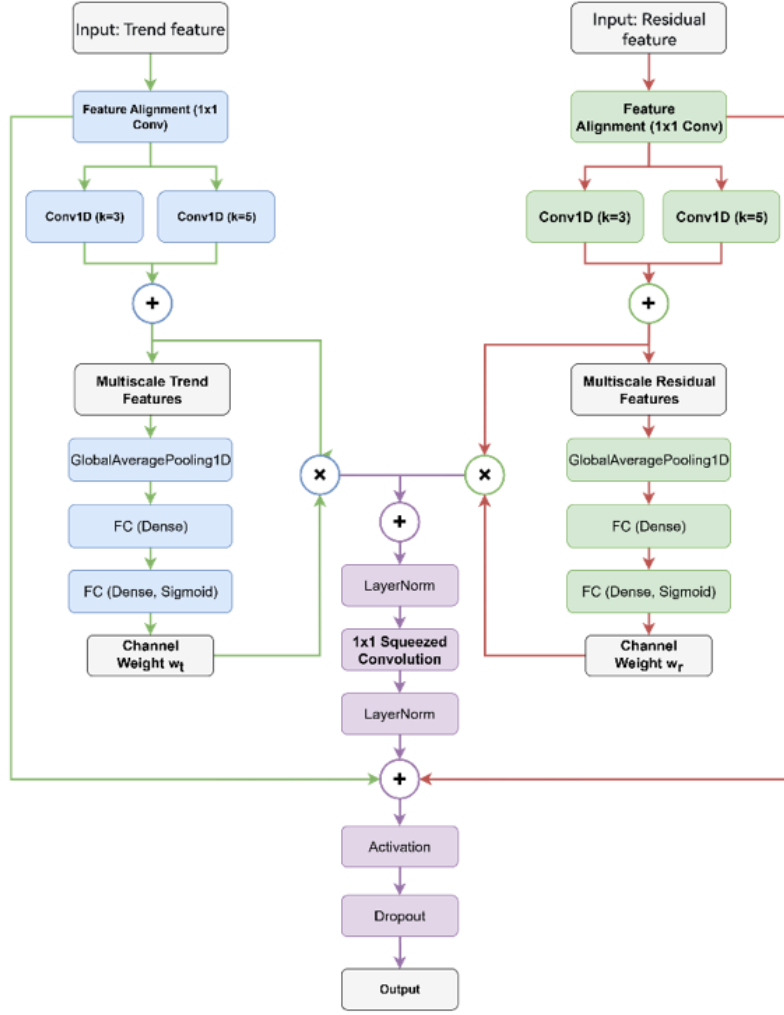
Following the dual-branch processing, the extracted trend features and residual features require effective integration. Simple operations such as feature concatenation or element-wise summation fail to adequately capture the nonlinear interactions between these two feature spaces [23]. To address this limitation, an MSIF module was designed, as illustrated in Figure 8, adopting the strategies below.

First, a  $1 \times 1$  convolution was applied to project the trend and residual branches into a unified feature dimensionality. Subsequently, multiple groups of parallel convolutions with different sizes were employed to extract deep features, after which the representations from both branches were aggregated to form a joint feature set. To emphasize key features, two fully independent channel attention mechanisms were introduced, enabling the adaptive reweighting of feature channels for the trend and residual branches, respectively.

Finally, the attention-enhanced trend and residual features were combined via element-wise summation, followed by a  $1 \times 1$  convolution to further compress and reduce the fused feature dimensionality, restoring it to the original feature shape. This operation facilitates residual connections and supports the generation of the final prediction output.



**Figure 7.** Architecture of the AMSNRLSTMCell



**Figure 8.** Architecture of the MSIF module

## 4 Experimental Design and Hyperparameter Optimization

### 4.1 Experimental Design

The experimental investigation is designed to comprehensively evaluate the performance and effectiveness of the proposed TRSNet framework in practical rockburst prediction tasks. Accordingly, three categories of experiments were conducted: comparative experiments, rockburst risk classification experiments, and ablation studies. These experiments are designed to assess the proposed model from complementary perspectives, thereby verifying both its predictive capability and its applicability to real engineering scenarios.

