Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 2 No.6 (2025)

Stock Price Forecasting Based on CEEMDAN-AM-BiLSTM Hybrid Model

Ruixue Bao[1], Junxiang Lu[1,2]*

1.School of Science, Xi'an Polytechnic University, Xi'an, Shaanxi 710600, China

2.International Science and Technology Cooperation Base for Big Data Analysis and Algorithms, Xi'an, Shaanxi 710048, China

Corresponding author: Junxiang Lu

Abstract: The fluctuation of stock prices is closely linked to a country's economic development. However, due to the significant non-linear and non-stationary characteristics of price fluctuations, traditional prediction methods struggle to capture their underlying patterns. A CEEMDAN-AM-BiLSTM prediction model is established to predict stock prices in this paper. The original data series is first decomposed to obtain Intrinsic Mode Functions (IMFs) and residual terms by using CEEMDAN. These components are then classified into high-frequency disturbance terms and low-frequency non-disturbance terms based on the spectral characteristics of IMFs. The attention mechanism is employed to identify and focus on key IMF components, which are subsequently input into a BiLSTM network for predicting non-disturbance terms. The prediction results of each IMF component are merged to derive the final predicted value. An empirical study using the minute-level closing prices of the CSI 300 index is conducted, with comparisons made against the traditional BiLSTM model and CEEMDAN-BiLSTM model. The results show that the proposed model achieves higher accuracy in high-frequency closing price prediction and is more effective in capturing the complex features of high-frequency financial data, providing a new methodological reference for improving the precision of financial market trend forecasting.

KeyWords: Stock Price Prediction; CEEMDAN; Attention Mechanism; BiLSTM; High-Frequency Financial Data

Published: Dec 30, 2025

DOI: https://doi.org/10.62177/apemr.v2i6.997

1.Introduction

Being the core of the modern economic system, the stock market's price fluctuations are influenced by multiple factors, including the macroeconomic environment, industry trends, market sentiment, policies, and international politics. Stock price forecasting remains a challenging research topic, characterized by strong non-linearity, dynamics, and uncertainty. Although China's stock market has a short history, it has experienced rapid growth with a huge scale, exerting significant impacts on the national economy. Accurate price prediction is of great practical significance for financial regulators to monitor risks and make policies, as well as for ordinary investors to optimize investment decisions. The objective of this study is to develop a systematic model to enhance market participants' understanding of price volatility and improve prediction accuracy.

In recent years, driven by the proliferation of big data and artificial intelligence, machine learning and deep learning methodologies have witnessed extensive application in stock price prediction research [1]. Jiang et al. [2] conducted a comparative study on the LSTM and RNN models with the Shanghai Composite Index and the Dow Jones Industrial Average as the research objects, and the results showed that the LSTM model had superior performance in stock price prediction. Samal et al. [3] compared the LSTM model with the BiLSTM model and found that the BiLSTM model outperformed the LSTM model in terms of prediction performance. Deep learning, a subfield of machine learning rooted in artificial neural network architectures, has demonstrated remarkable superiority over shallow machine learning models and traditional data analysis approaches across diverse domains. Neural network architectures exhibit diverse configurations, including backpropagation (BP) networks, recurrent neural networks (RNNs), and their variants. Among these, Long Short-Term Memory (LSTM) a pivotal framework in deep learning research has achieved notable advancements in financial forecasting, emerging as a cutting edge methodology in the field. L Federico et al.[4] improved the LSTM model by adding a learnable nonlinear projection of the cell state, which enhanced the capability of capturing long-term dependencies. This architectural innovation enables LSTM to effectively process long-range contextual information, distinguishing it as a preferred choice for modeling the complex dynamics inherent in financial time series data.

LSTM, a specialized variant of recurrent neural networks (RNNs), is uniquely suited for processing long input sequences by leveraging its inherent capability to model temporal dependencies and nonlinear dynamics in stock market data. Fischer and Krauss [5]demonstrated that LSTM outperforms traditional models in predicting S&P 500 stock prices, with profits derived from high volatility and short-term reversal characteristics. Notably, Moghar and Hamiche[6] introduced an LSTM-based RNN architecture for predicting the opening price trends of GOOGL and NKE, with empirical results validating the model's predictive efficacy. In a parallel line of research, Vidal and Krist Janpoller[7] proposed a hybrid CNN-LSTM framework for gold price volatility forecasting, demonstrating that the integrated model outperforms standalone CNN or LSTM architectures by virtue of its enhanced ability to extract multi-scale temporal features. Additionally, Ashy et al. [8]proposed a hybrid deep model integrating attention mechanism and LSTM for predicting India's stock market, and the results showed that the model achieved favorable prediction accuracy.

Bidirectional Long Short-Term Memory (BiLSTM), an extension of the traditional unidirectional LSTM, was proposed to enhance model prediction accuracy by enabling bidirectional temporal feature learning. Jia et al.[9] utilized a BiLSTM model for GREE stock price forecasting, demonstrating its predictive superiority over the unidirectional LSTM framework. Wang et al. [10]comparative study has demonstrated that bidirectional long short-term memory (BiLSTM) models with enhanced data training outperform traditional LSTM architectures in terms of prediction accuracy. Empirical results further indicate that BiLSTM surpasses both ARIMA and unidirectional LSTM models in capturing complex temporal dependencies, though its computational complexity leads to significantly slower convergence compared to LSTM-based counterparts. In a related line of research, Li et al.[11] proposed a CEEMDAN-SE-BiLSTM hybrid model-an advanced iteration of the BiLSTM framework for daily flow prediction at the Huayuankou Hydrological Station in the Lower Yellow River. Through comparative analysis with CEEMDAN-BiLSTM and standard BiLSTM models, the study concluded that the proposed architecture exhibits optimal performance in handling hydrological time series dynamics.

The attention mechanism in neural networks functions as a computational resource allocation strategy, directing processing capabilities toward critical information to address the challenge of information overload under constrained computational resources. Seo et al.[12] investigated the impact of attention mechanisms on stock price prediction, evaluated their contribution to improving predictive performance, and proposed an optimized CSI 300 stock prediction model based on neural networks. The extended model exhibited significant performance improvements over its baseline counterpart, highlighting the efficacy of attention mechanisms in enhancing temporal pattern recognition. Xian et al.[13] and others innovatively proposed a new fuzzy time series model (NFTSM) based on the improved sparrow search algorithm (ISSA) and CEEMDAN, which realized the accurate prediction of the closing price of the Nasdaq.

Notwithstanding the remarkable predictive potential of deep learning in stock price forecasting, existing studies have rarely integrated CEEMDAN mode decomposition, the AM attention mechanism, and BiLSTM architectures for stock trend prediction. By virtue of CEEMDAN's efficacy in financial data denoising, BiLSTM's bidirectional feature extraction capability, and the attention mechanism's capacity to mine temporal dependencies, this study constructs a CEEMDAN-AM-BiLSTM hybrid model. Empirical validation is performed using minute-level closing prices of the CSI 300 index spanning from December 2, 2024, to March 28, 2025. The model aims to enhance the accuracy, generalization capability, and fitting performance of financial stock price prediction by synergizing CEEMDAN's decomposition-denoising functionality, BiLSTM's bidirectional feature extraction, and the attention mechanism's focus on critical Intrinsic Mode Function (IMF) components. This integration is designed to address the information complexity in financial time series, enabling more precise capture of latent patterns and nonlinear dynamics.

