

Article

A Comparative Study of Deep Learning-Based Models for Mine Water Inflow Prediction

Lujun Chai ^{1,2,†}, Zhuolin Li ^{3,†} , Fanjun Wang ^{3,*}, Xin Huang ^{4,*}  and Kun Wu ^{1,2}

¹ Technology Innovation Center for Groundwater Disaster Prevention and Control Engineering for Metal Mines, Ministry of Natural Resources, Shijiazhuang 050024, China; sophyer@126.com (L.C.); shuluoye@126.com (K.W.)

² North China Engineering Investigation Institute Co., Ltd., Shijiazhuang 050024, China

³ Academy of Eco-Civilization Development for Jing-Jin-Ji Megalopolis, Tianjin Normal University, Tianjin 300387, China; 18713591593@163.com

⁴ College of Computer and Information Engineering, Tianjin Normal University, Tianjin 300387, China

* Correspondence: 13012297920@163.com (F.W.); huxpiero@tjnu.edu.cn (X.H.)

† These authors contributed equally to this work.

Abstract

To address the need for high-precision prediction of mine water inflow under complex geological conditions, this study proposes a hybrid deep learning framework for accurate water inrush forecasting. Using daily water inrush records and nine meteorological and hydrological variables collected from the Maoping mining area in northeastern Yunnan Province, China, in 2023, five deep learning models, namely, CNN, LSTM, Transformer, CNN-LSTM, and LSTM-Transformer, were systematically developed and comparatively evaluated. The results demonstrate that the LSTM-Transformer hybrid model achieved the best predictive performance, with an MAE of 0.196, an RMSE of 0.243, and an R^2 of 0.882 outperforming both the individual deep learning models and other hybrid architectures. By combining the temporal memory capability of LSTM with the global attention mechanism of Transformer, the proposed model effectively captures the nonlinear, non-stationary, and multi-scale temporal dependencies embedded in mine water inflow series, thereby substantially enhancing prediction accuracy. These findings indicate that hybrid deep learning architectures integrating local temporal memory with global attention mechanisms provide a robust and effective approach for complex engineering time-series forecasting, offering valuable support for high-precision early warning and intelligent prevention of mine water inflow hazards.

Keywords: mine water inflow prediction; time series analysis; deep learning; LSTM-Transformer hybrid model; mine water hazard prevention and control



Academic Editor: Liliana Lefticariu

Received: 29 June 2026

Revised: 4 August 2026

Accepted: 5 August 2026

Published: 7 August 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

As a critical technical measure for mine water hazard prevention and control, mine water inflow prediction plays a vital role in preventing water-related accidents, ensuring personnel safety, and maintaining the economic viability of mining operations [1,2]. Mine water inflow refers to the volume of water entering mine workings from sources such as surface water and groundwater during mining activities. A sudden increase in water inflow can lead to severe consequences, including roadway flooding, equipment damage, and even casualties. With the increasing depth of mining operations and the growing complexity of geological conditions, disturbances to aquifers and the development of

water-conducting fractures have intensified the nonlinear and uncertain characteristics of mine water inflow. Consequently, conventional prediction methods are often inadequate to meet the requirements of accurate and timely water hazard prevention and control [3,4]. Therefore, developing efficient and accurate mine water inflow prediction models is of considerable theoretical and practical significance for establishing effective water hazard early-warning systems and ensuring safe mining operations [5].

In the development of mine water inflow prediction methods, two major technical approaches have gradually emerged: traditional physics-based models and modern data-driven models. Among these, machine learning and deep learning methods have attracted increasing attention because of their powerful nonlinear fitting capabilities and ability to capture complex temporal dependencies [6]. Early machine learning approaches, such as Backpropagation (BP) neural networks and Support Vector Regression (SVR), have been applied to mine water-related prediction tasks [7]. For example, Yu et al. employed a BP neural network to accurately predict the depth of floor failure induced by underground mining pressure, thereby providing technical support for aquifer water hazard prevention and the identification and control of potential water-conducting channels [8]. Meenal et al. proposed an SVR-based machine learning approach for prediction; however, its practical performance remains highly dependent on the appropriate selection and optimization of hyperparameters [9,10].

With the rapid development of deep learning, various neural network architectures have been increasingly applied to mine water inflow and related hydrological prediction tasks. Long Short-Term Memory (LSTM) networks have been widely adopted because of their ability to learn long-term dependencies and temporal patterns in sequential data [11]. As a simplified variant of LSTM, the Gated Recurrent Unit (GRU) employs a more streamlined gating structure while retaining the ability to model long-term temporal dependencies and has demonstrated competitive predictive performance in various time-series forecasting tasks [12,13]. In contrast to recurrent neural networks, the Transformer architecture processes sequential data through self-attention mechanisms, enabling parallel computation and efficient modeling of long-range dependencies [14]. In particular, its multi-head self-attention mechanism can capture complex interactions among variables and temporal features, demonstrating considerable potential in a wide range of time-series forecasting applications [15,16]. For instance, Liu et al. developed a dual-encoder Transformer model for long-term, high-precision runoff forecasting in the Yangtze River Basin and achieved promising predictive performance [17].

In recent years, hybrid modeling and multi-technology integration have emerged as important strategies for further improving prediction accuracy and model robustness. Hybrid models are designed to integrate two or more complementary algorithms or neural network architectures, thereby exploiting their respective strengths—such as combining local feature extraction with global dependency modeling—while mitigating the limitations inherent in single models [18]. For example, Xu et al. integrated particle swarm optimization with an LSTM model, significantly improving the accuracy and forecasting horizon of flood prediction and demonstrating the advantages of the hybrid approach over conventional hydrological models [19]. Choubin et al. developed a hybrid ANN-ARIMA model for predicting groundwater-level fluctuations in semi-arid regions and introduced the gamma test and M-test to optimize input variables and training data length, thereby substantially enhancing prediction accuracy [20]. Feng et al. proposed an LSTM-based network (AGCLSTM) that integrates graph convolution and spatiotemporal attention mechanisms, achieving excellent performance in both flood peak prediction and flow-process fitting [21]. These studies demonstrate that the integration of complementary modeling strategies can

effectively enhance the ability of deep learning models to characterize complex nonlinear relationships and temporal dependencies in hydrological systems.

Despite these advances, a significant research gap remains regarding the optimal integration of advanced deep learning architectures specifically for the unique hydrological challenges of deep mining environments. Accurate prediction of mine water inflow under complex geological and meteorological-hydrological conditions remains challenging. Mine water inflow is influenced by multiple interacting factors and exhibits pronounced nonlinear, non-stationary, and multi-scale temporal characteristics. Current literature has not sufficiently explored how to effectively couple the local feature extraction capabilities of Convolutional Neural Networks (CNN) with the global sequence modeling strengths of Transformers and LSTMs to address these specific complexities. Therefore, further investigation is needed to determine whether specific hybrid deep learning architectures can effectively integrate these distinct capabilities to improve prediction performance in mining scenarios.

Addressing this gap, this study presents a novel comparative analysis and development of multiple deep learning frameworks for daily mine water inflow prediction in the Maoping mining area, Yiliang County, Zhaotong City, Yunnan Province, China, utilizing the Python 3.10 platform. The primary innovation of this research lies in the systematic construction and evaluation of hybrid architectures that combine CNN, LSTM, and Transformer models. Specifically, the predictive performance of standalone CNN, LSTM, and Transformer models is compared against their hybrid counterparts to identify the most effective configuration. The objective of this study is to establish a robust deep learning framework capable of accurately characterizing the complex temporal patterns of mine water inflow. By doing so, this research aims to provide reliable technical support for dynamic water hazard monitoring and early warning, thereby contributing to safer and more intelligent mining operations.