The primary objective of the comparative experiments is to benchmark the proposed model against five representative deep learning models that are widely adopted in the field of rockburst prediction. These baseline models include CNN–LSTM [12], LSTM [24], Deep Neural Network (DNN) [25], CNN–Bidirectional Long Short-Term Memory (BiLSTM)–Attention [26], and CNN–Bidirectional Gated Recurrent Unit (BiGRU)–Attention [27]. Performance comparisons were conducted under identical datasets and evaluation metrics to assess model effectiveness. To guarantee evaluation objectivity, the dataset was strictly partitioned into training, validation, and testing subsets according to a 7:1:2 ratio. All input data were processed using min–max normalization, facilitating faster model convergence.

To ensure experimental rigor, a unified training strategy was adopted for all models. During training, the Adam optimizer and the mean squared error (MSE) loss function were employed, while ReLU was used as the activation function. The batch size was fixed at 32 for all experiments. In addition, an early-stopping mechanism was introduced to mitigate overfitting, whereby training was terminated if the validation loss did not improve for 10 consecutive epochs. Hyperparameters for all models were automatically optimized using a Bayesian algorithm, with the maximum number of trials set to 50.

## 4.2 Evaluation Metrics

To evaluate model performance, four quantitative evaluation metrics were selected to assess both regression prediction accuracy and classification performance, namely, the Mean Absolute Error (MAE), MSE, recall, and the False Positive Rate (FPR). In general, lower values of MAE and MSE indicate superior regression prediction accuracy, whereas higher recall coupled with lower FPR reflects better classification performance.

## 4.3 Comparative Analysis of Model Performance

The quantitative performance of all models is summarized in Table 2 and visually illustrated in Figure 9.

Based on the above experimental results, it is evident that the TRSNet model outperforms all other models across all evaluation metrics. TRSNet achieves the lowest MAE and MSE, while also attaining a recall of 88.37% and an FPR of 3.88%.

At the same time, the five baseline models exhibit a trend where simpler models outperform more complex models. Both the DNN and LSTM models show significantly better performance compared to the three more complex baseline models. This is likely due to the larger number of parameters in the complex models, which makes them more prone to overfitting when trained on noisy data.

## 4.4 Risk Level Classification Performance Evaluation

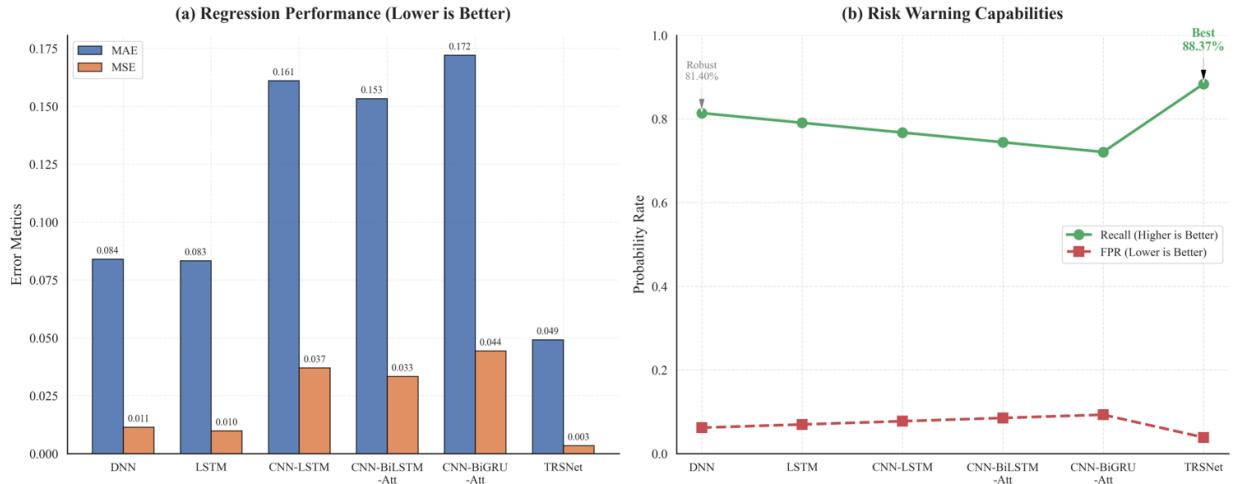
To visually assess the ability of each model to identify different risk levels, Figure 10 presents the confusion matrices for all models.