2.Relevant model

2.1 CEEMDAN

CEEMDAN（Complete Ensemble Empirical Mode Decomposition with Adaptive Noise）is an advanced signal decomposition technique designed for nonlinear and nonstationary signals, including vibration signals, biomedical data, and financial time series. As an optimized version of traditional EMD (Empirical Modal Decomposition) and its variants EEMD (Ensemble EMD) and CEEMD, it addresses critical limitations such as modal aliasing, noise residue, and computational inefficiency. In financial markets, high-frequency data typically contain substantial noise and exhibit strong nonlinear and nonstationary characteristics. Traditional feature extraction methods like Fourier transform struggle to capture local data structures, while linear models fail to characterize the complex dynamics of financial markets. CEEMDAN's adaptive decomposition capability enables it to break down high-frequency data into intrinsic mode function (IMF) components across diverse frequency bands based on data-driven characteristics, facilitating precise extraction of effective information. Through its noise-assisted decomposition mechanism, high-frequency noise energy is distributed across multiple IMFs, allowing for efficient noise reduction via thresholding or selective reconstruction, thereby enhancing data quality and stability. Moreover, by leveraging adaptive noise injection and staged decomposition, CEEMDAN significantly mitigates modal aliasing issues, achieving accurate modal separation and demonstrating robust performance. The algorithm's decomposition of data into interpretable IMF components also enhances model transparency, providing clear feature dimensions for subsequent analysis. As such, CEEMDAN offers a powerful solution for feature extraction and noise processing in high-frequency financial datasets.

The main decomposition process is as follows:

(1) White noise $n_0(t)$ with zero mean and standard deviation σ_0 is added to the original signal $x(t)$, then, $x_0(t)=x(t)+n_0(t)$, which is decomposed to obtain the first Intrinsic Mode Function (IMF) with component $c_1(t)$ and the residual component $\mathrm{r}_1(t)$;

(2) The standard deviation σ_1 of $c_1(t)$ is calculated. White noise $n_1(t)$ with zero mean and standard deviation σ_1 is added to the residual component $\mathrm{r}_1(t)$, resulting in $x_1(t)=r_1+n_1(t)$. Then, EMD on $x_1(t)$ is performed to obtain the second IMF with component $c_2(t)$ and the new residual component $\mathrm{r}_2(t)$;

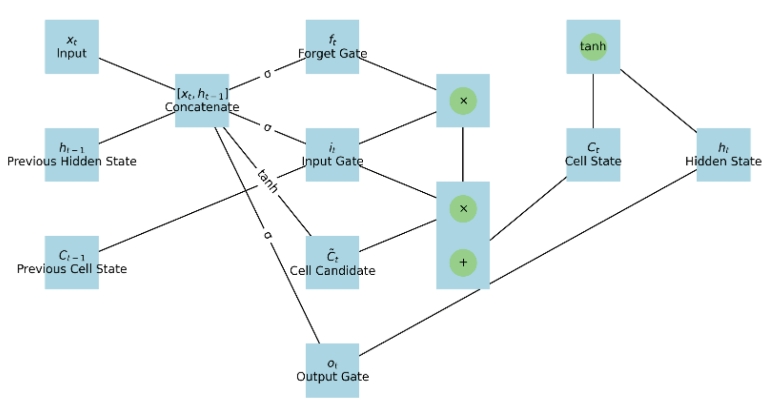
(3) The above steps are repeated until a predefined stopping criterion is met, typically when the residual component satisfies certain stationarity conditions or a preset number of decomposition layers is reached. Finally, a series of IMF components $c_1(t)$, $c_2(t)$, … $c_n(t)$ and the residual component $\mathrm{r}_n(t)$ are obtained.

2.2 LSTM and BiLSTM

2.2.1 Long Short-Term Memory

The LSTM (Long Short-Term Memory) model, as an advanced variant of Recurrent Neural Networks (RNNs), effectively addresses the gradient vanishing exploding issues that plague traditional RNNs when processing long sequential data.

Figure 1 Schematic of LSTM



The LSTM model primarily comprises forgetting gates, input gates, output gates, and cell states, as illustrated in Fig. 1. The forgetting gate selectively retains or discards historical state information, the input gate precisely regulates the current cell state update, and the output gate controls the transmission of internal state to the external hidden state h_t. Given the current input x_t, the final output h_t is derived through the following equations:

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

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

$$\tilde{C}_t = \tanh\left(W_C \bullet \left[h_{t-1}, x_t\right] + b_C\right) \tag{3}$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{4}$$

$$o_t = \sigma\left(W_o \bullet \left[h_{t-1}, x_t\right] + b_o\right) \tag{5}$$

$$h_t = o_t * \tanh\left(C_t\right) \tag{6}$$

Here, f_t, i_t, and o_t represent the forgetting gate, input gate, and output gate, respectively; $\tilde{C}_t$ denotes the candidate cell state being updated; C_{t-1} is the previous cell state; σ signifies the sigmoid activation function; tanh represents the hyperbolic tangent function; W_f、W_i、W_c、W_o are weight matrices; and b_f、b_i、b_c、b_o are bias terms.

However, the unidirectional architecture of LSTM exhibits inherent limitations in analyzing high-frequency data. This structure can only predict current stock prices based on historical information, failing to leverage potential influencing factors embedded in future data-yet in financial markets, the impact of future events or information on current stock prices is often non-negligible. Moreover, the unidirectional design constrains its capability to deeply excavate and extract features from complex data patterns.

2.2.2 Bidirectional Long Short-Term Memory

The BiLSTM model fundamentally relies on LSTM cell states to retain key information over long sequences, while enhancing its capability to capture cross-cycle features through bidirectional information flow.

In processing high-frequency stock data, BiLSTM outperforms unidirectional LSTM models significantly. Unlike LSTM, BiLSTM's bidirectional architecture overcomes the constraints of one-way information processing: it not only extracts historical trend patterns but also captures the potential influence of future events on current stock prices via backward propagation, thereby integrating temporal dependencies from both past and future contexts.

BiLSTM leverages its gating mechanisms and diverse activation functions to effectively model high-dimensional nonlinear mappings, overcoming the limitations of linear models and shallow networks to accurately characterize complex stock price fluctuations. Its forward and backward LSTM layers independently extract temporal features from opposite directions, and their integrated outputs capture comprehensive sequence context dependencies, enabling deep analysis of complex correlations in time-series data. When combined with CEEMDAN, BiLSTM further enhances noise suppression and signal pattern separation capabilities, establishing itself as a critical technical tool for stock price prediction.

2.3 Attention Mechanism

The attention mechanism (AM) enables automatic identification and emphasis on critical event periods. Unlike traditional LSTM models, which suffer from long-range information memory decay, AM dynamically assigns weights to strengthen associations with key historical nodes while suppressing high-frequency noise and extreme outliers, thus effectively mitigating model overfitting.

In stock closing price prediction, integrating AM with CEEMDAN and BiLSTM offers distinct advantages. When combined with CEEMDAN, AM accurately focuses on key IMF components derived from decomposition, efficiently filtering data noise, highlighting core trend information, and significantly enhancing feature extraction efficiency and signal purity in complex financial data. This strengthens the model's capability to identify price fluctuation patterns. When paired with BiLSTM, AM breaks the traditional equal-processing mode of bidirectional information by dynamically allocating time-step weights, enhancing the model's ability to capture long-range dependencies. This allows BiLSTM to more flexibly focus on critical information in stock price sequences and optimize the fusion of historical and future context.

