Short-Term Mine Water Inflow Forecasting at Sandao Coal Mine Using an Adaptive Weighted Ensemble Model

Shaokai Hu[1], Xuelei Zhang[2]

[1] North China Institute of Science and Technology, Hebei State Key Laboratory of Mine Disaster Prevention, Beijing 101601, China

[2] Shaanxi Changwu Tingnan Coal Industry Co., Ltd., Xianyang, Shaanxi 712000, China

Abstract

The water inflow in mines exhibits dynamic characteristics influenced by multiple coupled factors, and a single predictive model often cannot meet the requirements for precise forecasting. This study proposes an adaptive weighted ensemble for rolling forecasts at 10-day intervals. A continuous series of 117 observations from January 2022 to March 2025 was extracted from the water inflow ledgers of Sandao Coal Mine. Persistence, a three-point damped trend model, and a seasonal-naive model were used as complementary base learners. Their weights were updated from the absolute errors over the latest 18 periods. Data from 2022-2023 were used for model construction and parameter selection, 2024 was retained as an independent test set, and the first quarter of 2025 served as an external validation set. On the 2024 test set, the ensemble achieved an MAE of 10.563 m^3/h, an RMSE of 15.881 m^3/h, a MAPE of 10.690%, and an R^2 of 0.653. Relative to persistence, MAE, RMSE, and MAPE decreased by 4.89%, 9.45%, and 6.40%, respectively. The external-validation MAPE was 9.386%. Forecasts for the first, middle, and last 10-day periods of April 2025 were 73.4, 74.9, and 76.0 m^3/h, respectively. The proposed method is computationally light, traceable, and suitable for short-term mine water inflow forecasting when monitoring samples are limited.

Keywords

Mine Water Inflow; 10-day Interval Forecasting; Adaptive Weighting; Rolling Forecast; Small Sample.

1. Introduction

Mine water inflow is a fundamental indicator for drainage-system operation, water-hazard prevention works, and risk assessment. Mining disturbance changes the development of water-conducting fractures, the connectivity of goafs, and aquifer recharge conditions. Consequently, inflow may exhibit gradual evolution, periodic variation, and abrupt stage transitions. Forecasts based only on empirical values or fixed-parameter models can develop persistent bias during working-face transitions, concentrated drainage, or changes in recharge. Improving short-term adaptability without sacrificing interpretability is therefore directly relevant to mine operations[1-2].

Existing approaches to mine water inflow forecasting can be grouped into hydrogeological analysis and numerical simulation, statistical time-series models, and machine learning. Analogy, analytical, and numerical methods can represent aquifer structures and boundary conditions, but require complete parameters and defensible conceptualization[3]. LSTM, GRU, and attention-based models can extract nonlinear temporal features and perform well when sufficiently rich multi-factor data are available [4-6]. However, when data are limited, the predictive results of these models are easily affected by training randomness and distributional changes. Persistence models, moving average

ISSN: 2414-1895 DOI: 10.6919/ICJE.202609_12(9).0010

models, and seasonal naive models have relatively stable structures; yet, they struggle to simultaneously track peaks and preserve cyclical information. Therefore, appropriate weights can be assigned to combine different models, thus exploiting the complementarity in their errors. [7-9].

2. Mine Overview and Dataset

2.1 Mine Overview

Sandao Gou Coal Mine is located in the northern part of the Loess Plateau in northern Shaanxi. The terrain is highly undulating, with numerous ravines, representing typical Loess Plateau landforms. Vegetation in the area is sparse, and soil erosion is severe. The bedrock and red soil are extensively exposed along both sides of the valleys, while the tops of some ravines and ridges are covered with Quaternary loess or eolian sand. Overall, the terrain slopes from west to east, with a maximum elevation difference of 252.4 meters. The strata, from oldest to youngest, are as follows: Wayaobu Formation of the Upper Triassic (T3w), Fuxian Formation of the Lower Jurassic (J1f), Yan'an Formation of the Middle to Lower Jurassic (J1-2y), Jingle Formation of the Upper Neogene (N2j), Lishi Formation of the Middle Pleistocene (Q2l), and Holocene alluvial deposits (Q41al pl, Q42al pl) as well as eolian deposits (Q42eol) of the Quaternary. The currently operating working face is 85203, with the main coal seam being 5-2. The historical average water inflow is 76 m^3/h, with a maximum inflow of 101 m^3/h.

