ISSN 2616-5902 Vol. 7, Issue 7: 168-175, DOI: 10.25236/AJBM.2025.070722

A Multi-Source Time Series Forecasting Framework Based on Informer for Climate Finance and Industrial Economy

Haoyu Wu[1], Jingwen Zhang[2,*]

[1]International Business College, Dongbei University of Finance and Economics, Dalian, 116025, China
[2]School of Accounting, Dongbei University of Finance and Economics, Dalian, 116025, China
[*]Corresponding author

Abstract: *As the impact of climate change on financial and industrial systems intensifies, building a high-precision and strong generalization prediction model has become an important research direction for intelligent economic management. This paper proposes a multi-step time series prediction method based on the Informer model, which integrates multi-source heterogeneous data such as meteorology, finance and industry to achieve medium- and short-term predictions of key variables such as carbon prices and industrial output value. The constructed model retains the advantages of the Transformer structure and introduces the ProbSparse Attention mechanism and the time embedding module, which significantly improves the efficiency of long sequence modeling and the ability to identify nonlinear relationships. Experimental results show that Informer is superior to traditional machine learning and deep learning models in terms of prediction accuracy, stability and multi-step fitting ability, and is suitable for multi-variable and multi-scale time series modeling needs in complex systems. This study provides theoretical support and algorithmic basis for the application of intelligent prediction models in scenarios such as green finance, energy scheduling and industrial early warning.*

Keywords: *Informer Model, Time Series Forecasting, Climate Finance, Sparse Attention Mechanism, Industrial Economy*

1. Introduction

As global climate change intensifies, frequent extreme weather events have caused a continuous impact on multiple key industries such as energy, agriculture, and manufacturing, triggering violent fluctuations in the financial market [1]. Especially under the promotion of carbon neutrality and carbon peak policies, the coupling relationship between climate factors and the financial system and industrial economy has become increasingly complex, and "climate finance" has gradually become an important indicator for measuring systemic risks and green transformation capabilities. In this context, building a prediction model that can accurately capture the impact of climate variables on economic operations is not only of practical significance for risk prevention, but also provides data support for industrial optimization, policy formulation, and investment decisions [2].

In recent years, with the rapid development of artificial intelligence and electronic information technology, machine learning and deep learning methods have demonstrated powerful modeling capabilities in financial market modeling and economic trend forecasting [3]. However, traditional methods mostly rely on short-term historical data and have difficulty processing complex sequence data with multimodal and long-dependent structures including climate factors, and both prediction accuracy and explanatory power are limited. To address this problem, the Informer model optimized based on Transformer has become an important breakthrough direction in long-term, multivariate forecasting tasks because of its sparse attention mechanism and efficient long-sequence modeling capabilities [4].

This study is aimed at the field of climate finance, integrating three types of multi-source time series data: meteorology, finance, and industry, and constructing a multi-step prediction model based on Informer to explore the impact of climate variables on industrial economic operations and prediction patterns [5]. Through empirical modeling and comparative experiments on typical indicators such as carbon prices and industrial output value, the effectiveness of this method in climate-driven prediction tasks is verified, providing quantifiable technical support for green financial risk control and industrial policy optimization [6].

2. Related research review

As climate change deeply penetrates financial markets and the real economy, climate finance has gradually become an important intersection of green economy and financial stability research. As a core tool for measuring carbon emission costs and policy guidance, the transaction price of carbon assets is affected by climate policies, energy structure and macroeconomic factors [7]. Existing literature has conducted research from the perspectives of carbon financial market volatility modeling, carbon price forecasting, and carbon risk premium assessment, and proposed a series of statistical modeling and machine learning methods including VAR, GARCH, and SVR to detect the dynamic changes in the carbon market [8]. However, few studies have comprehensively modeled the long-term impact of climate-driven mechanisms on the evolution of carbon prices from the perspective of multi-source climate variables [9].