2.4 Indicators for model evaluation

This paper employs the Mean Absolute Error (MAE), Mean Squared Error (MSE), and Coefficient of Determination

(R-squared) as evaluation metrics to assess the model's fitting and prediction performance. The formulations for these indices are as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right| \tag{7}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2 \tag{8}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \overline{y}\right)^2} \tag{9}$$

Here, y_i denotes the true value of the i -th sample, $\hat{y}_i$ represents the predicted value, and n is the sample size. Smaller MAE and MSE values indicate better model-data fit. R^2 measures the model's explanatory power for real variations, ranging in [0,1]: the closer its value is to 1, the better the model's fitting performance.

3.Model Building

In this paper, CEEMDAN-AM-BiLSTM and CEEMDAN-BiLSTM-AM models are established based on BiLSTM, to predict the stock price and the model results are compared further.

(1). The univariate BiLSTM model utilizes a single historical feature column as input for prediction. Its core principle lies in modeling the temporal patterns of univariate data by capturing contextual dependencies between past and future states in the time series.

(2). The CEEMDAN-BiLSTM model incorporates preprocessing steps to enhance prediction accuracy. First, the Mann-Kendall trend test ($p < 0.05$) confirmed significant trends in the data. Leveraging CEEMDAN decomposition, the original dataset was decomposed into IMFs. IMFs were then filtered through three criteria: exclusion of components with excessive volatility (indicating noise or anomalies), removal of those with negligible magnitude and predictive contribution (based on mean/variance analysis), and identification/elimination of white noise components via the Ljung-Box test. This process effectively isolated noise from underlying trends, improving signal purity. By leveraging CEEMDAN's modal decomposition, the model minimizes noise interference, enabling BiLSTM to more accurately capture temporal trends and cyclical patterns. This synergistic approach enhances the model's analytical and predictive capabilities for time-series data.

(3). The CEEMDAN-AM-BiLSTM model enhances predictive accuracy by integrating adaptive decomposition, attention weighting, and bidirectional temporal modeling. Following CEEMDAN decomposition of the original time series into IMFs, the AM prioritizes forecasting-relevant components, enabling the model to: Dynamically assign higher weights to informative IMFs to focus on key trends; Suppress noise and redundant signals via differential weighting; Optimize input representations for BiLSTM by emphasizing salient temporal patterns. This synergy among CEEMDAN's noise-filtering decomposition, AM's adaptive feature selection, and BiLSTM's bidirectional context capture strengthens the model's capability to discern complex financial trends.

(4). To address BiLSTM's memory constraints in processing long sequences, the CEEMDAN-BiLSTM-AM model introduces an attention mechanism for optimization. By computing attention weights, the model breaks the equal-processing paradigm for time-step information, enabling flexible focus on critical sequence segments, effective capture of long-range dependencies, and enhanced long-sequence processing capabilities. Additionally, the attention mechanism dynamically adjusts bidirectional contextual information weights to further optimize information fusion.

3.1 Data sources and pre-processing

3.1.1 Data sources

The minute-level closing prices of the CSI 300 index during trading days from December 2, 2024, to March 28, 2025, were retrieved via the Tomtom financial terminal, yielding an original high-frequency dataset comprising 21,594 data points. To visually characterize the data, a line graph is plotted, as shown in Figure 2.

3.1.2 Data descriptive statistical analysis.

Descriptive statistical analysis of the raw data is presented in Table 1. Results show an extreme deviation of approximately 386.62 and a standard deviation of 70.39, accounting for 1.8% of the mean. The skewness value of -0.8169 indicates a left-skewed price series with strong volatility.

Figure 2 Line graph of raw data

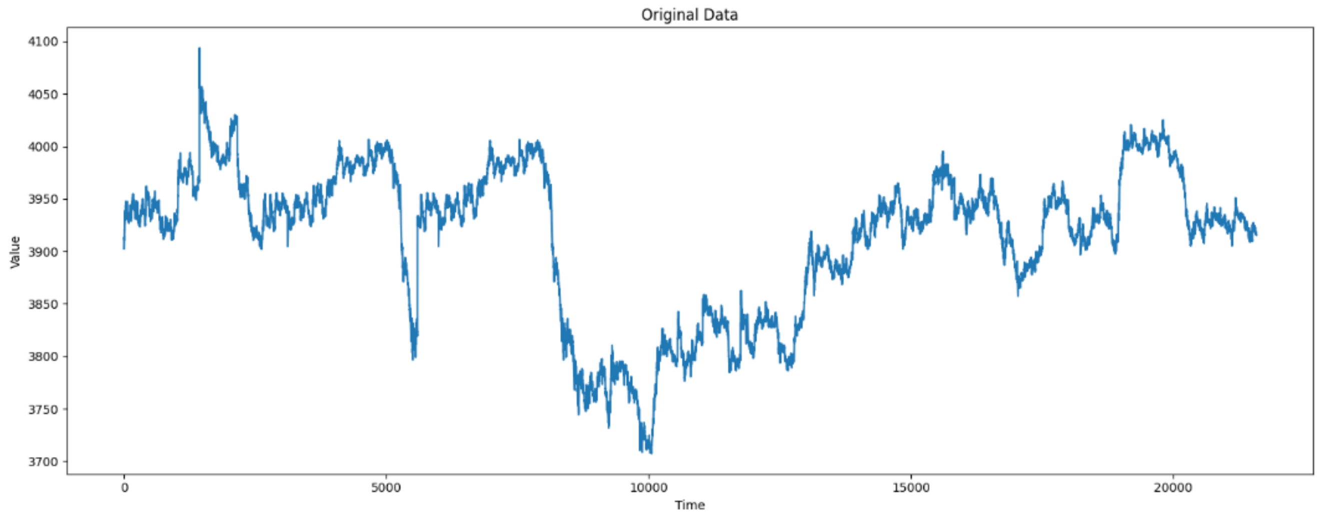


Table 1 Descriptive statistics of raw data

	volume	**maximum**	**minimum**	**average**	**statistics**	**kurtosis**	**skewness**
1-minute close	21594	4093.82	3707.20	3913.89	70.39	2.8405	-0.8169

3.1.3 Data pre-processing

To enhance algorithm performance and accelerate training, the data were normalized using the Z-score method to conform to a distribution with a mean of 0 and a standard deviation of 1. The normalization formula is as follows:

$$\hat{x}_t = \frac{x_t - \mu}{\sigma} \tag{10}$$

Here, $\hat{x}_t$ denotes the normalized stock price data, x_t represents the original data prior to normalization, and μ and σ signify the mean and standard deviation of the original series, respectively.

3.2 Empirical Research

3.2.1 BiLSTM model training and results analysis

The model employs a sliding window approach with a window size of 20 for data processing. The architecture consists of two layers, each containing 64 neurons, with dropout layers added to both to mitigate overfitting. The dataset is partitioned into an 80% training set and a 20% testing set. For training, the model is configured with 100 epochs and a batch size of 64. Early stopping is implemented to monitor validation set loss, terminating training if no improvement is observed for 30 consecutive epochs.

As depicted in Figure 3, the trained prediction results are visualized for analysis, where the blue line represents the original data and the orange line denotes the prediction results of the BiLSTM model. The orange curve closely aligns with the trend of the blue curve, demonstrating that the regularized model exhibits superior prediction performance.