2.2 Dataset

Taking the Sandao Gully Coal Mine as the research subject, this study systematically integrates multi-source hydrological, meteorological, geological, and mine monitoring data. The data used are primarily categorized into two types: static data and dynamic data. Static data include a 1:10,000 hydrological and geological map of the study area, a coal stratigraphic column chart, a report on hydrological and geological type classification, and drilling data. Dynamic data mainly consist of (1) river water level and flow monitoring data from 15 flow monitoring stations within the mine area, used to characterize the overall hydrological dynamics of the watershed; (2) a meteorological station, established to collect rainfall data at one-hour intervals; and (3) monthly mine water inflow records from 2022 to 2025, including the total discharge of the main drainage system, with domestic water and industrial recycled water interferences excluded.

The study adopted a one-step rolling forecast. When the model predicts for period t, it can only use observations prior to period t. After the observation at t arrived, historical errors and the weight for the next period were updated. The 72 observations in 2022-2023 supported model construction, with the 18 observations from July to December 2023 used for hyperparameter selection. The 36 observations in 2024 formed the independent test set, and nine observations from January to March 2025 formed the external validation set. This chronological design prevents future information from leaking into training.

Table 1. Chronological division of the dataset

Stage	Period	Samples	Purpose
Modeling	2022-01 to 2023-06	54	Base-model fitting and rolling updates
Parameter selection	2023-07 to 2023-12	18	Select error window and weight exponent
Independent test	2024-01 to 2024-12	36	Compare out-of-sample performance
External validation	2025-01 to 2025-03	9	Assess the new low-inflow state

ISSN: 2414-1895 DOI: 10.6919/ICJE.202609_12(9).0010

Table 2. Annual descriptive statistics of mine water inflow

Year	Samples	Mean (m^3/h)	SD (m^3/h)	Minimum (m^3/h)	Maximum (m^3/h)
2022	36	88.75	8.85	68.0	106.0
2023	36	93.74	12.62	76.0	126.0
2024	36	98.44	26.98	71.0	170.0
2025	9	78.78	3.99	72.0	84.0

A consistency audit showed that the total inflow for every record in the continuous 2022-2025 interval equaled the sum of the whole-mine clean-water and wastewater entries. Stored water volume was recorded separately and was not counted as instantaneous inflow. The arithmetic relationships among clean water, wastewater, and total inflow were also consistent in the high-value worksheets from September to November 2024; the peak of 170 m^3/h was therefore not a spreadsheet-formula or cell-reading error. Because reused water was not recorded independently in every version of the ledger, the analysis follows the published ledger definition: whole-mine clean water plus wastewater.

A further audit of the 2024 high-inflow stage confirmed the following clean-water, wastewater, and total values: 53, 67, and 120 m^3/h on September 20; 48, 88, and 136 m^3/h on October 20; 43, 106, and 149 m^3/h on October 30; 34, 136, and 170 m^3/h on November 10; 26, 130, and 156 m^3/h on November 20; and 26, 120, and 146 m^3/h on November 30. No summation inconsistency was found.

As shown in Table 2, mean annual inflow increased from 88.75 m^3/h in 2022 to 98.44 m^3/h in 2024. The standard deviation reached 26.98 m^3/h in 2024, substantially exceeding those of the preceding two years. In the first quarter of 2025, the mean declined to 78.78 m^3/h and the standard deviation to 3.99 m^3/h, indicating a new, relatively stable low-inflow state. The complete series is shown in Fig. 1.