The confusion matrix of TRSNet exhibits the most pronounced diagonal dominance. Overall, TRSNet demonstrates a clear and significant advantage over the other baseline models. Misclassifications produced by TRSNet are primarily concentrated between adjacent risk categories. Consequently, if the task is simplified to a binary classification problem distinguishing between risk and no risk, the performance advantage of TRSNet would become even more pronounced.

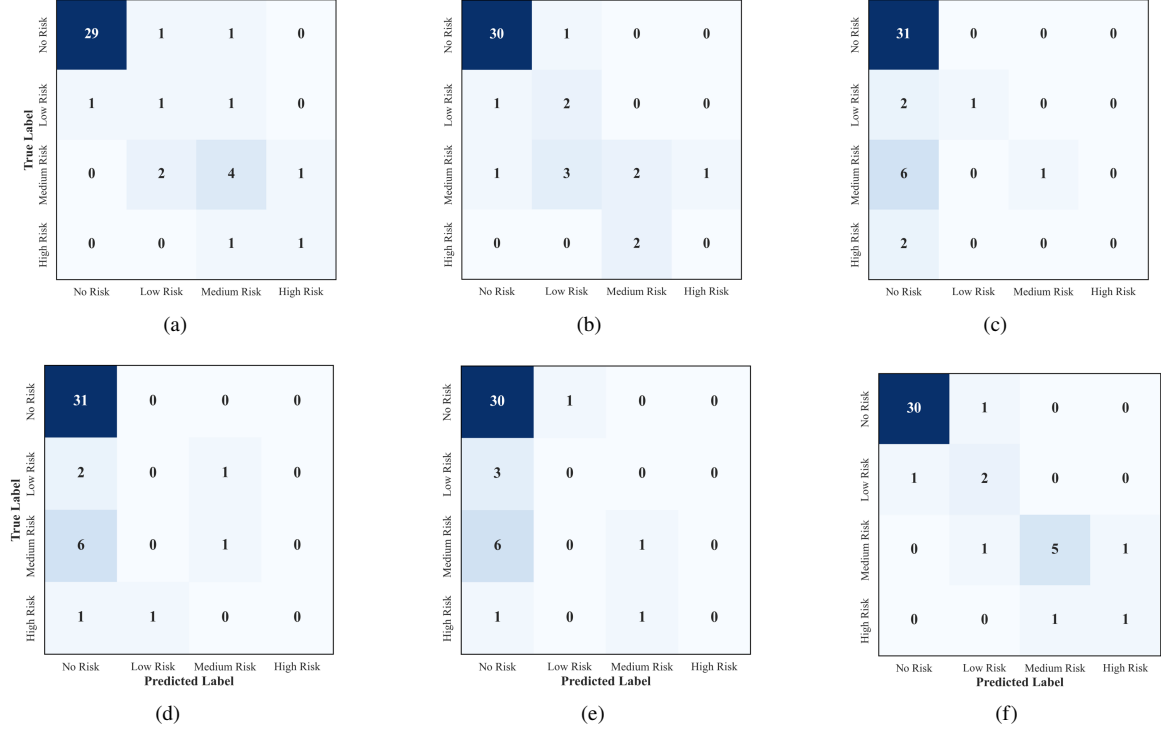
By contrast, several baseline models exhibit potentially hazardous prediction behaviors when handling high-risk samples. In particular, the CNN-BiGRU-Attention model fails to accurately identify high-risk events and further displays overly conservative prediction tendencies across other risk levels. As a result, events associated with elevated risk are frequently misclassified as no-risk cases.