2. Materials and Methods

2.1. Study Area Overview

The Maoping lead–zinc mining area is located in Yiliang County, northeastern Yunnan Province, China, between $103^{\circ}51'$ E and $104^{\circ}45'$ E longitude and $27^{\circ}16'$ N and $27^{\circ}57'$ N latitude, covering an area of approximately 8.27 km². The study area is situated within the Yunnan–Guizhou Plateau and has been strongly influenced by regional tectonic activity. The terrain is characterized by pronounced topographic dissection, with a well-developed network of valleys and gullies. In general, the elevation gradually increases across the study area, giving rise to distinctive high-plateau canyon landscapes. The geographical location of the study area is presented in Figure 1.

2.2. Data Sources

All variables used in this study were obtained from conventional hydrological and meteorological monitoring systems, without relying on geological parameters that are difficult or costly to acquire. The mine water inflow data were obtained from daily monitoring records of mine inflow in the Maoping mining area during 2023. Daily hydrological data were collected from observations at the Longtan Spring hydrological station in the Maoping mining area, including runoff, water level, and discharge. Daily precipitation data were obtained from real-time observations at the Yiliang Meteorological Station ($27^{\circ}38'$ N, $104^{\circ}4'$ E) in Yunnan Province. Other meteorological variables, including average relative humidity, average wind speed, average surface temperature, average station pressure, and sunshine duration, were obtained from the China Meteorological Data Service Center (<https://data.cma.cn/>, accessed on 1 November 2025). To ensure data integrity and scien-

tific rigor, a rigorous preprocessing protocol was applied to the mine water inflow data and the nine meteorological/hydrological variables. This process involved outlier detection and removal, followed by linear interpolation to address missing values.

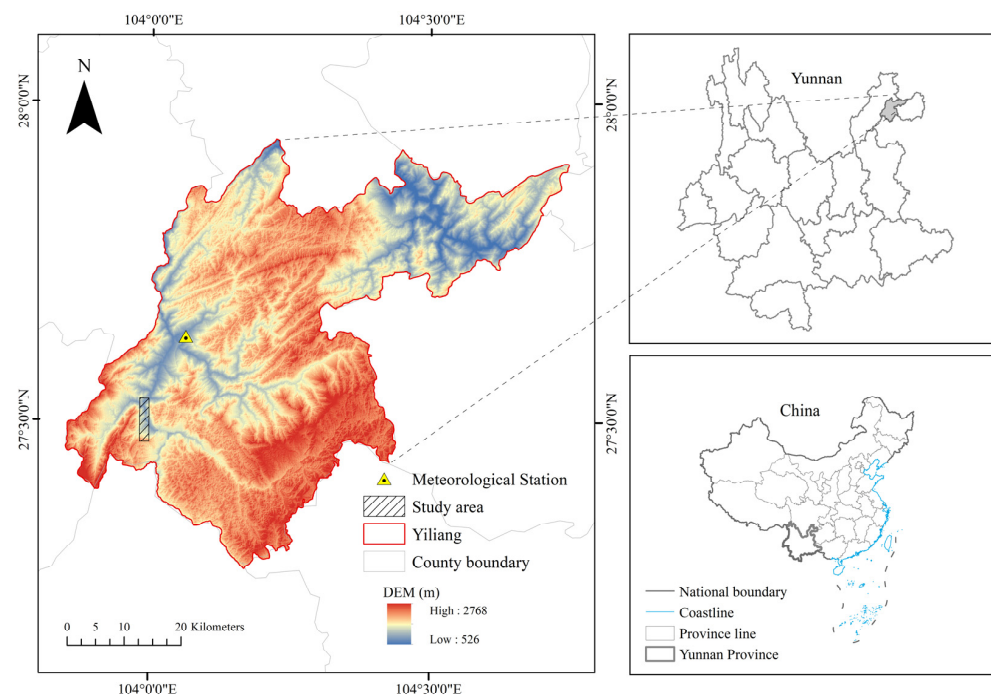


Figure 1. Location and setting of the study area.

Figure 2 presents the temporal variation in mine water inflow in the Maoping mining area during 2023. The mine water inflow exhibits distinct seasonal fluctuations, reaching its annual minimum in July and maximum in September. Figures 3 and 4 show the temporal variations in the hydrological and meteorological variables, respectively. Based on continuous observations from January to December 2023, the temporal variation characteristics of nine key variables, including average relative humidity, average wind speed, runoff, water level, discharge, precipitation, average surface temperature, average station pressure, and sunshine duration, were systematically analyzed. Overall, all nine variables exhibited pronounced temporal heterogeneity and seasonal variability, reflecting the complex temporal dynamics of the hydrological and meteorological conditions in the study area. These variations provide important environmental information for characterizing the temporal dynamics of mine water inflow and serve as the input variables for subsequent deep learning-based prediction.

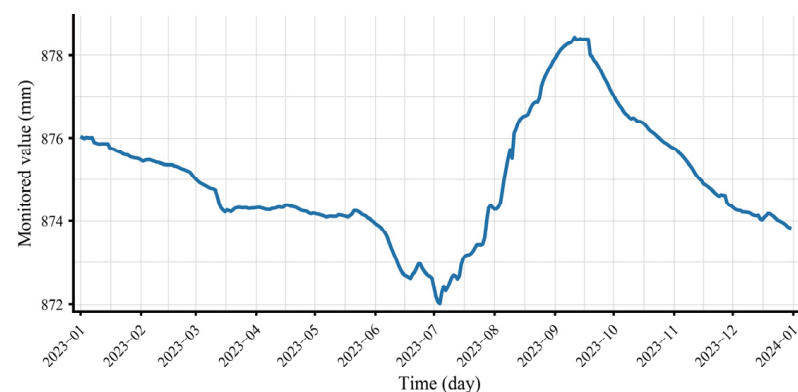


Figure 2. Daily observation data of mine water inflow in Maoping Mining Area in 2023.

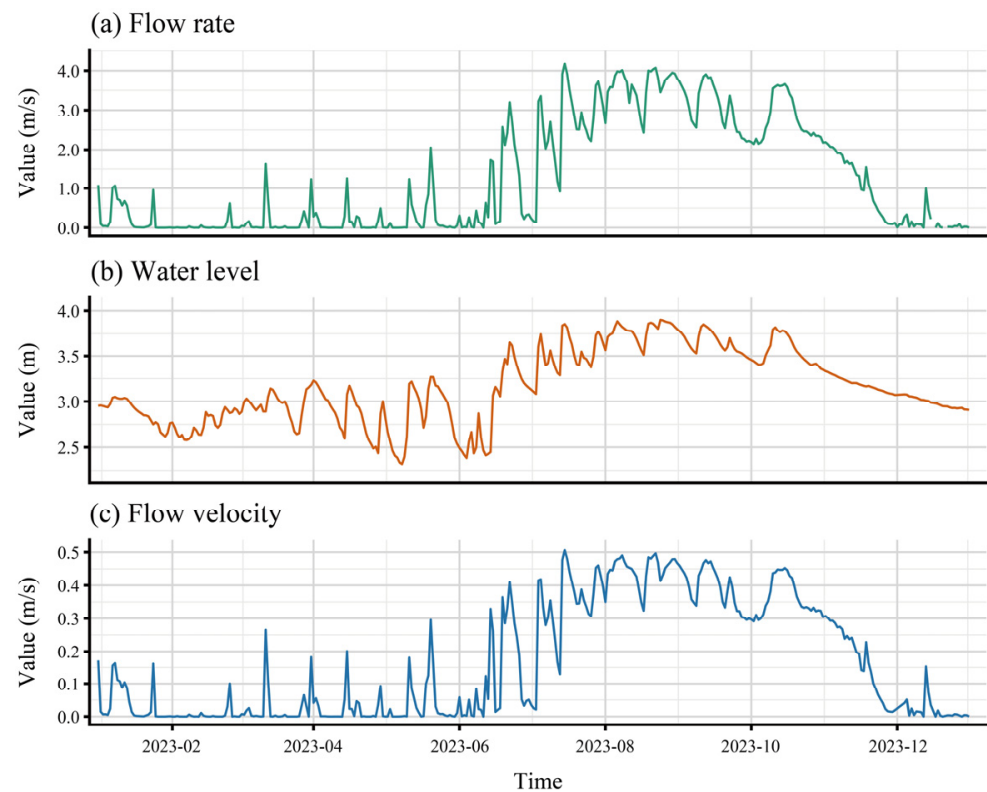


Figure 3. Hydrological parameters used in model construction: (a) Flow rate, (b) Water level, (c) Flow velocity.