In the field of industrial economics and climate adaptability research, mainstream work focuses on highly climate-sensitive industries such as agriculture and energy [10]. By establishing an empirical model of the relationship between climate shocks and output, the impact of climate variability on fluctuations in the upstream and downstream of the industrial chain is revealed [11]. For example, variables such as temperature, precipitation, and drought have been proven to significantly interfere with agricultural output value, electricity demand, and manufacturing costs. Some studies have quantitatively analyzed the differences in adaptability in different regions based on panel data, gray prediction, and ecological economic models [12]. However, due to the limitations of data structure and modeling methods, these studies usually use annual or quarterly scales, making it difficult to achieve short- and medium-term forecasts and early warnings for industrial economics [13].

In recent years, machine learning and deep learning methods have been widely used in economic forecasting. For example, recurrent neural networks such as LSTM and GRU perform better than traditional linear models in stock market forecasting and energy consumption modeling [14]. In particular, deep learning has shown significant advantages in dealing with complex problems such as nonlinear relationships, long-term dependencies, and multivariate sequences [15]. Transformer and its variant models have further improved the ability to model multidimensional sequences due to their self-attention mechanism and have been applied to financial time series modeling, weather forecasting and other fields. However, existing research has mostly focused on a single task or a single data modality, lacking a collaborative modeling method for the "climate-finance-industry" complex system, and the prediction targets are mostly single-point short-term predictions, which make it difficult to cover the multi-scale prediction needs in complex economic systems.

In summary, although current research has made positive progress in climate finance, industrial adaptability and forecasting methods, it still has the following shortcomings: (1) Insufficient fusion of multi-source heterogeneous data has failed to effectively build a cross-modal forecasting system; (2) Most models are limited to single variable or single field forecasts, which makes it difficult to reveal the cross-system transmission effect of climate variables; (3) The modeling capabilities of multi-step forecasts and multi-dimensional risk indicators are limited, which is difficult to meet the needs of actual policy making and industrial decision-making. Therefore, building a multi-modal, long-term, and multi-step forecasting model based on deep learning has important research value and application prospects.

3. Model and Method

3.1 Data source and preprocessing

This study selected multi-source time series data closely related to climate financial risks and industrial economic fluctuations, covering meteorological factors, financial market indicators and macro-industry variables, and constructed a unified data set for model training and forecasting analysis. The time span of the original data is from 2016 to 2023, with a time resolution of daily. The data is collected through multiple channels and time series alignment and feature cleaning are performed. Table 1 lists the main variable types, data sources and sampling frequencies used in this study.

Since the original data has problems such as multi-source heterogeneity, missing values, and inconsistent frequencies, it needs to be processed uniformly. First, all data are aligned to daily frequencies. For monthly data (such as industrial output value), linear interpolation is used to fill in daily valuations to make them consistent with high-frequency meteorological data. For missing and outliers in the data, linear interpolation is used to handle small-scale missing values, and for mutation points, the sliding

ISSN 2616-5902 Vol. 7, Issue 7: 168-175, DOI: 10.25236/AJBM.2025.070722

window mean is combined for smoothing and repair.

Table 1: Variable Types and Data Sources

No.	Variable Type	Example Variables	Data Source	Sampling Frequency
1	Meteorological Data	Daily Avg Temperature, Precipitation, Wind Speed	National Meteorological Center, NOAA	Daily
2	Carbon Finance Data	EU ETS Carbon Price	Investing, Carbon Monitor	Daily
3	Commodity Prices	Crude Oil Price, Coal Price	WIND Database, EIA	Daily
4	Financial Indices	VIX Index, Stock Returns	Yahoo Finance, Wind Info	Daily
5	Industrial Economic Data	Agricultural Output, Electricity Consumption	National Bureau of Statistics, Industry Reports	Monthly (Interpolated to Daily)

In order to avoid the influence of the numerical scale difference of each variable on the model training effect, all input variables are normalized and normalized to (0,1) using minimum-maximum scaling. The normalization formula is as follows:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{1}$$

Subsequently, in order to meet the requirements of the Informer model for input structure, a fixed-length sliding time window is constructed for modeling. The length of the historical input sequence is set to L, the prediction step is set to T, and each sample contains an input data matrix with a shape of "L rows × D columns" (where D represents the number of feature variables) and a predicted target variable sequence with a shape of "T rows × 1 column". The sliding window sample structure is shown in Table 2.