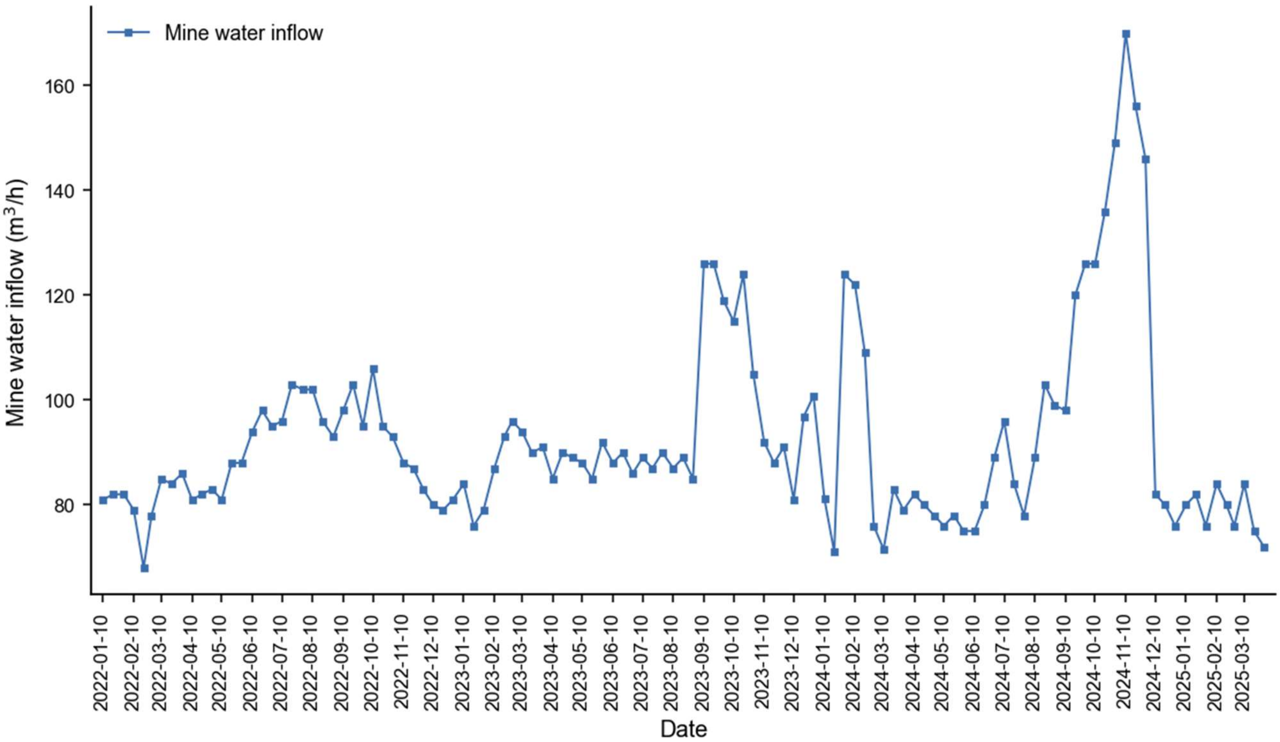

Fig. 1 Variation of 10-day mine water inflow from 2022 to 2025

ISSN: 2414-1895 DOI: 10.6919/ICJE.202609_12(9).0010

3. Adaptive Weighted Ensemble Forecasting Method

3.1 Base Learners

This study uses the persistence model, the three-period damped trend model, and the seasonal naive model as the baseline models. The persistence model focuses on the most recent level, the damped trend model considers short-term directional changes, and the seasonal naive model relies on information from the same period in the previous year. This combination does not require any additional external data. Each baseline model performs one-step-ahead forecasting in the same chronological order, using only the data available prior to period t when forecasting period t. This approach ensures a fair comparison among models and prevents future information from entering the forecasting process.

The persistence model uses the most recent measured value as the forecast for the next period. The basic idea of this model is that the drainage conditions and overall water inflow levels between two adjacent periods usually have some continuity. When the sequence is in a stable phase, this simple approach can often give fairly consistent results. It doesn't require parameter estimation, nor does it show obvious fitting fluctuations with a small sample size. The calculation formula for the persistence model[10] is:

$$\hat{y}_{P,t} = y_{t-1} \tag{1}$$

Here, y represents the amount of water inflow, t represents the t-th ten-day period, and P represents Persistence model.