**Table 2.** Performance comparison of different models

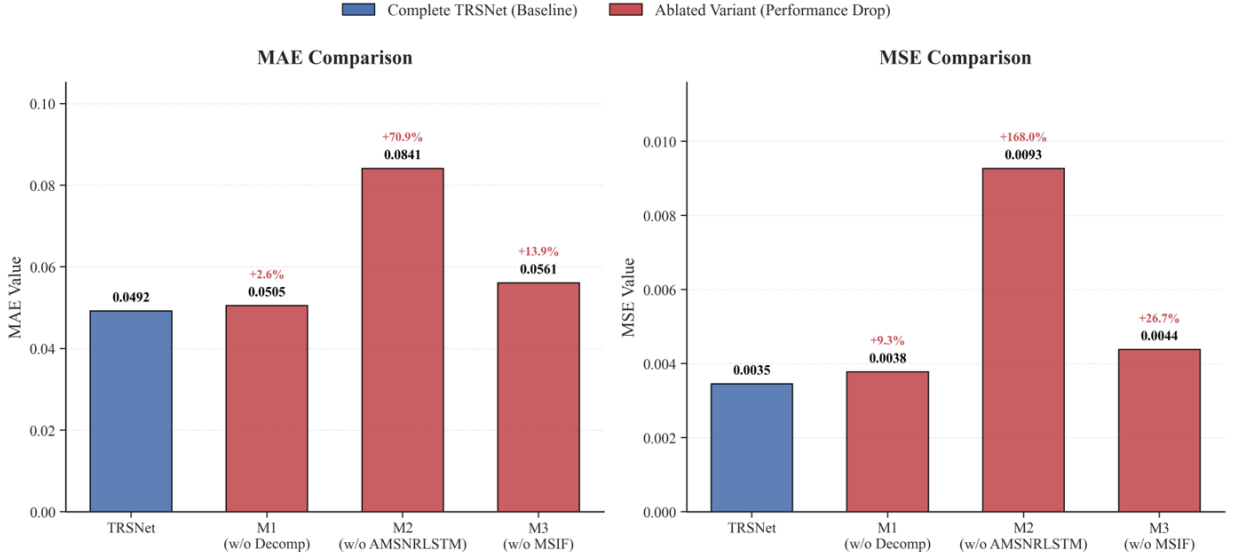
Model	MAE	MSE	Recall	FPR
LSTM	0.083380	0.009895	79.07%	6.98%
DNN	0.084041	0.011477	81.40%	6.20%
CNN-LSTM	0.161116	0.037055	76.74%	7.75%
CNN-BiLSTM-Attention	0.153319	0.033328	74.42%	8.53%
CNN-BiGRU-Attention	0.172084	0.044407	72.09%	9.30%
TRSNet	0.049231	0.003459	88.37%	3.88%



**Figure 9.** Quantitative performance evaluation of different prediction models



**Figure 10.** Confusion matrix comparison across models: (a) DNN, (b) LSTM, (c) CNN-LSTM, (d) CNN-BiLSTM-Attention, (e) CNN-BiGRU-Attention, and (f) TRSNet (Ours)



**Figure 11.** Comparison results of ablation experiments

#### 4.5 Ablation Study Analysis

To quantify the contribution of each functional module within the proposed framework, a series of ablation experiments was conducted. The following ablation settings were designed:

- M1: The decomposition module is removed, and a single-branch input structure is adopted.
- M2: The AMSNRLSTM is replaced with a standard LSTM model.
- M3: The MSIF module is removed and replaced with simple element-wise summation.

The experimental results of the ablation studies are presented in Figure 11.