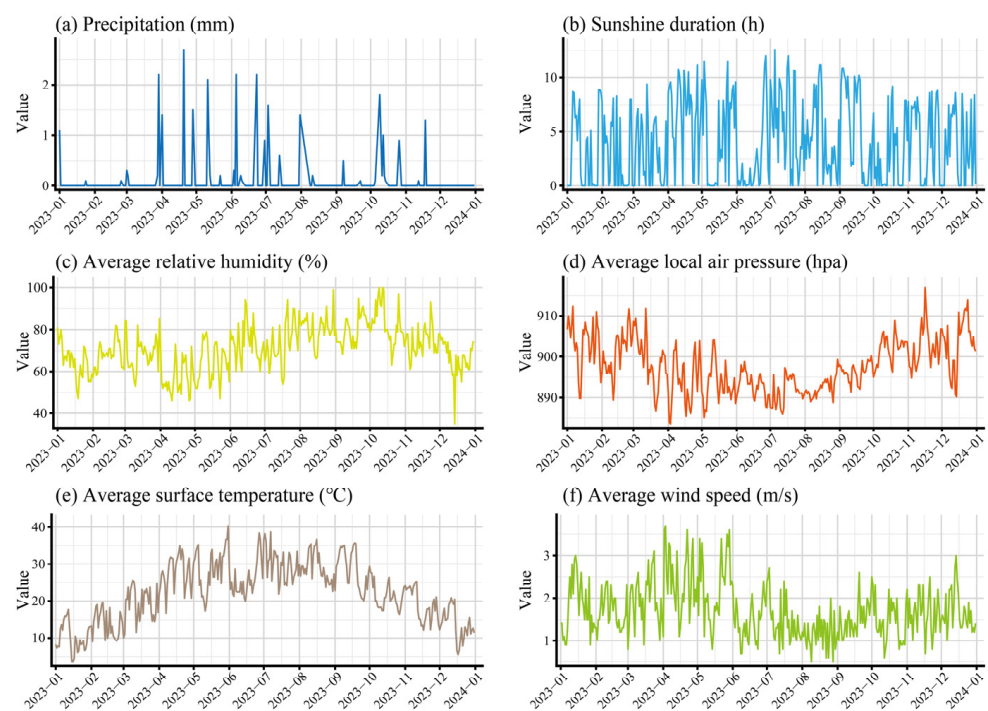


Figure 4. Meteorological parameters used in model construction: (a) Precipitation, (b) Sunshine duration, (c) Average relative humidity, (d) Average local air pressure, (e) Average surface temperature, (f) Average wind speed.

2.3. Deep Learning Model Construction

2.3.1. Convolutional Neural Network (CNN) Model

The Convolutional Neural Network (CNN) extracts local dependencies and short-term fluctuations in time series through convolutional layers. Pooling layers perform feature aggregation and dimensionality reduction to enhance model robustness, while fully connected layers integrate the extracted features and generate the final prediction [22].

In the prediction of mine water inflow, CNNs can effectively identify local patterns and periodic signals, maintaining strong representation and generalization capabilities even with limited training data [23].

2.3.2. Transformer Model

The Transformer is a neural network architecture built entirely on the self-attention mechanism. Its encoder–decoder structure enables it to efficiently capture long-term dependencies within sequences [24,25].

The self-attention mechanism serves as the core feature extraction component of the Transformer. By computing attention weights among different positions in the sequence, the model can assign varying levels of attention to different parts of the input, thereby capturing long-range dependency relationships.

2.3.3. LSTM Model

Long Short-Term Memory (LSTM) is an important variant of the Recurrent Neural Network (RNN). By introducing a gating mechanism, it effectively addresses the problems of gradient vanishing and gradient explosion that occur in traditional RNNs when learning long sequences [26].

An LSTM unit mainly consists of three core gate structures and a cell state responsible for information storage. The internal operational mechanism is illustrated in Equations (1)–(6) and Figure 5.

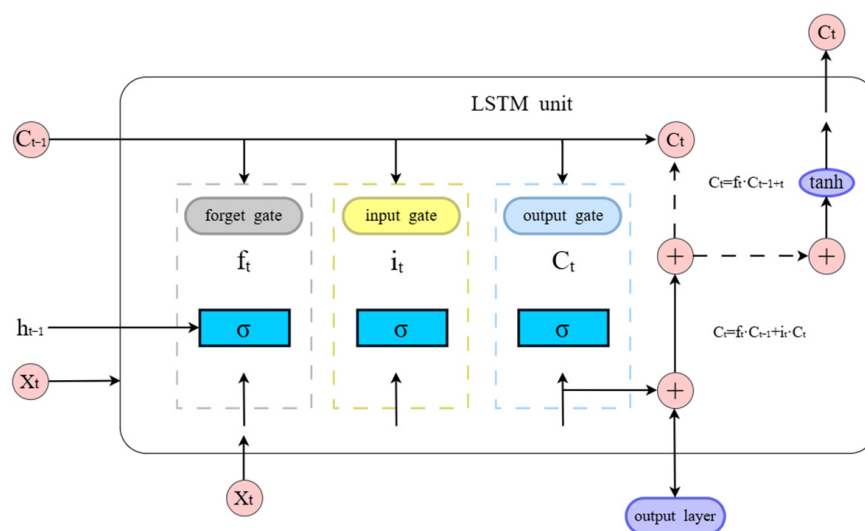


Figure 5. Structural diagram of the LSTM model.

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (1)$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (2)$$

$$\tilde{c}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (3)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (4)$$

$$o_t = (W_o[h_{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = o_t \odot \tanh(c_t) \quad (6)$$

where

$f_t, i_t, \tilde{c}_t, c_t, o_t, h_t$ represent the forget gate, input gate, and output gate, respectively;

c_t denotes the cell state;

h_t represents the hidden state;

σ is the sigmoid function, and $\tanh$ is the hyperbolic tangent function.

2.3.4. CNN-LSTM Hybrid Model

The CNN-LSTM model combines the advantages of CNN in spatial feature extraction and LSTM in temporal sequence modeling. Its core idea is to break through the limitations of a single model by integrating features from both spatial and temporal dimensions, thereby achieving accurate modeling of complex systems.

The front-end CNN module and the back-end LSTM module realize the integration of spatial-temporal information through feature transmission, forming a “spatial → temporal” information fusion process. The model usually consists of five parts: an input layer, a data preprocessing layer, a CNN layer, an LSTM layer, and an output layer [27]. The structural diagram of the CNN-LSTM model is illustrated in Figure 6.