Based on the comparison between model predictions and real values, evaluation metrics are calculated, with results presented in Table 2. For the test set error metrics: MAE of 1.2244 indicates minimal average prediction bias per sample; MSE of 2.9952 suggests strong performance in predicting extreme values; An R^2 close to 1 demonstrates excellent data fit. Collectively, these metrics highlight the model's effectiveness while indicating potential for further optimization to enhance predictive performance.

Figure 3 BiLSTM Price Forecast

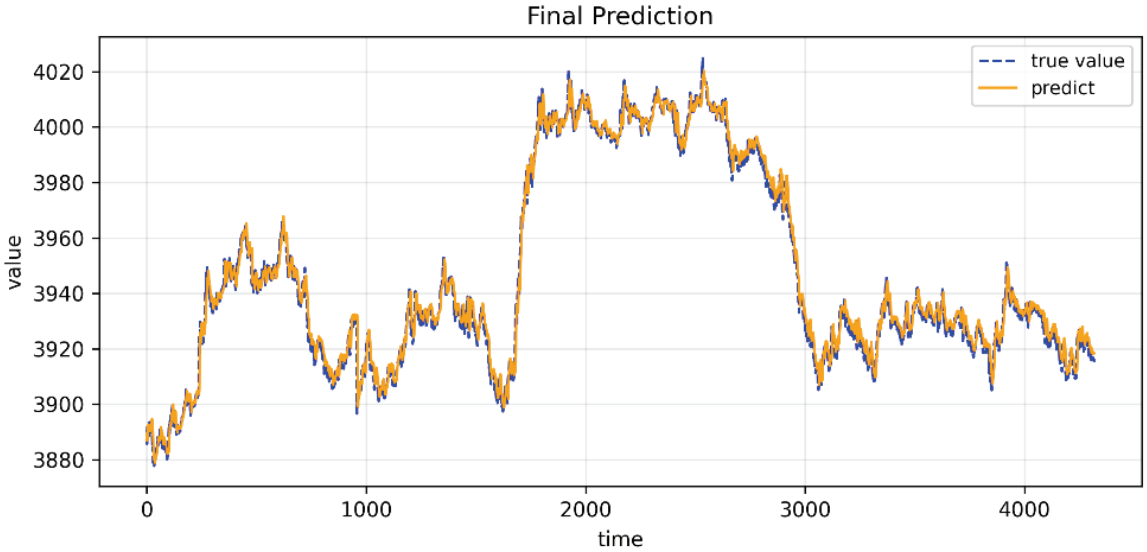


Table 2 BiLSTM model evaluation metrics

	MAE	MSE	R^2
BiLSTM	1.2244	2.9952	0.9873

3.2.2 CEEMDAN -BiLSTM model training and result analysis

The dataset underwent CEEMDAN decomposition, with results shown in Figure 4. IMF1 was excluded due to excessive volatility, while the remaining IMFs were confirmed as non-white noise. Further analysis revealed a clear trend in IMF13, prompting the application of polynomial fitting-based on fitting coefficients—to predict its behavior.

This model configures a time step of 30 and partitions the dataset into an 8:2 training-test split. The architecture includes: First Bidirectional LSTM Layer: With `units=64` and `return-sequences=True`, it preserves sequential outputs to connect subsequent LSTM layers. A Dropout layer (20% neuron dropout rate) is added to mitigate overfitting. Second Bidirectional LSTM Layer: Set to `units=64` and `return-sequences=False` to output only the final time step's result, followed by another Dropout layer for further regularization. The model is trained with `epochs=100` and a batch size of 16.

The trained model was applied to predict the test set data, generating predicted values for each IMF component. Non-interfering terms were processed using the BiLSTM model, with results visualized in Figure 6. As shown in Figure 6, the BiLSTM model effectively predicts each IMF component, largely aligning with the general trend of actual values. Figure 7 illustrates the polynomial fitting results for IMF13, where the fitted curve broadly captures the downward trend of the actual data. This indicates the model's capability to identify IMF13's overall change pattern, though opportunities for improving the precision of detailed fitting remain. By element-wise summing all IMF predictions, the total forecast reflecting the overall trend is obtained. As shown in Figure 8, the overall trends of the two curves are highly consistent, indicating that the model's forecasts effectively fit the actual values in terms of trend. Table 3 reveals that the CEEMDAN-BiLSTM model outperforms the BiLSTM model in prediction accuracy. With an MAE of 1.0304, the CEEMDAN-BiLSTM model exhibits a reduced average deviation from true values. Its MSE of 2.1431 indicates minimized prediction result fluctuations and enhanced data fit. While both models yield high R^2 values, signaling a strong linear relationship between predicted and true values, the CEEMDAN-BiLSTM model's R^2 is closer to 1, underscoring its superior linear correlation with actual data.