Based on the ablation results, when the core AMSNRLSTM unit is replaced with a standard LSTM unit, the MAE increases by 70.86%, while the MSE rises sharply by 168.03%, providing strong evidence that the AMSNRLSTM

module constitutes the foundational component enabling the superior performance of TRSNet.

When the MSIF module is removed (M3) and the two branches are fused using simple element-wise summation, noticeable increases in both MAE and MSE are also observed. This outcome indicates that the relationship between the trend branch and the residual branch is not a simple linear superposition. Element-wise addition alone is insufficient to effectively integrate the information from the two branches. For M1, in which the decomposition module is removed, the resulting performance degradation is smaller than that observed in M2 and M3. Nevertheless, the performance decline remains evident, thereby confirming the effectiveness of the divide-and-conquer modeling strategy.

## 5 Conclusions

To address the core challenges in rockburst prediction—namely the low signal-to-noise ratio and strong non-stationarity of microseismic data—an integrated TRSNet framework was developed by combining physics-inspired decomposition principles with deep learning techniques. Through comprehensive analysis and experimental validation using field microseismic data from a northern coal mine, the following main conclusions were drawn:

- Experimental results demonstrate that independent modeling of long-term trends and short-term residuals significantly reduces learning difficulty for predictive models. Ablation studies confirm that removal of the trend–residual decomposition module leads to a clear degradation in prediction performance.
- The proposed AMSNRLSTM cell, by introducing a dynamic noise coefficient and multiscale memory states, effectively addresses the limitations of standard LSTM architectures, including noise sensitivity and single-scale representation. This component is identified as a key contributor to performance improvement, as its removal results in a sharp 168.03% increase in MSE.
- In comparative evaluations against baseline models, TRSNet consistently achieves superior performance in both regression accuracy and classification effectiveness. In particular, for engineering-critical rockburst risk level assessment, a high recall of 88.37% is achieved alongside a low FPR of 3.88%.
- Despite the superior performance demonstrated by TRSNet in the conducted experiments, several limitations remain and should be addressed in future research.
- The dataset utilized in this study is derived from a single longwall mining face in a northern coal mine. Given the substantial variability in geological structures and mining depths across different mining regions, the generalization capability of the proposed model requires further validation. Future studies should therefore investigate validation and transfer learning strategies using microseismic data collected under diverse geological conditions.
- At present, the proposed framework primarily relies on time-series data derived from microseismic energy and event frequency. However, the occurrence of rockburst is also closely associated with acoustic emission, electromagnetic radiation, in-situ stress monitoring, and coal seam geological structures. The incorporation of multimodal data fusion techniques is therefore recommended in future work to further enhance the robustness of rockburst prediction.

## Author Contributions

Conceptualization, Q.X. and H.W.; methodology, Q.X.; software, Q.X.; validation, Q.W., B.Z. and J.G.; formal analysis, Q.X. and H.G.; investigation, J.G. and C.X.; resources, Q.W. and B.Z.; data curation, Q.X.; writing—original draft preparation, Q.X.; writing—review and editing, H.W.; visualization, Q.X.; supervision, H.W.; project administration, H.W.; funding acquisition, H.W. All authors have read and agreed to the published version of the manuscript.

## Data Availability

The data presented in this study are available from the corresponding author upon reasonable request and with approval from our industry partner.

## Conflicts of Interest

The authors declare no conflict of interest.

## References

- [1] F. Jiang, Y. Feng, and Y. Liu, “Dynamic evaluation method for rockburst risk before stopping,” *Chin. J. Rock Mech. Eng.*, vol. 33, no. 10, pp. 2101–2106, 2014. <https://doi.org/10.13722/j.cnki.jrme.2014.10.017>
- [2] Y. Lei, “Study on establishment and application of hazard evaluation model for rock-burst,” Ph.D. dissertation, China Coal Res. Inst., Beijing, China, 2005.
- [3] L. Wojtecki, S. Iqbal, M. Dyduch, M. Bukowska, and J. Makówka, “Use of machine learning algorithms to assess the state of rockburst hazard in underground coal mine openings,” *J. Rock Mech. Geotech. Eng.*, vol. 14, no. 3, pp. 703–713, 2022. <https://doi.org/10.1016/j.jrmge.2021.10.011>