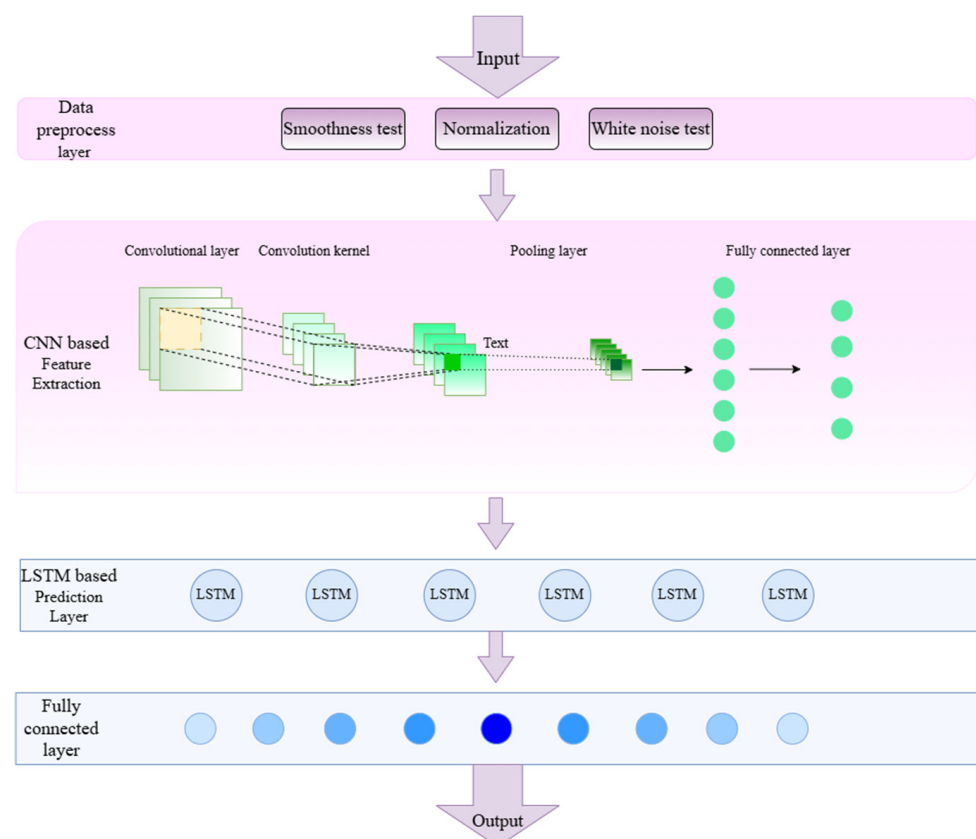


Figure 6. Structural diagram of the CNN-LSTM model.

2.3.5. LSTM-Transformer Hybrid Model

The combination of LSTM and Transformer enables complementary extraction of global and local features. The LSTM layer processes the input sequence sequentially according to time steps, updating the hidden state at each step to capture short-term dependencies. The Transformer layer explores the global correlations among different time steps in the sequence through the self-attention mechanism [28].

The core calculation formula of Transformer is:

$$\text{Attention}(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V \quad (7)$$

Here, Q , K , and V represent the query, key, and value vectors, respectively, and d_k is the feature dimension. This mechanism extracts the global features of the time series by calculating correlation weights across different time steps, thereby achieving weighted aggregation of information.

In the multi-head attention mechanism, the input is mapped into multiple subspaces, each head independently performs attention calculations, and the results are finally concatenated:

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head1}, \text{head2}, \dots, \text{headh}) \text{WO} \quad (8)$$

Multi-head attention mechanisms can capture multi-dimensional features of the input sequence in parallel across different subspaces, enhancing the model's global modeling capability [24]. Simultaneously, this establishes a complementary mechanism of “global long-term dependencies $\rightarrow$ local short-term features,” providing more comprehensive data support for complex time series prediction [29]. The structural diagram of the LSTM-Transformer hybrid model is illustrated in Figure 7.

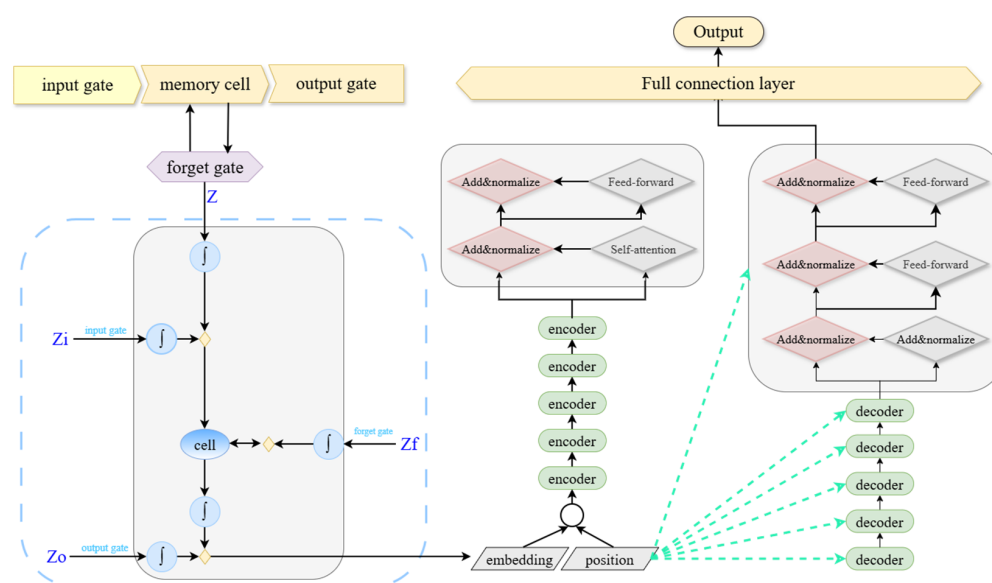


Figure 7. Structural diagram of the LSTM-Transformer hybrid model.

2.4. Model Evaluation Methods

MAE (Mean Absolute Error) is a fundamental metric for evaluating the predictive accuracy of regression models. It directly calculates the average of the absolute differences between predicted and true values, making it more robust to outliers. A smaller value of this metric indicates more precise model predictions.

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (9)$$

RMSE (Root Mean Square Error) is a commonly used metric for evaluating prediction errors in regression models. It is calculated by taking the square root of the average of the squared differences between predicted and true values, reflecting the average deviation of

predictions from the actual values. A lower value of this metric indicates higher prediction accuracy of the model.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (10)$$

R^2 (coefficient of determination) is a core metric for assessing the performance of regression models. It reflects the proportion of variance in the target variable explained by the model. Its value ranges from 0 to 1, with higher values indicating better model fit.

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (11)$$

3. Results

3.1. Analysis of Model Prediction Results

The daily observed mine water inflow data from the Maoping mining area in 2023 were chronologically divided into training and validation sets, with the first 80% of the time series used for training and the remaining 20% for validation. Nine key hydrological and meteorological variables—specifically flow velocity, water level, discharge, precipitation, average relative humidity, average wind speed, average surface air temperature, average station pressure, and sunshine duration—were used as input features, with mine water inflow defined as the target variable. The fitting and predictive performance of five deep learning models—CNN, LSTM, Transformer, CNN-LSTM, and LSTM-Transformer—was systematically evaluated. Figures 8 and 9 present the fitting curves and scatter plots of the predicted versus observed values, respectively. Overall, while all five models successfully captured the general temporal variation pattern of mine water inflow, substantial differences were observed in their fitting accuracy, local adaptability, and ability to reproduce short-term fluctuations.