Figure 4 Decomposing the IMF chart

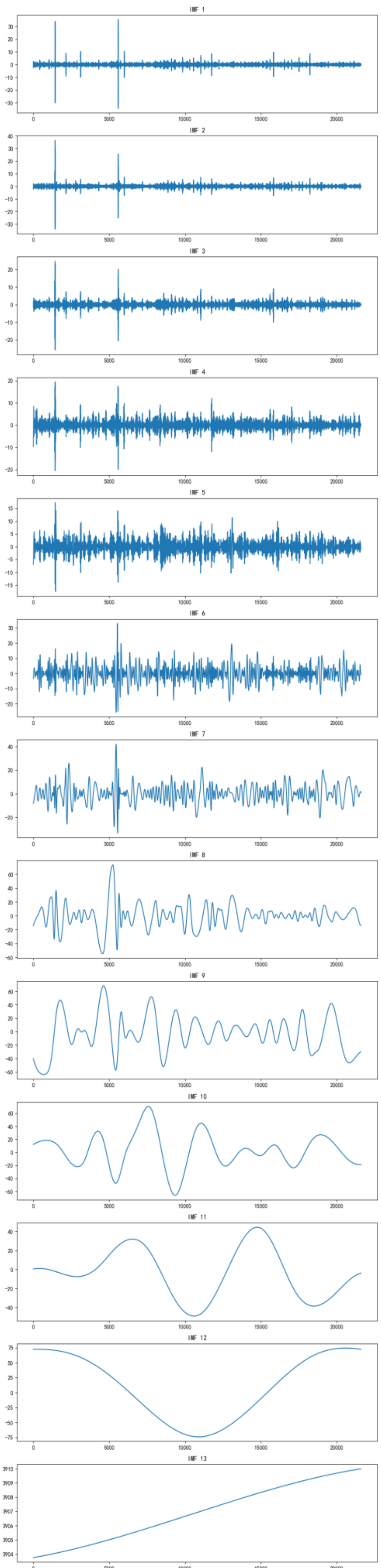


Figure 5 IMF instantaneous rate plot

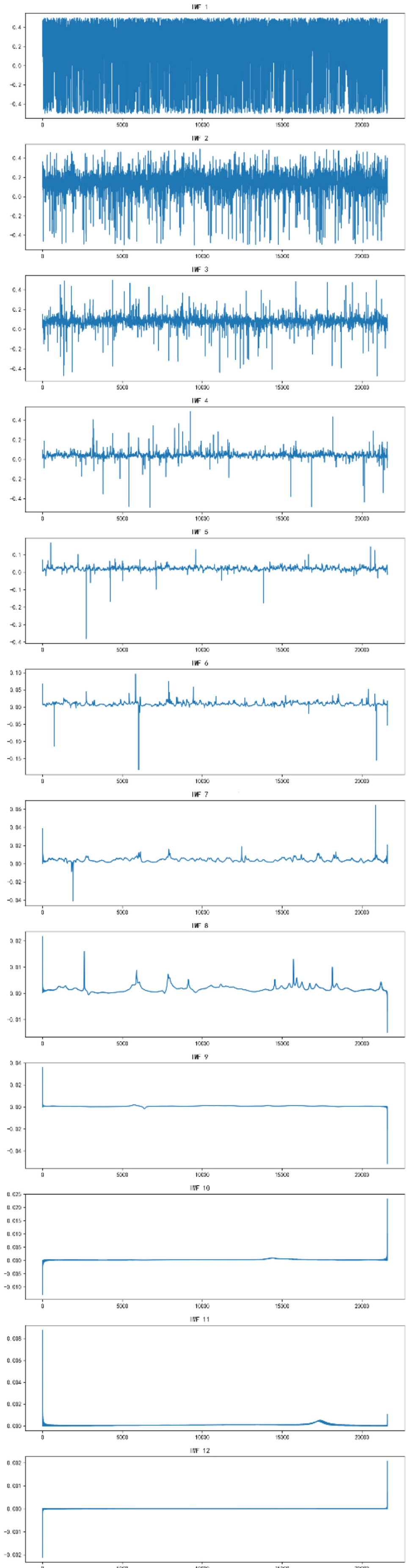


Figure 6 BiLSTM prediction IMF plot

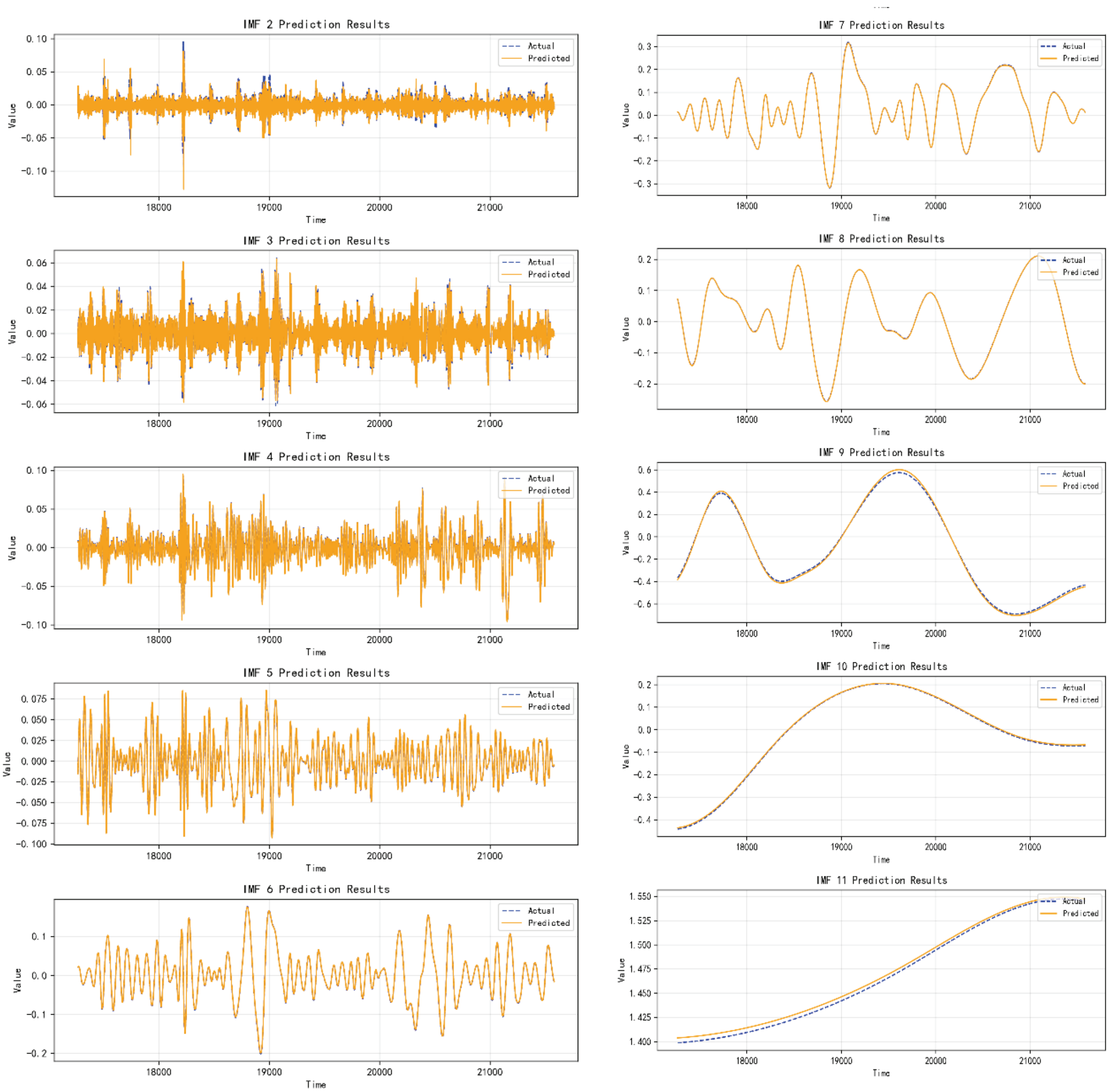


Figure 7 IMF13 polynomial fitting effects

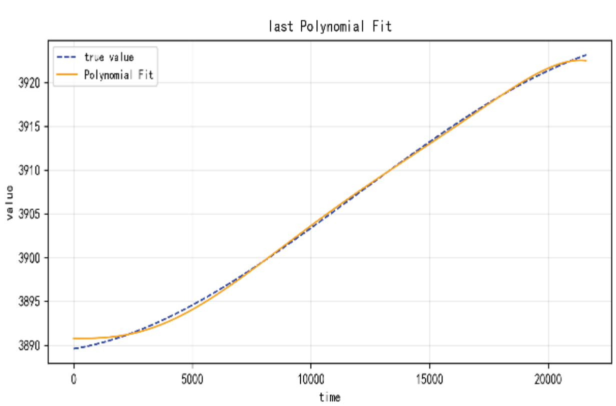


Figure 8 CEEMDAN-BiLSTM prediction effect

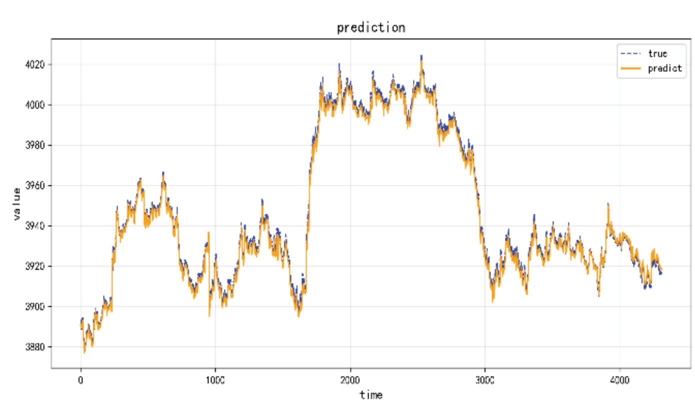


Table 3 Comparison of model evaluation indicators

	MAE	MSE	R^2
BiLSTM	1.2244	2.9952	0.9873
CEEMDAN-BiLSTM	1.0304	2.1431	0.9983

3.2.3 Temporal Attention CEEMDAN-BiLSTM-AM Model Training and Result Analysis

The model configuration is as follows: Data Preparation: A sliding window with `look-back=10` is applied, and the dataset is split into 80% training and 20% test sets. CEEMDAN Decomposition: Parameters set to `noise-width=0.25` and `trials=100`. BiLSTM Architecture: Two hidden layers each with 64 neurons, interspersed with Dropout layers (`dropout-rate=0.2`). Attention Mechanism: Integrated to dynamically weight temporal dependencies. Training Protocol: Maximum `epochs=100`, `batch-size=32`, and early stopping with validation monitoring. Post-processing: 6th-order polynomial fitting applied to the trend-dominant IMF component. Figure 9 presents the attention weight heatmap for IMF2, based on 100 test set samples with a look-back window of 10 time steps (representing the last 10 minutes). The color scale on the right indicates that lighter shades correspond to higher weights. The heatmap reveals: Temporal Focus: Lighter colors dominate the latter time steps, demonstrating the model's preference for recent stock price data. Sample-Specific Variability: Significant color differences across samples at the same time step highlight that critical moments for price movement judgment vary across periods, with certain time steps in specific samples carrying disproportionately high weights. These results indicate the model's tendency to prioritize recent trends while adapting to time-varying key influence points.

Figure 9 Heat map of IMF2 weights

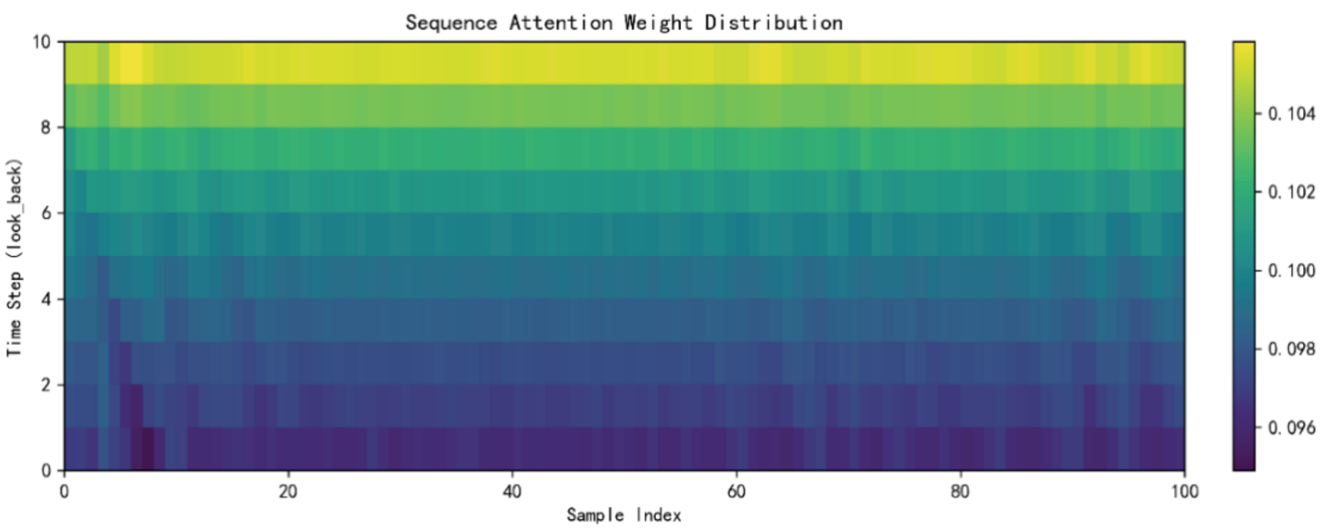