- [4] H. Guo and C. Y. Yang, "Prediction of rock burst based on PCA-RF model," *J. Harbin Univ. Commer. (Nat. Sci. Ed.)*, vol. 39, no. 4, pp. 413–418, 2023. <https://doi.org/10.19492/j.cnki.1672-0946.2023.04.016>
- [5] C. Shi, F. Gao, and L. Chen, "Prediction of pressure bump in coal mine by PCA-GRNN," *China Saf. Sci. J.*, vol. 26, no. 7, pp. 119–124, 2016. <https://doi.org/10.16265/j.cnki.issn1003-3033.2016.07.022>
- [6] X. Jiang, X. Liu, J. Li, P. Feng, B. Song, H. Zhu, and C. Zong, "Prediction method of rock burst in drivage roadway in steeply inclined ultra-thick coal seam based on machine learning," *J. Min. Strata Control Eng.*, vol. 6, no. 3, pp. 95–107, 2024. <https://doi.org/10.13532/j.jmsce.cn10-1638/td.20240520.002>
- [7] X. Fu, S. Chen, and T. Zhang, "Prediction of rock bursts based on microseismic energy change: Application of bayesian optimization–long short-term memory combined model," *Appl. Sci.*, vol. 14, no. 20, p. 9277, 2024. <https://doi.org/10.3390/app14209277>
- [8] Y. Ding, E. Wang, Z. Li, X. Liu, T. Huang, and J. Yao, "Comprehensive early warning method of microseismic, acoustic emission, and electromagnetic radiation signals of rock burst based on deep learning," *Int. J. Rock Mech. Min. Sci.*, vol. 170, p. 105519, 2023. <https://doi.org/10.1016/j.ijrmms.2023.105519>
- [9] F. Cui, C. Zong, X. P. Lai, S. F. He, S. L. Zhang, and C. Jia, "Intelligent prediction of time series and grade of rock burst in steeply inclined ultra-thick coal seam excavation roadway," *J. China Coal Soc.*, vol. 50, no. 2, pp. 845–861, 2025. <https://doi.org/10.13225/j.cnki.jccs.2023.1762>
- [10] Y. H. Liang, F. B. Shen, Z. D. Xie, and J. F. Wu, "Research on prediction method of rock burst based on LSTM model," *China Min. Mag.*, vol. 32, no. 5, pp. 88–95, 2023. <https://doi.org/10.12075/j.issn.1004-4051.2023.05.017>
- [11] A. Y. Cao, Y. Q. Liu, X. Yang, L. S., C. B. Wang, X. Q. Bai, and Y. P. Liu, "Physical index and data fusion-driven method for coal burst prediction in time sequence," *J. China Coal Soc.*, vol. 48, no. 10, pp. 3659–3673, 2023. <https://doi.org/10.13225/j.cnki.jccs.2022.0680>
- [12] H. M. Liu, F. Y. Xu, B. J. Liu, and M. Deng, "Time-series prediction method for risk level of rockburst disaster based on CNN–LSTM," *J. Cent. South Univ. (Sci. Technol.)*, vol. 52, no. 3, pp. 1–12, 2021. <https://doi.org/10.11817/j.issn.1672-7207.2021.03.001>
- [13] Y. Gong, R. S. Jia, X. M. Lu, Y. J. Peng, W. D. Zhao, and X. L. Zhang, "To suppress the random noise in microseismic signal by using empirical mode decomposition and wavelet transform," *J. China Coal Soc.*, vol. 43, no. 11, pp. 1–9, 2018. <https://doi.org/10.13225/j.cnki.jccs.2017.1667>
- [14] Z. Zhang, Y. Ye, B. Luo, G. Chen, and M. Wu, "Investigation of microseismic signal denoising using an improved wavelet adaptive thresholding method," *Sci. Rep.*, vol. 12, no. 1, p. 22186, 2022. <https://doi.org/10.1038/s41598-022-26576-2>