The persistent model only retains information from the most recent period. When the inflow rate continuously rises, continuously falls, or suddenly changes direction, its forecast will experience a lag of about one decadal period. Therefore, it is used only as a component of the combined model to reflect short-term inertia.

The three-period damped trend model fits a least-squares line using observations from the most recent three decadal periods and derives the local slope b_t from the line. A positive slope indicates that the inflow rate has generally increased in the recent period, while a negative slope indicates that it has generally decreased. Using only three observations allows the model to quickly identify short-term trends; however, the short window may also be easily affected by a single anomalous fluctuation. Especially near peaks or troughs, directly extrapolating according to the slope may cause overshooting. Therefore, the local slope is multiplied by a damping factor φ, and φ=0.5. The model computation formula is:

$$\hat{y}_{T,t} = y_{t-1} + \varphi b_t \tag{2}$$

The seasonal-naive model incorporates the possible annual hydrogeological cycle and broadly repeated mining rhythm. At a 10-day resolution, one year contains 36 observations:

$$\hat{y}_{S,t} = y_{t-36} \tag{3}$$

For comparison, a three-point weighted moving average and a ridge-regression time-series model were also evaluated. Ridge inputs included lags of 1, 2, 3, 6, 9, 12, and 18 periods; moving means over 3, 6, and 9 periods; the six-period standard deviation; annual sine and cosine terms; and a time-trend term. The regularization coefficient was selected from 0.1-100 by validation RMSE, giving an optimum of 1.0.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202609_12(9).0010

The complete rolling procedure is shown in Fig. 2. Each forecast uses only observations and historical errors available before the forecast origin; test observations are not used for advance parameter fitting or weight updates.