Figure 11 indicates notable biases in the model's predictions for IMF2 and IMF12. In contrast, Figure 13 demonstrates a strong polynomial fit for IMF13. Figure 14 shows that the model's total prediction curve aligns closely with the actual value curve in overall trend, effectively tracking the actual trends at peak and valley points. These results suggest that the BiLSTM model integrated with the attention mechanism can effectively capture dynamic changes in data trends.

3.2.4 Modal Attention CEEMDAN-AM-BiLSTM Model Training and Result Analysis

This model shares identical parameter settings with the CEEMDAN-BiLSTM-AM model, with the sole distinction being the application location of the attention mechanism. Figure 10 displays a histogram of IMF attention weights (IMF1 is excluded from prediction, and 0-11 in the figure correspond to IMF2-IMF13, respectively). Bar heights indicate that most IMFs have positive weights, with IMF6 exhibiting the highest weight, followed by significant contributions from IMF9 and IMF4. Notably, IMF2 carries a negative weight, representing market volatility components the model "discounts." Tests show that removing negatively weighted IMFs does not enhance forecasting performance, so they are retained to preserve data integrity.

Figure 10 IMF weighting chart

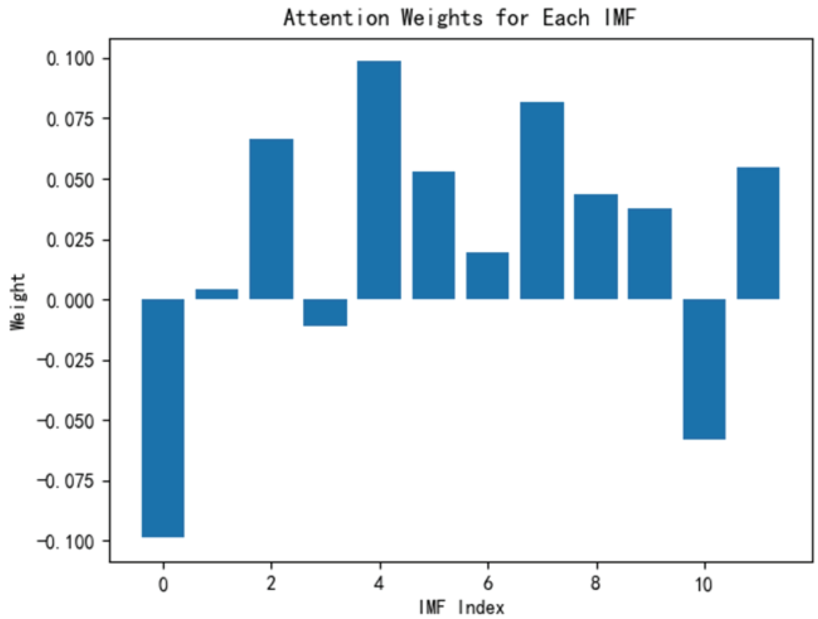


Figure 12 demonstrates the model's superior performance in predicting individual IMF components. Figure 15 highlights a strong polynomial fit for IMF13, while Figure 16 shows that the model's total prediction curve closely aligns with the overall trend of the actual data. Specifically: In the 0-2000 horizontal axis range, both curves exhibit upward-then-downward fluctuations; In the 2000-3000 range, both curves show synchronized and significant downward trends.

These results indicate that integrating the attention mechanism at the CEEMDAN decomposition layer enables the model to effectively capture the data's overall trend. Collectively, this mechanism enhances the model's capability to discern data trends and fluctuations by filtering key IMF components and eliminating redundant information, thereby improving the efficiency of time-series feature extraction.