The CNN model, structured with four convolutional layers and a ReLU activation function (Table 1), effectively captured the overall downward trend in mine water inflow, demonstrating a basic ability to extract temporal patterns from the input variables. However, during the middle period of the validation interval, particularly from mid-November to early December 2023, noticeable discrepancies emerged between the predicted and observed values, with the model generally overestimating mine water inflow. Quantitatively, the CNN model achieved an MAE of 0.250, an RMSE of 0.329, and an R^2 of 0.782 on the validation set, indicating relatively strong predictive performance among the three individual deep learning models.

Table 1. Main structures and parameters of the five deep learning models.

Model	Conv Layers	LSTM Layers	Transformer Layers	Attention Heads	Activation Function	Dropout
CNN	4	/	/	/	ReLU	0.1
LSTM	/	2	/	/	No explicit activation	/
Transformer	/	/	4 Encoder layers	8	Default internal	0.1
CNN-LSTM	1	1	/	2	ReLU	0.2
LSTM-Transformer	/	1	3Encoder layers	8	Tanh,ReLU	0.1

The fitting performance between the predicted inflow values from the five deep learning models and the observed values is shown in Figure 10, which provides a visual comparison of the model predictions against the actual measurements.

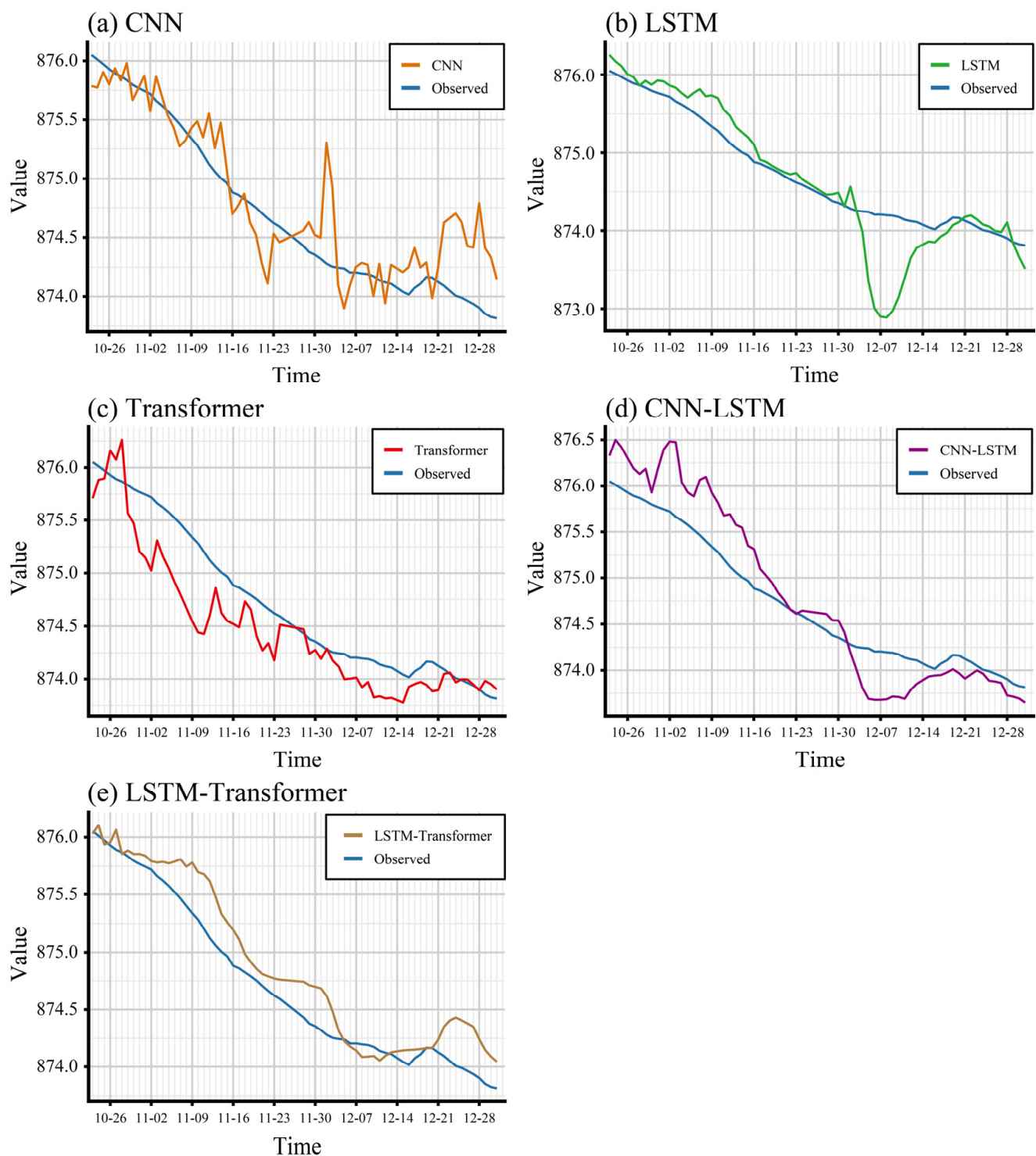


Figure 8. Fitting plot of prediction results for the five models.