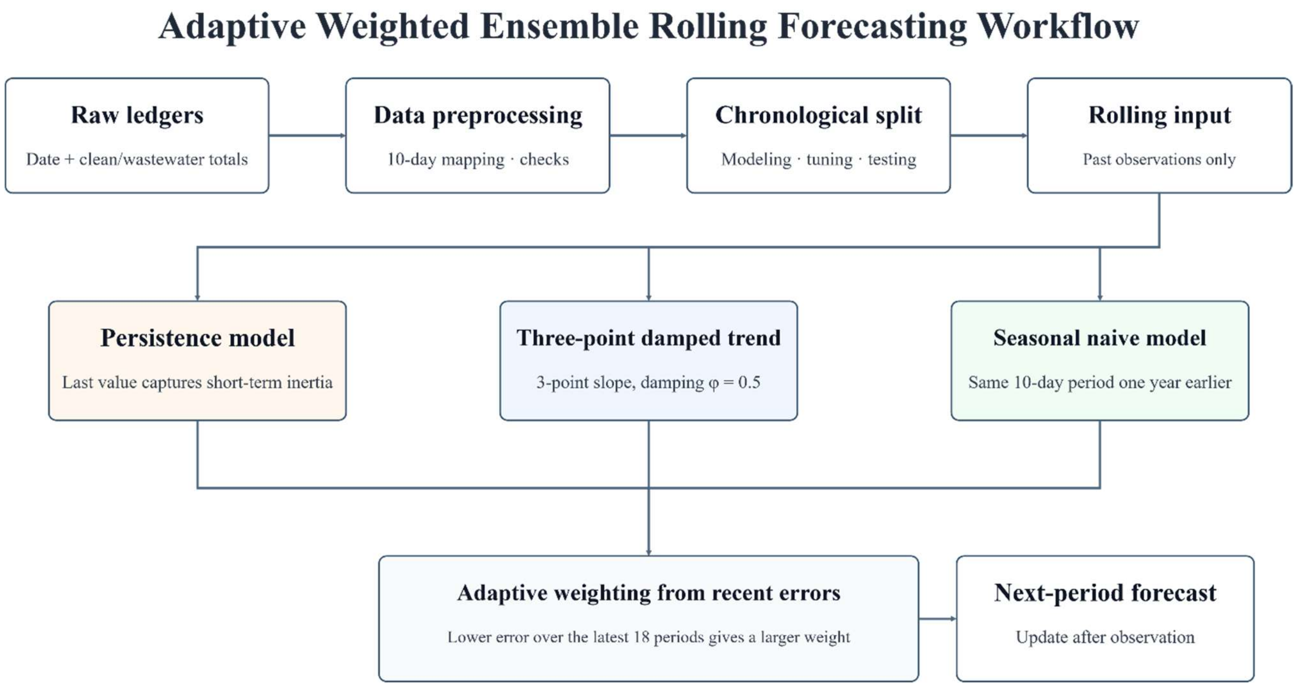

Fig. 2 Rolling forecasting procedure of the adaptive weighted ensemble model

3.2 Adaptive Weights

Fixed weights cannot readily adapt to changes caused by working-face transitions or concentrated drainage. The recent absolute error over a window of L periods is used to quantify the current reliability of each base learner. For model i immediately before period t:

$$E_{i,t} = \frac{1}{L}\sum_{k=t-L}^{t-1}\left|y_k - \hat{y}_{i,k}\right|, i \in \{P,T,S\} \tag{4}$$

A power function of inverse error then produces nonnegative normalized weights:

$$W_{i,t} = \frac{\left(E_{i,t}+\varepsilon\right)^{-p}}{\sum_{j\in\{P,T,S\}}\left(E_{j,t}+\varepsilon\right)^{-p}} \tag{5}$$

Here $\varepsilon = 0.5$ prevents excessive weight concentration when an error approaches zero, and p controls sensitivity to differences among model errors. The ensemble forecast is:

$$\hat{y}_t = w_{P,t}\hat{y}_{P,t} + w_{T,t}\hat{y}_{T,t} + w_{S,t}\hat{y}_{S,t} \tag{6}$$

During parameter selection, $L \in \{6, 9, 12, 18\}$ and $p \in \{0.5, 1.0, 1.5, 2.0\}$ were compared by grid search using RMSE. The selected values were $L = 18$ and $p = 0.5$. This combination changes the weights gradually according to approximately six months of recent performance and prevents a single-period error from causing rapid switching. Mean 2024 weights for persistence, damped trend, and seasonal naive were 0.362, 0.337, and 0.301, respectively.