3.3 Analysis of results

Based on the model, evaluation metrics are calculated and presented in Table 4. Results show that incorporating the attention mechanism into CEEMDAN-BiLSTM significantly improves MAE, MSE, and R^2, demonstrating enhanced model performance. When comparing attention mechanism placements (CEEMDAN decomposition layer vs. BiLSTM layer), modal attention achieves an MAE of 0.2828, which is lower than temporal attention-indicating superior control over the average absolute deviation between predicted and actual values, thus reflecting higher prediction accuracy. In terms of MSE, modal attention achieves a value of 0.1422, smaller than that of temporal attention, indicating that the model under modal attention handles errors more effectively, with a smaller mean squared deviation between predicted and actual values, reflecting higher accuracy. Regarding R^2, both values are very close to 1, with modal attention reaching 0.9998, which slightly higher than temporal attention signifying a marginal advantage in data fitting, though the difference between the two is negligible.

In summary, integrating the attention mechanism into the CEEMDAN-BiLSTM framework significantly enhances model performance for both modal and temporal attention variants. Among them, the modal attention mechanism demonstrates more prominent improvements in reducing prediction errors and enhancing goodness-of-fit, enabling more accurate predictions compared to the temporal attention mechanism and the original CEEMDAN-BiLSTM model.

Table 4 Comparison of model evaluation indicators

	MAE	**MSE**	**R**
BiLSTM	1.2244	2.9952	0.9873
CEEMDAN-BiLSTM	1.0304	2.1431	0.9983
CEEMDAN-AM-BiLSTM	0.2828	0.1422	0.9998
CEEMDAN-BiLSTM-AM	0.4450	0.2860	0.9997