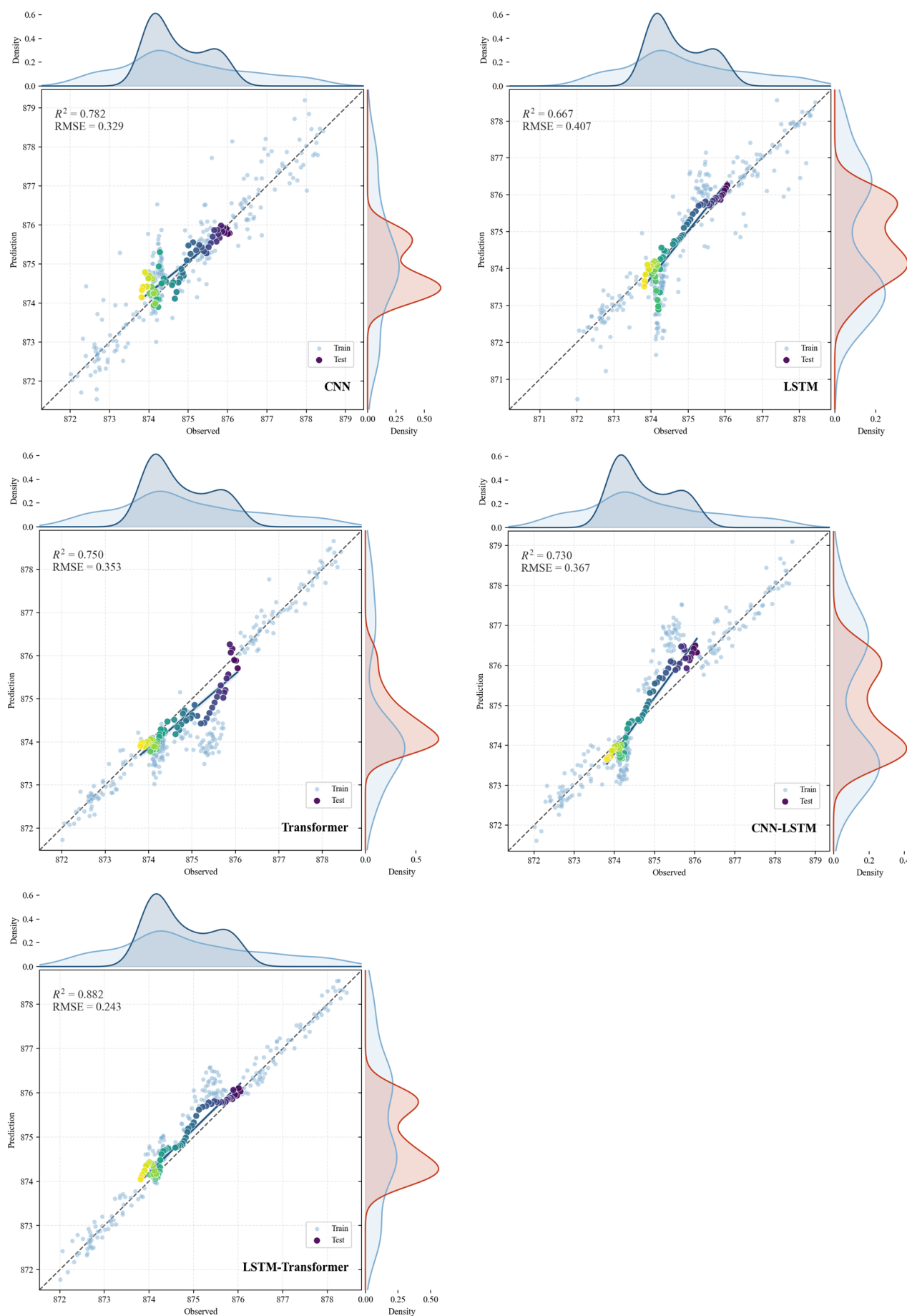


Figure 9. Scatter plot of prediction results for the five models.

The LSTM model, configured with two stacked LSTM layers, generally reproduced the gradual downward trend of the observed mine water inflow, and its predictions were relatively consistent with the observations during the first half of the validation period. However, during the middle period, the prediction curve exhibited an anomalous sharp decline followed by a rapid increase, forming a pronounced “deep-V” pattern. This behavior suggests that the model may have been overly sensitive to short-term variations in the input sequence, resulting in spurious local fluctuations and reduced prediction stability. The LSTM model yielded an MAE of 0.256, an RMSE of 0.408, and an R^2 of 0.667 on the validation set, representing the lowest goodness of fit among the five models evaluated.

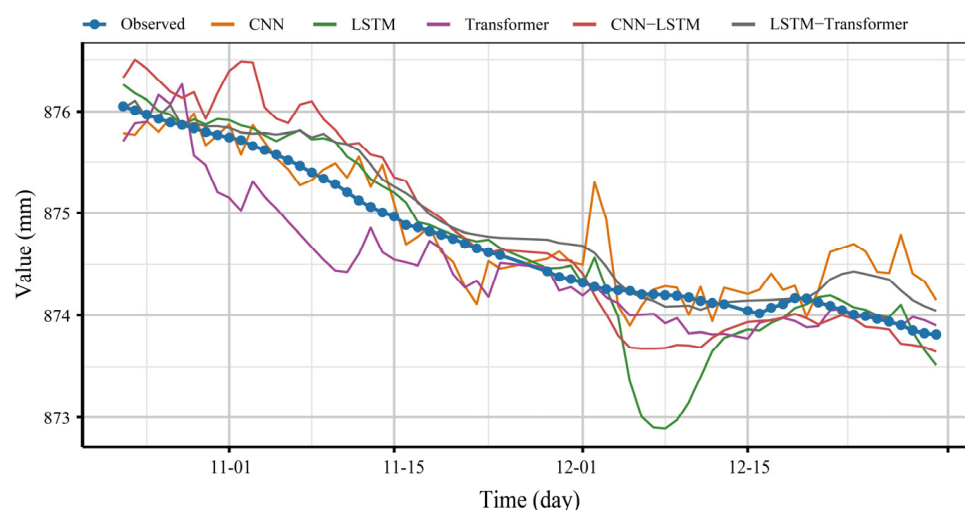


Figure 10. Fitting plot of prediction results of the five models over time.

The Transformer model employed a deeper architecture consisting of four encoder layers and eight attention heads, with a dropout rate of 0.1 to prevent overfitting. This model successfully captured the overall temporal variation pattern of mine water inflow and demonstrated competitive predictive performance. Nevertheless, during several short periods, such as early November and mid-to-late December 2023, the predicted values were lower than the corresponding observations. This indicates that despite its global attention mechanism, the model may not have fully captured certain abrupt local variations in mine water inflow. The Transformer model achieved an MAE of 0.280, an RMSE of 0.353, and an R^2 of 0.750 on the validation set, demonstrating a moderate predictive capability but a lower overall accuracy than the CNN model.

The CNN-LSTM model combined one convolutional layer with one LSTM layer and utilized a higher dropout rate of 0.2. Despite this regularization, the model exhibited noticeable local fluctuations in its prediction curve, indicating a relatively high sensitivity to variations in the input features during specific periods. Its predictive performance was characterized by an MAE of 0.305, an RMSE of 0.367, and an R^2 of 0.730 on the validation set. Although the hybrid architecture combines the local feature extraction capability of CNN with the sequential modeling ability of LSTM, it did not outperform the better-performing individual models in this experiment. This finding suggests that simply integrating multiple neural network architectures does not necessarily lead to improved prediction accuracy, and that the effectiveness of a hybrid model depends strongly on the complementarity of the underlying feature extraction and temporal modeling mechanisms.

The LSTM-Transformer model exhibited the strongest performance, leveraging a hybrid architecture comprising one LSTM layer and three Transformer encoder layers with eight attention heads. This model integrates the temporal memory capability of LSTM with the global dependency modeling ability of the Transformer self-attention mechanism.

Specifically, the LSTM layer sequentially encodes the input data to capture temporal dependencies, while the Transformer layer employs multi-head self-attention to identify long-range correlations among different time steps. The extracted features are subsequently integrated to generate the final prediction. Among the five models, the LSTM-Transformer model exhibited the strongest ability to reproduce the temporal dynamics of mine water inflow. Its predicted curve closely followed the trajectory of the observed values, with only minor deviations at several peak points. The model achieved an MAE of 0.196, an RMSE of 0.242, and an R^2 of 0.882 on the validation set, representing the best overall predictive performance among all five models. These results demonstrate that the complementary integration of LSTM-based temporal memory and Transformer-based global attention mechanisms can effectively improve the model's ability to capture complex temporal dependencies and enhance the accuracy and stability of mine water inflow prediction.

The superior performance of the LSTM-Transformer model is primarily attributed to the complementary strengths of its hybrid architecture. By leveraging the LSTM layer to effectively capture short-term temporal dependencies and integrating the Transformer's multi-head self-attention mechanism to identify long-range global correlations, the model combines the temporal memory capability of LSTM with the global dependency modeling ability of the Transformer. This integration significantly enhances the model's capacity to capture complex dynamic features, ultimately resulting in the highest prediction accuracy and stability.

3.2. Comparison of Model Prediction Results

The Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and coefficient of determination (R^2) were selected as evaluation metrics to assess the predictive performance of the five models, namely CNN, Transformer, LSTM, CNN-LSTM, and LSTM-Transformer. The results are presented in Figure 11. The evaluation results reveal substantial differences in predictive performance among the five models. Overall, the LSTM-Transformer hybrid model achieved the best performance, with an MAE of 0.196, an RMSE of 0.243, and an R^2 of 0.882, indicating the highest prediction accuracy and strongest goodness of fit among the evaluated models.

Among the remaining models, CNN achieved the second-lowest RMSE and the second-highest R^2 , with an MAE of 0.250, an RMSE of 0.329, and an R^2 of 0.782, demonstrating relatively strong predictive performance. The Transformer model also exhibited competitive performance, achieving an MAE of 0.280, an RMSE of 0.353, and an R^2 of 0.750. In contrast, the CNN-LSTM model did not consistently outperform the stronger individual models, with an MAE of 0.305, an RMSE of 0.367, and an R^2 of 0.730. The LSTM model exhibited the weakest overall performance, with an MAE of 0.256, an RMSE of 0.408, and an R^2 of 0.667, indicating the lowest goodness of fit among the five models.