Volume 12 Issue 9, 2026
DOI: 10.6919/ICJE.202609_12(9).0010

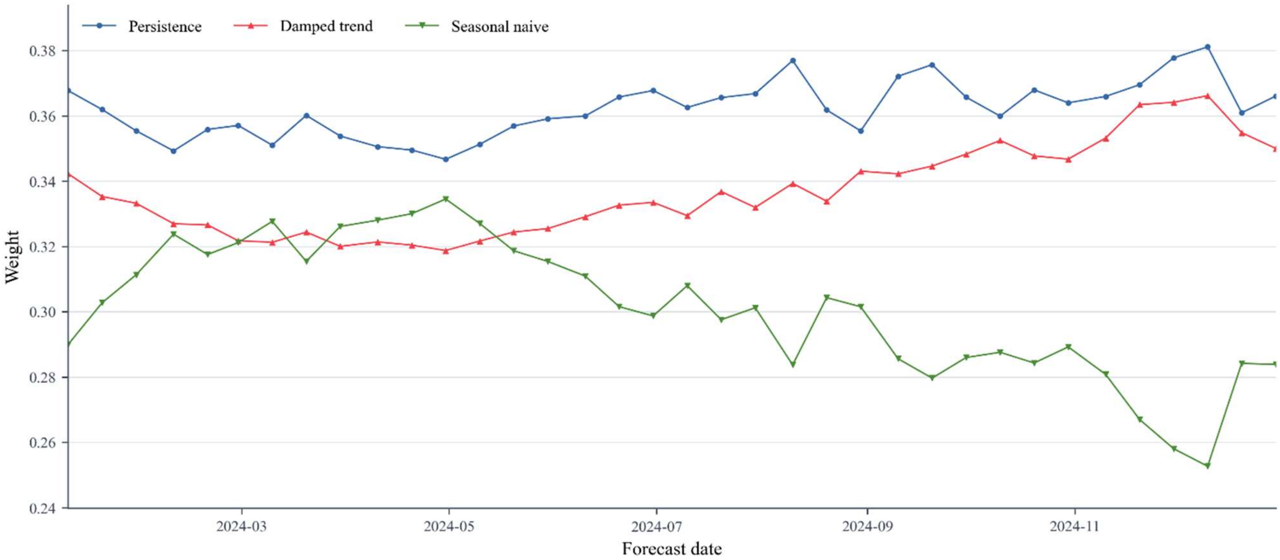

Fig. 3 Dynamic weights of the three base learners during the 2024 test period

Fig. 3 shows gradual adjustment in all three weights and no persistent dominance by a single learner. The weights of the persistence model are mostly at a relatively high level, indicating that recent observed values still have a strong reference effect for the next ten-day period. The weight adjustment of the damped trend model is intended to compensate for the lag of the persistence model. Overall, the weight of the seasonal naïve model has decreased, suggesting that the similarity between the state in 2024 and that of the same period in 2023 is weak, but information from the previous year has not been entirely discarded.

4. Results and Analysis

4.1 Parameter Determination and Test Results

The study selected parameters using 18 observations from July to December 2023 and evaluated the predictive performance of different error windows L and weight index p through grid search. Validation results indicate that when L=18 and p=0.5, the model's RMSE is minimized at 11.542 m^3/h; at this point, MAE, MAPE, and R are 7.424 m^3/h, 7.079%, and 0.415, respectively. The model uses L=18 to calculate errors over the past six months and does not significantly adjust weights immediately due to a single prediction deviation. Setting p to 0.5 allows for smoother weight changes among the three base models and prevents any single model from being assigned excessive weight due to smaller short-term errors. Once the parameters are determined, this set of fixed parameters will be used in the tests in 2024 and the external validation in 2025.

Table 3. Performance of Independent Models

Model	MAE (m^3/h)	RMSE (m^3/h)	MAPE (%)	R^2
Persistence	11.106	17.538	11.421	0.577
Three-point weighted moving average	13.969	20.479	14.784	0.424
Three-point damped trend	13.090	18.907	13.472	0.509
Seasonal naive	18.192	25.781	16.858	0.087
Ridge-regression time series	13.809	19.231	14.592	0.492
Adaptive weighted ensemble	10.563	15.881	10.690	0.653