- [15] B. Gutenberg and C. F. Richter, "Frequency of earthquakes in California," *Bull. Seismol. Soc. Am.*, vol. 34, no. 4, pp. 185–188, 1944. <https://doi.org/10.1038/156371a0>
- [16] J. H. Xie, Y. Z. Zhang, Y. Zhang, G. L. Ding, C. H. Shi, and R. Yao, "Optimization of microseismic energy early-warning index based on energy level and frequency analysis," *Coal Eng.*, vol. 53, no. 10, pp. 67–72, 2021. <https://doi.org/10.11799/ce202110014>
- [17] J. Kan, L. Dou, J. Li, X. Li, J. Bai, and M. Wang, "Characteristics of microseismic waveforms induced by underground distress blasting: Comparison with those induced by ground blasting and coal mining," *Front. Earth Sci.*, vol. 10, no. 10, p. 797358, 2022. <https://doi.org/10.3389/feart.2022.797358>
- [18] K. Ma, H. Wang, Z. Liao, Y. Peng, and K. Wang, "Precursor of microseismic energy and stress evolution induced by rockburst in coal mining: A case study from Xiashijie, Shaanxi, China," *Geomech. Geophys. Geo-Energy Geo-Resour.*, vol. 8, no. 5, p. 134, 2022. <https://doi.org/10.1007/s40948-022-00435-w>
- [19] Z. Zhang, Y. Ye, B. Luo, G. Chen, and M. Wu, "Investigation of microseismic signal denoising using an improved wavelet adaptive thresholding method," *Sci. Rep.*, vol. 12, no. 1, p. 22186, 2022. <https://doi.org/10.1038/s41598-022-26576-2>
- [20] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997. <https://doi.org/10.1162/neco.1997.9.8.1735>
- [21] H. Yang, Z. Pan, and Q. Tao, "Robust and adaptive online time series prediction with long short-term memory," *Comput. Intell. Neurosci.*, vol. 2017, pp. 1–9, 2017. <https://doi.org/10.1155/2017/9478952>
- [22] J. Liu, F. Rahmani, K. Lawson, and C. Shen, "A multiscale deep learning model for soil moisture integrating satellite and in situ data," *Geophys. Res. Lett.*, vol. 49, no. 7, p. e2021GL096847, 2022. <https://doi.org/10.1029/2021GL096847>
- [23] Y. Dai, F. Gieseke, S. Oehmcke, Y. Wu, and K. Barnard, "Attentional feature fusion," 2020. <https://doi.org/10.48550/arXiv.2009.14082>
- [24] X. Fu, S. Chen, and T. Zhang, "Prediction of rock bursts based on microseismic energy change: Application of bayesian optimization–long short-term memory combined model," *Appl. Sci.*, vol. 14, no. 20, p. 9277, 2024. <https://doi.org/10.3390/app14209277>

- [25] Y. Zhang, M. Zhang, J. Li, and G. Chen, "Rockburst intensity grade prediction model based on batch gradient descent and multi-scale residual deep neural network," *Comput. Syst. Sci. Eng.*, vol. 47, no. 2, pp. 1987–2006, 2023. <https://doi.org/10.32604/csse.2023.040381>
- [26] F. Ren, Z. Tian, K. Ma, and J. Gu, "Relationship between rock brittleness and rockburst early warning time," *Int. J. Rock Mech. Min. Sci.*, vol. 195, p. 106305, 2025. <https://doi.org/10.1016/j.ijrmms.2025.106305>
- [27] H. Liu, T. Ma, Y. Lin, K. Peng, X. Hu, S. Xie, and K. Luo, "Deep learning in rockburst intensity level prediction: Performance evaluation and comparison of the NGO-CNN-BiGRU-attention model," *Appl. Sci.*, vol. 14, no. 13, p. 5719, 2024. <https://doi.org/10.3390/app14135719>