These results demonstrate that hybrid modeling does not necessarily guarantee improved predictive performance, and its effectiveness largely depends on the complementarity of the integrated components. The superior performance of the LSTM-Transformer model can be attributed to the effective integration of LSTM's capability to capture sequential temporal dependencies with the Transformer's self-attention mechanism for modeling long-range and global dependencies. This complementary architecture enables the model to extract more comprehensive temporal information from complex mine water inflow series, thereby improving prediction accuracy and overall model performance.

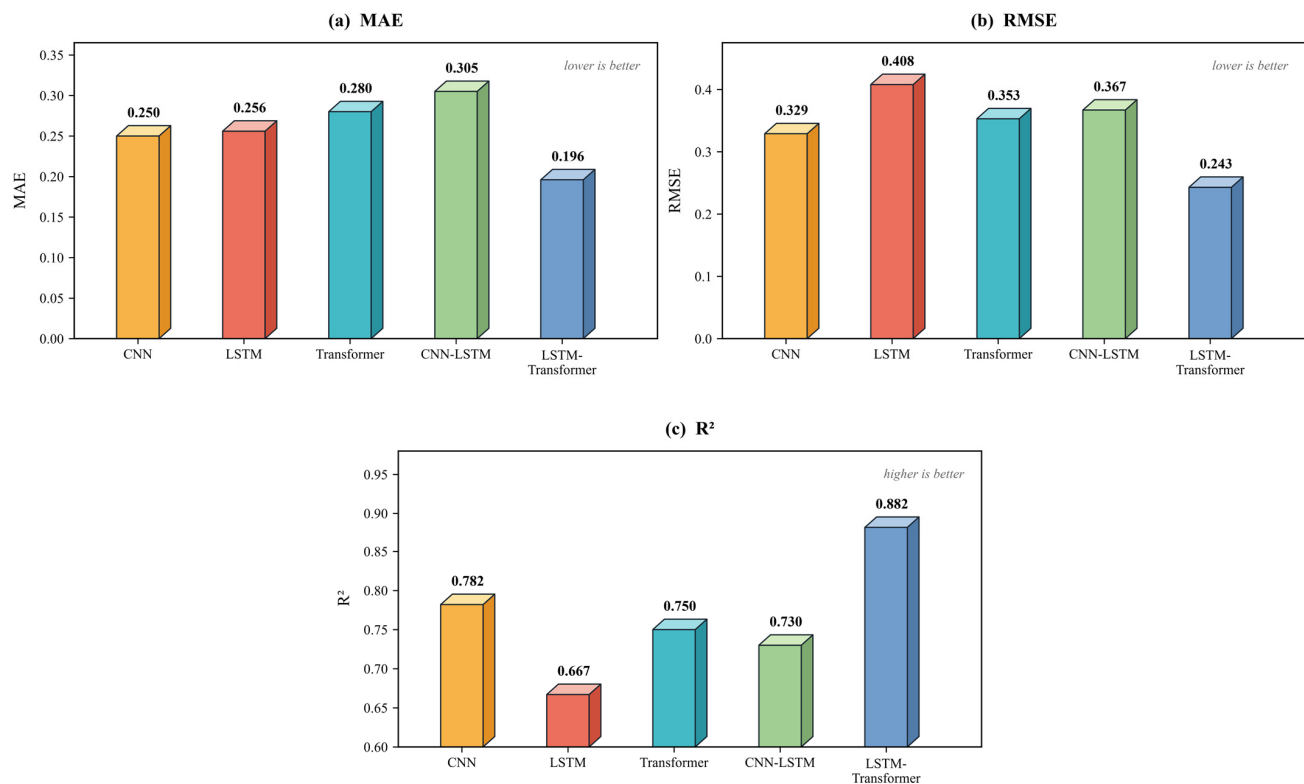


Figure 11. Prediction results of the five deep learning models.

4. Discussion

The results demonstrate the effectiveness of hybrid architectures that integrate local feature extraction (CNN/LSTM) with global dependency modeling (Transformer). This architecture is well suited to mine water inflow prediction, which is jointly influenced by localized geological and hydrological processes and broader meteorological conditions, providing an effective pathway for improving prediction accuracy. The gated memory mechanism of LSTM is suitable for capturing the lagged responses and cumulative effects of mine hydrological processes, while the self-attention mechanism of the Transformer can model complex short- and long-range dependencies among multiple water sources and meteorological-hydrological variables. Their complementary capabilities contribute to the superior performance of the LSTM-Transformer model. The nine meteorological and hydrological variables used in this study enabled satisfactory predictive performance; however, incorporating additional geological and structural parameters, such as dynamic aquifer properties and fault-fracture characteristics, as well as advanced feature engineering methods such as multi-scale decomposition, may further improve model stability and generalization under complex mining conditions. The high-precision predictive capability of the LSTM-Transformer model provides reliable technical support for dynamic mine water hazard early warning. Future research could develop physics-informed deep learning models by incorporating groundwater flow and seepage constraints into the LSTM-Transformer framework through physical loss functions or soft constraints. This approach could combine the nonlinear fitting capability of deep learning with physical consistency, thereby improving extrapolation accuracy and prediction reliability, particularly under conditions of limited, missing, or noisy observations.

5. Conclusions

This study constructed various deep learning models for the daily prediction of mine water inflow in the Maoping mining area and systematically evaluated the performance of CNN, LSTM, Transformer, and their hybrid models. The main conclusions are as follows:

(1) Superiority of the LSTM-Transformer Architecture

Among the evaluated models, the LSTM-Transformer hybrid model demonstrated the most robust predictive performance ($MAE = 0.196$, $RMSE = 0.243$, $R^2 = 0.882$), significantly outperforming single models (such as the baseline LSTM with $R^2 = 0.667$) and other hybrid configurations. This superiority is attributed to the model's ability to synergize distinct temporal features: the LSTM layer effectively captures short-term dependencies and lagged responses to continuous meteorological events, while the Transformer's self-attention mechanism successfully models long-period, cross-seasonal global trends. This complementary architecture resolves the instability often observed in single models when handling the non-linear and non-stationary characteristics of hydrological time series.

(2) Practical Implications for Mine Safety

From an application perspective, the high precision and stability of the LSTM-Transformer model offer significant value for mining safety management. Unlike models that merely fit historical trends, this hybrid approach provides reliable early warnings for abnormal water inflow, allowing for proactive disaster prevention. Its ability to minimize local prediction errors ensures that decision-makers can formulate more accurate drainage schedules, thereby optimizing energy consumption and reducing the risk of water inrush accidents in complex mining environments.

(3) Limitations and Future Directions

Despite its promising performance, this study has certain limitations. The model was trained primarily on historical inflow and meteorological data, potentially overlooking other hydrogeological factors such as real-time groundwater levels or geological structural changes. Furthermore, the computational complexity of the Transformer module may pose challenges for deployment on low-power edge devices in underground mines. Future research should focus on:

- (1) **Multi-source Data Fusion:** Integrating hydrogeological monitoring data and InSAR surface deformation data to enhance the model's physical interpretability.
- (2) **Model Lightweighting:** Exploring model compression techniques to facilitate real-time deployment on embedded mining systems.
- (3) **Generalization Verification:** Validating the model's transferability across different mining areas with varying hydrogeological conditions to establish a more universal prediction framework.

Author Contributions: Conceptualization, F.W.; methodology, L.C., Z.L. and X.H.; software, K.W.; validation, L.C.; formal analysis, L.C. and Z.L.; investigation, L.C. and Z.L.; resources, F.W.; data curation, X.H.; writing—original draft preparation, Z.L.; writing—review and editing, L.C.; visualization, X.H.; supervision, K.W.; project administration, F.W. L.C. and Z.L. contributed equally to this work and share first authorship. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Science Foundation of China, No. 42272308, 61703306.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding authors.