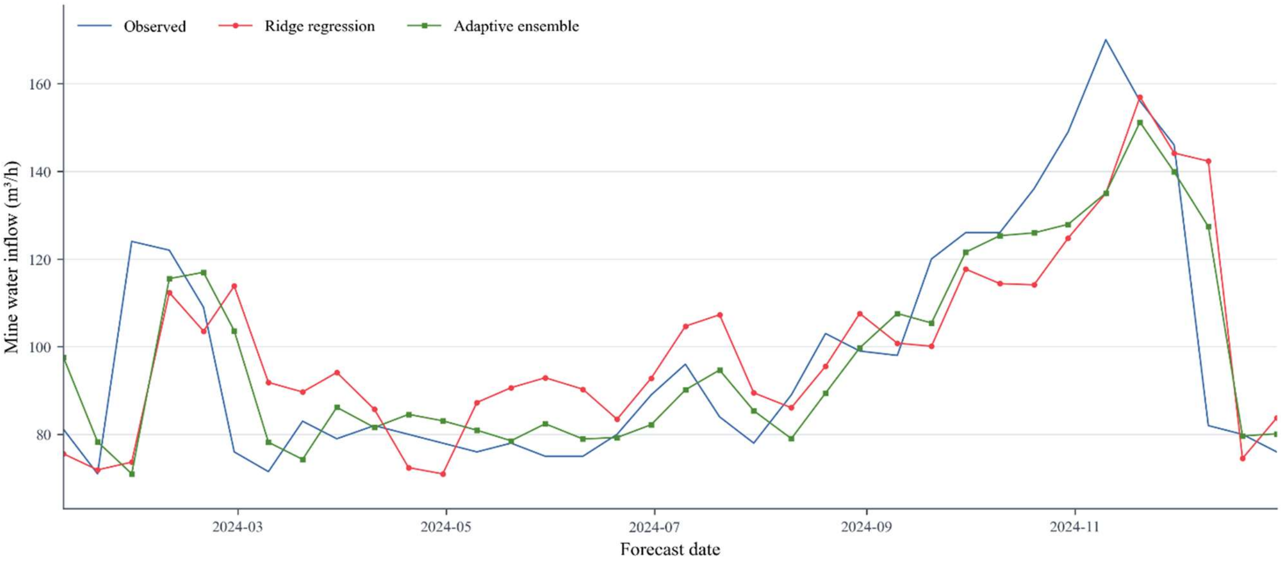

Fig. 4 Model Rolling Prediction Results

4.2 Peak Error

Fig. 4 shows the prediction results of the combined model. The model generally followed the overall upward trend from September to November but did not capture the sudden high values in a timely manner. In the test set, there were 10 observed values reaching or exceeding 120 m^3/h; the model underestimated all these high values, with an MAE of 15.629 m^3/h and a mean error of -15.629 m^3/h. For the remaining 26 observations, the model's MAE was 8.614 m^3/h, with a mean error of 5.044 m^3/h; the prediction errors of the model within the normal range of inflow were smaller. The maximum absolute error in the test set occurred in late January 2024; at that time, the observed value was 124.0 m^3/h, and the predicted value was 71.0 m^3/h, resulting in an absolute error of 53.0 m^3/h. Since there were no obvious historical signals prior to the recent increase in water inrush, models based solely on past water inrush data cannot predict this high value in advance.

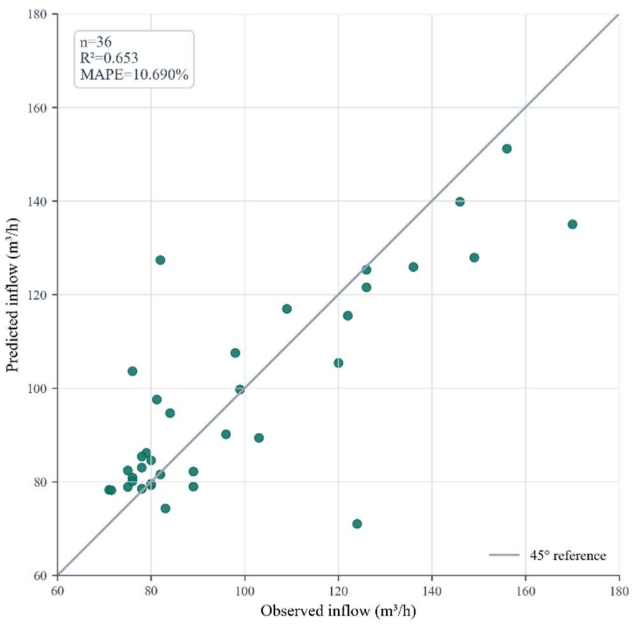

Fig. 5 Scatter Plot of Actual and Predicted Water Inflow

As shown in Fig. 5: for most points with low water inflow, the model's predicted values are relatively close to the measured values, with scatter points mainly distributed near the 1:1 line. As the measured

ISSN: 2414-1895 DOI: 10.6919/ICJE.202609_12(9).0010

water inflow increases, the model is more likely to underestimate, with more scatter points falling below the 1:1 line; the average error during high-value periods is also negative. The model's errors mainly occur during sudden increases in water inflow and peak periods, rather than being evenly distributed throughout the testing period. Therefore, MAE and MAPE can only measure the overall error of the model and cannot determine whether the model misses peaks. In practical model evaluation, RMSE, error direction, and statistical results of high-value samples should also be considered.

4.3 Discussion

The model only used the historical total mine inflow data and did not incorporate variables such as rainfall, water levels, drainage volume, mining location, and face advancement; therefore, it is unable to use external signals to predict peaks in advance. Although the additive relationship between clear water volume, sewage volume, and total inflow was verified, changes in statistical criteria could still affect the results. The existing loss functions and evaluation metrics treat underestimation of high values and overestimation of low values in the same way, making it impossible to emphasize the safety risks associated with missed peak reports.