Figure 11 C-B-A Forecast chart

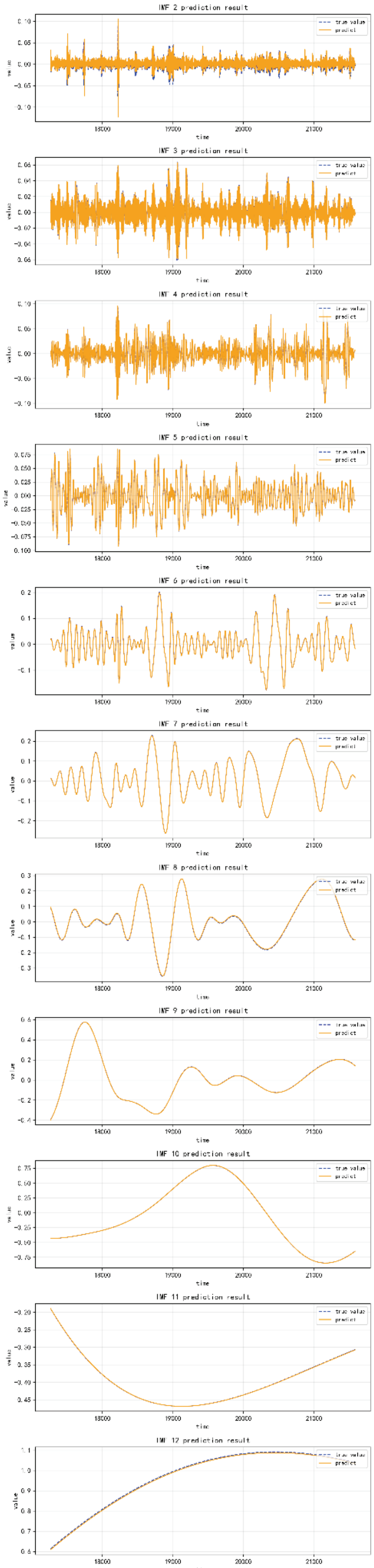


Figure 12 C-A-B Forecast Chart

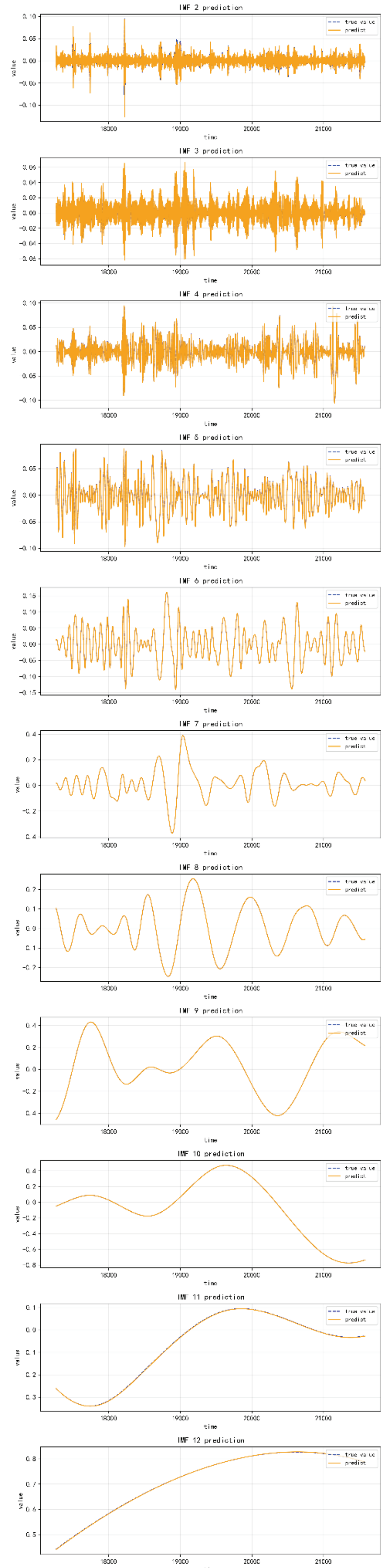


Figure 13 C-B-A IMF13 projection chart

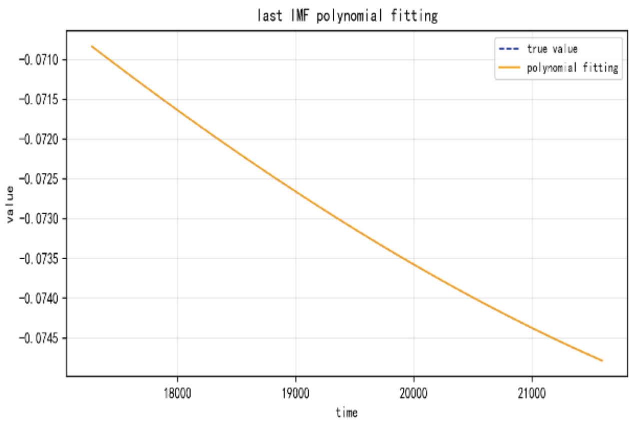


Figure 14 C-B-A General Forecast Map

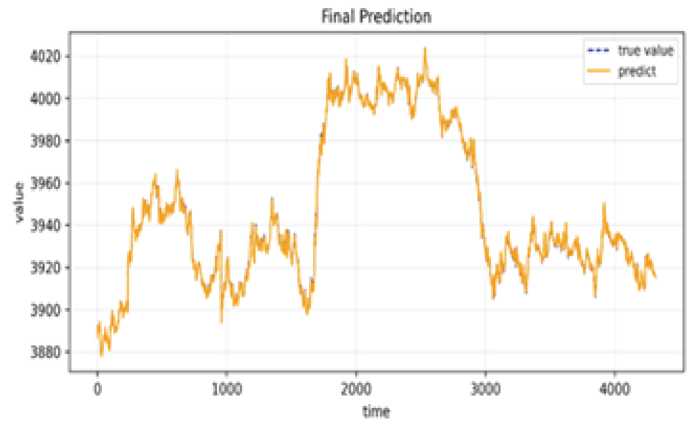


Figure 15 C-A-B IMF13 projection chart

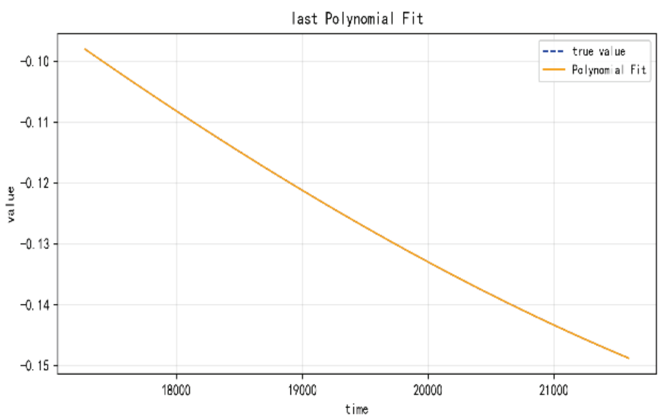
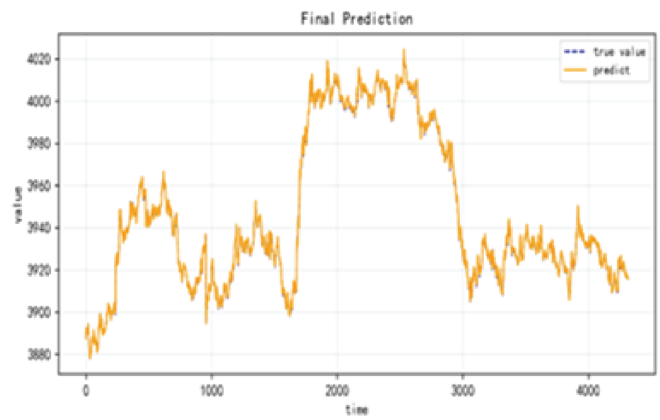


Figure 16 C-A-B General Forecast Map



4.Summaries

To address high-frequency sequence prediction challenges, we propose a novel CEEMDAN-AM-BiLSTM hybrid model. The framework integrates CEEMDAN, BiLSTM networks, and AM. Key steps include: Multi-scale Decomposition: CEEMDAN decomposes the original sequence into Intrinsic Mode Functions (IMFs), reducing noise and extracting multi-scale features. Bidirectional Temporal Learning: BiLSTM networks capture bidirectional temporal dependencies within each IMF component. Dynamic Feature Weighting: An attention mechanism dynamically assigns weights to IMF components, prioritizing key features and filtering redundant information. This hybrid approach effectively addresses limitations of traditional models in financial time-series analysis, achieving efficient integration of financial data characteristics and enhancing predictive accuracy.

Empirical results demonstrate that the hybrid model outperforms the traditional BiLSTM model across multiple evaluation metrics, significantly reducing prediction errors and validating the effectiveness of algorithmic advantage complementarity. However, the study has two notable limitations requiring improvement: Neglected Transaction Costs: In high-frequency trading scenarios, the model focuses solely on price trend prediction without incorporating transaction costs (e.g., commissions, stamp duties, slippage costs) into the modeling framework. This omission may cause discrepancies between predicted results and actual trading returns, potentially undermining the validity of investment decisions. Suboptimal Hyperparameter Tuning: The model's hyperparameter optimization relies primarily on manual trial-and-error adjustments, lacking systematic and global optimization. It fails to leverage intelligent algorithms such as Particle Swarm Optimization (PSO), Grid Search, or Bayesian Optimization for automatic parameter tuning, preventing the model from achieving its optimal performance state.

To address the above limitations, future research could explore the following directions: Integration of Transaction Costs: Develop a "prediction-cost" joint optimization model by incorporating transaction cost functions (e.g., commissions, slippage) into the objective function or constraints. This would enable a balance between prediction accuracy and trading revenue, enhancing the model's applicability to real-world high-frequency trading scenarios. Intelligent Hyperparameter Optimization: Employ global optimization algorithms (e.g., Particle Swarm Optimization, Bayesian Optimization) to systematically tune hyperparameters. Additionally, integrating reinforcement learning or evolutionary algorithms could dynamically adjust model structures and parameters, further improving adaptive capacity and predictive performance. These advancements would strengthen the hybrid model's practical utility in financial markets.

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Guresen, E., Kayakutlu, G., & Daim, U. T. (2011). Using artificial neural network models in stock market index prediction. Expert Systems with Applications, 38(8), 10389–10397. https://doi.org/10.1016/j.eswa.2010.12.083

[2] Jiang, W. W. (2021). Applications of deep learning in stock market prediction: Recent progress. Expert Systems with Applications, 184, 115537. https://doi.org/10.1016/j.eswa.2021.115537

[3] Samal, U. (2025). Evolution of machine learning and deep learning in intelligent manufacturing: A bibliometric study. International Journal of System Assurance Engineering and Management, 1–17. https://doi.org/10.1007/s13198-024-02521-8

[4] Federico, L., Lorenzo, B., Marcella, C., et al. (2021). Working memory connections for LSTM. Neural Networks, 144, 334–341. https://doi.org/10.1016/j.neunet.2021.09.010

[5] Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research, 270(2), 654–669. https://doi.org/10.1016/j.ejor.2018.03.023

[6] Moghar, A., & Hamiche, M. (2020). Stock market prediction using LSTM recurrent neural network. Procedia Computer Science, 170, 1168–1173. https://doi.org/10.1016/j.procs.2020.03.246

[7] Andrés, V., & Werner, K. (2020). Gold volatility prediction using a CNN-LSTM approach. Expert Systems with Applications, 157, 113481. https://doi.org/10.1016/j.eswa.2020.113481

[8] Sebastian, A., & Tantia, V. D. (2025). Multi-variate LSTM with attention mechanism for the Indian stock market. International Journal of Information Management Data Insights, 5(2), 100350. https://doi.org/10.1016/j.ijimdi.2024.100350

[9] Jia, M., Huang, J., Pang, L., et al. (2019). Analysis and research on stock price of LSTM and bidirectional LSTM neural network. In Proceedings of the International Conference on Computational Intelligence and Applications (ICCIA) (pp. 1–6). https://doi.org/10.2991/iccia-19.2019.72

[10] Hao, W. (2025). Enhancing stock price forecasting accuracy using LSTM and Bi-LSTM models. ITM Web of Conferences, 47, 01002. https://doi.org/10.1051/itmconf/20254701002

[11] Li, H., Zhang, X., Sun, S., et al. (2023). Daily flow prediction of the Huayuankou hydrometeorological station based on the coupled CEEMDAN-SE-BiLSTM model. Scientific Reports, 13(1), 18915. https://doi.org/10.1038/s41598-023-45863-y

[12] Seo, K., Dong, Z., & Song, Y. (2025). Neural network-based CSI300 stock prediction: Feature importance and attention mechanism analysis. Electronics, 14(9), 1729. https://doi.org/10.3390/electronics14091729

[13] Xian, S., Lei, H., Chen, K., et al. (2023). A novel fuzzy time series model based on improved sparrow search algorithm and CEEMDAN. Applied Intelligence, 53(9), 11300–11327. https://doi.org/10.1007/s10489-022-04281-6