Conflicts of Interest: Author Lujun Chai and Kun Wu were employed by the company North China Engineering Investigation Institute Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Dong, W.W.; Jie, L.S.; Xi, H.J. Analysis of the main global coal resource countries' supply-demand structural trend and coal industry outlook. *Min. Mag.* **2015**, *24*, 5–9.
2. Li, B.; Wu, H.; Liu, P.; Fan, J.; Li, T. Construction and application of mine water inflow prediction model based on multi-factor weighted regression: Wulunshan Coal Mine case. *Earth Sci. Inform.* **2023**, *16*, 1879–1890. [\[CrossRef\]](#)
3. Qu, X.; Shi, L.; Han, J. Preventing water-inrush from floor in coal working face with paste-like backfill technology. *Sci. Rep.* **2023**, *13*, 15947. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Liu, H.; Liu, G.; Ning, D.; Fan, J. Mine water inflow prediction method based on VMD-DBN model. *Coal Geol. Explor.* **2023**, *51*, 13–21.
5. Xie, J.K.; Xu, Y.P.; Gao, C.; Xuan, W.; Bai, Z. Total Basin Discharge from GRACE and Water Balance Method for the Yarlung Tsangpo River Basin, Southwestern China. *J. Geophys. Res. Atmos.* **2019**, *124*, 7617–7632.
6. Li, B.; Li, T.; Zhang, W.; Liu, Z.; Yang, L. Multisource information risk evaluation technology of mine water inflow based on VWM: A case study of Weng'an coal mine. *Geofluids* **2021**, *2021*, 8812144. [\[CrossRef\]](#)
7. Zhang, X.; Guo, Y.H.; Chen, Y.K. Forecasting of mine discharge based on phase space reconstruction and SimpleMKL. *Chem. Eng. Trans.* **2016**, *51*, 1027–1033.
8. Yu, X.G.; Han, J.; Shi, L.Q.; Wang, Y.; Zhao, Y. Application of a BP neural network in predicting destroyed floor depth caused by underground pressure. *Environ. Earth Sci.* **2017**, *76*, 535.
9. Meenal, R.; Selvakumar, A.I. Assessment of SVM, empirical and ANN based solar radiation prediction models with most influencing input parameters. *Renew. Energy* **2018**, *121*, 324–343. [\[CrossRef\]](#)
10. Meenal, R.; Selvakumar, A.I. Solar mapping of India using support vector machine. *J. Phys. Conf. Ser.* **2018**, *1142*, 012010. [\[CrossRef\]](#)
11. Duan, H.S.; Meng, X.; Tang, J.; Qiao, J.-F. Time-series prediction using a regularized self-organizing long short-term memory neural network. *Appl. Soft Comput.* **2023**, *145*, 110553. [\[CrossRef\]](#)
12. Salloom, T.; Kaynak, O.; He, W. A novel deep neural network architecture for real-time water demand forecasting. *J. Hydrol.* **2021**, *599*, 126353. [\[CrossRef\]](#)
13. Jabeen, G.; Rahim, S.; Afzal, W.; Khan, D.; Khan, A.A.; Hussain, Z.; Bibi, T. Machine learning techniques for software vulnerability prediction: A comparative study. *Appl. Intell.* **2022**, *52*, 17614–17635. [\[CrossRef\]](#)
14. Wang, N.; Zhao, X. Time series forecasting based on convolution transformer. *IEICE Trans. Inf. Syst.* **2023**, *E106.D*, 976–985. [\[CrossRef\]](#)
15. Yu, Y.; Si, X.S.; Hu, C.H.; Zhang, J. A review of recurrent neural networks: LSTM cells and network architectures. *Neural Comput.* **2019**, *31*, 1235–1270. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Yin, H.; Zhu, W.; Zhang, X.; Xing, Y.; Xia, R.; Liu, J.; Zhang, Y. Runoff predictions in new-gauged basins using two transformer-based models. *J. Hydrol.* **2023**, *622*, 129684. [\[CrossRef\]](#)
17. Liu, C.; Liu, D.; Mu, L. Improved transformer model for enhanced monthly streamflow predictions of the Yangtze River. *IEEE Access* **2022**, *10*, 58240–58253. [\[CrossRef\]](#)
18. Di Nunno, F.; De Marinis, G.; Granata, F. Short-term forecasts of streamflow in the UK based on a novel hybrid artificial intelligence algorithm. *Sci. Rep.* **2023**, *13*, 7036. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Xu, Y.; Hu, C.; Wu, Q.; Jian, S.; Li, Z.; Chen, Y.; Zhang, G.; Zhang, Z.; Wang, S. Research on particle swarm optimization in LSTM neural networks for rainfall-runoff simulation. *J. Hydrol.* **2022**, *608*, 127553. [\[CrossRef\]](#)
20. Choubin, B.; Malekian, A. Combined gamma and M-test-based ANN and ARIMA models for groundwater fluctuation forecasting in semiarid regions. *Environ. Earth Sci.* **2017**, *76*, 538. [\[CrossRef\]](#)
21. Feng, J.; Wang, Z.; Wu, Y.; Xi, Y. Spatial and temporal aware graph convolutional network for flood forecasting. In Proceedings of the 2021 International Joint Conference on Neural Networks (IJCNN), New York, NY, USA, 18–22 July 2021.
22. Lü, X.; Nurmamet, I.; Xiao, S.; Zhao, J.; Yu, X.; Aili, Y.; Li, S. Spatial-temporal simulation and prediction of root zone soil moisture based on Hydrus-1D and CNN-LSTM-attention models in Yutian Oasis, southern Xinjiang, China. *Pedosphere* **2025**, *35*, 846–857. [\[CrossRef\]](#)
23. Ding, Y.Y.; Yin, S.X.; Lian, H.Q.; Liu, W.; Li, Q. Multi-factor water inflow prediction in coal mining face based on SSA-CG-Attention model. *Coal Geol. Explor.* **2024**, *52*, 111–119.
24. Li, Z.H.; Jiang, Y.F.; Du, F.; Wang, W. Prediction of mine water inflow under water inrush condition based on LSTM-Transformer model. *J. Henan Polytech. Univ. (Nat. Sci.)* **2026**, *45*, 77–85.

25. Huang, D.; Liu, X.; Ma, A.G.; Wu, Y.; Cai, J.; Wang, L. Daily runoff forecasting for cascade reservoirs based on Transformer-GRU model. *China Rural. Water Hydropower* **2025**, *7*, 8–16.
26. Landi, F.; Baraldi, L.; Cornia, M.; Cucchiara, R. Working Memory Connections for LSTM. *Neural Netw.* **2021**, *144*, 334–341. [[CrossRef](#)] [[PubMed](#)]
27. Liu, K.Z.; Gou, J.Q.; Luo, Z.; Wang, K.; Xu, X.; Zhao, Y. Prediction method for dissolved gas concentration in transformer oil based on particle swarm optimization-long short-term memory network model. *Power Syst. Technol.* **2020**, *44*, 2778–2785.
28. Shi, J.W.; Wang, S.; Qu, P.; Shao, J. Time series prediction model using LSTM-Transformer neural network for mine water inflow. *Sci. Rep.* **2024**, *14*, 1828. [[CrossRef](#)] [[PubMed](#)]
29. Li, Z.X. Integrated learning method for ground pressure prediction based on variable time-shift Transformer-LSTM. *Ind. Mine Autom.* **2023**, *49*, 92–98.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.