Subsequent research needs to supplement data corresponding to dates, including rainfall, water level, mining area flow, drainage volume, working face position, and drainage events. Only after relevant data form a continuous sequence can the multi-factor model be further tested to see whether it can improve peak prediction. Subsequent validation should continue to use a time-ordered rolling approach, with the persistent model retained as a baseline; randomly partitioning the data may result in an overestimation of model accuracy. As the sample size increases, nonlinear models such as GRU or attention mechanisms can be incorporated into the existing combination framework, and it should be tested whether these models can consistently reduce errors during independent test periods. For safety-related applications, model evaluation should separately consider underestimation of high values; model outputs should also attempt to include upper quantile predictions or prediction intervals, rather than providing only a single point estimate.

5. Conclusion

A complete 117-observation series at 10-day intervals was constructed from the Sandao Coal Mine inflow ledgers for January 2022 to March 2025. Mean inflow in 2024 was 98.44 m^3/h with a standard deviation of 26.98 m^3/h, indicating substantially stronger variability than in 2022-2023. Component-level checks found no summation errors in the high ledger values. The first quarter of 2025 shifted to a low state with a mean of 78.78 m^3/h.

The proposed adaptive weighted ensemble combines persistence, a three-point damped trend, and a seasonal-naive learner using weights derived from the latest 18 absolute errors. Both the calculation path and the dynamic weights are traceable. On the independent 2024 test set, MAE, RMSE, MAPE, and R^2 were 10.563 m^3/h, 15.881 m^3/h, 10.690%, and 0.653, respectively. Relative to persistence, the first three errors decreased by 4.89%, 9.45%, and 6.40%.

The ensemble tracked routine variation reasonably well but underestimated abrupt peaks. The external-validation MAPE for the first quarter of 2025 was 9.386%, showing that rapid recalibration is still needed after a state change. Trial forecasts for the three 10-day periods of April 2025 were 73.4, 74.9, and 76.0 m^3/h. Candidate exogenous-variable classes and state-change thresholds are deployment examples only and do not enter the reported model or accuracy calculations.

References

[1] Yin, S. X., Xu, B., Yin, H. C., & et al. (2023). Basic framework and connotation of mine water prevention and control. *Coal Science and Technology*, 51(7), 24–35. (in Chinese)

[2] Zeng, Y. F., Yu, C., Wu, Q., & et al. (2024). Three-zone management theory for coal-mine water prevention and control and its significance. *Journal of China Coal Society*, 49(8), 3605–3618. (in Chinese)

ISSN: 2414-1895 DOI: 10.6919/ICJE.202609_12(9).0010

[3] Qin, Z. J. (2023). Application and evaluation of the analogy method for mine water inflow prediction. *China Metal Bulletin*, (11), 222–224. (in Chinese)

[4] Yang, S. L., Lian, H. Q., Xu, B., & et al. (2023). Application of robust deep learning models to predict mine water inflow: Implication for groundwater environment management. *Science of the Total Environment*, 871, 162056.

[5] Lian, H. Q., Li, Q. X., Wang, R., & et al. (2024). Mine water inflow prediction using a deep-learning LSTM-GRU ensemble model. *Safety in Coal Mines*, 55(9), 166–172. (in Chinese)

[6] Ding, Y. Y., Yin, S. X., Lian, H. Q., & et al. (2024). Multi-factor coal-face water inflow prediction based on an SSA-CG-Attention model. *Coal Geology & Exploration*, 52(4), 111–119. (in Chinese)

[7] Bates, J. M., & Granger, C. W. J. (1969). The combination of forecasts. *Operational Research Quarterly*, 20(4), 451–468.

[8] Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1), 55–67.

[9] Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3rd ed.). OTexts.

[10] Zhou, R., Wei, Z. Y., Zhang, Y. B., & et al. (2018). Time series prediction of tomato yield based on LSTM recurrent neural network. *Water Saving Irrigation*, (8), 66–70.