Table 2: Input-Output Sample Structure

Sample No.	Input Range (L days)	Prediction Target Range (T days)	Number of Features (D)	Output Dimension
1	Day 1 ~ Day 60	Day 61 ~ Day 67	8	1 × 7 days
2	Day 8 ~ Day 67	Day 68 ~ Day 74	8	1 × 7 days
…	…	…	…	…

3.2 Informer model structure and design

In order to deal with the high computational cost and information sparsity problems in long-sequence, multi-variable, and multi-step prediction tasks, this study uses the Informer model based on the improved Transformer as the backbone structure to model and predict climate finance and industrial economic time series data. The Informer model retains the powerful representation ability of the Transformer and introduces the ProbSparse Self-Attention mechanism, efficient decoding strategy, and position time encoding module, which significantly improves the computational efficiency and prediction ability of the model, and is suitable for the cross-modal long-time series prediction problem involved in the current study.

The Informer model adopts the Encoder-Decoder architecture as a whole, where the encoder is used to extract key features in the historical sequence, and the decoder is used to gradually generate future prediction values. Unlike the standard Transformer, the Informer attention module uses a probabilistic sparse attention mechanism in the encoding stage, that is, only the first K query-key pairs with the largest information entropy are retained, reducing redundant calculations and improving the ability to model long-distance dependencies. The attention calculation method under this mechanism is as follows:

$$\text{Attention}(Q,K,V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V \Rightarrow \text{ProbSparseAttention}(Q,K,V) \tag{2}$$

In Informer, ProbSparseAttention performs sparse processing on the formula, retaining only the most informative combination Q-K. Q represents the query matrix (Query), K represents the key matrix (Key), V represents the value matrix (Value), and d is the scaling factor of the feature dimension.

The model input adopts a "value + time" embedding structure (Value Embedding + Time Embedding), that is, the original multivariate sequence is embedded, and the time information (such as date, week, month, solar term, etc.) is encoded into a high-dimensional vector and added to the input to enhance the model's ability to understand periodicity and trend. The overall model structure is shown in Figure 1:

ISSN 2616-5902 Vol. 7, Issue 7: 168-175, DOI: 10.25236/AJBM.2025.070722

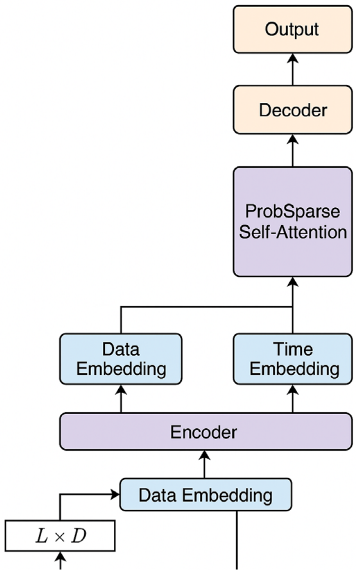

Figure 1: Informer structure diagram

Informer supports predicting multiple time steps at a time (multi-step prediction), avoiding the error accumulation problem of the traditional recurrent network model's step-by-step rolling prediction. Its training uses the standard mean square error loss function (MSE):

$$\mathcal{L} = \frac{1}{T}\sum_{i=1}^{T}(y_i - \hat{y}_i)^2 \tag{3}$$

3.3 Model training and evaluation indicators

After completing data preprocessing and structure definition, the Informer model constructed in this study needs to go through a complete process of model training, validation, and testing to ensure the reliability and generalization ability of the prediction performance. During the training process, the original data is divided into a training set, a validation set, and a test set, with the ratio set to 70%, 15%, and 15%. The training set is used to update model parameters, the validation set is used to adjust the participants to prevent overfitting, and the test set is used for final performance evaluation.

In terms of the choice of optimizer, the Adam optimization algorithm is used, which has the ability to adjust the learning rate adaptively and is suitable for processing non-stationary time series data. The initial learning rate is set to 0.0001, the batch size is set to 32, the maximum training round (epoch) is 100, and the Early Stopping mechanism is enabled during the training process to avoid overfitting. In order to comprehensively evaluate the prediction effect of the model, the following four common indicators are used to quantify the performance:

Mean Absolute Error (MAE):

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right| \tag{4}$$

Root Mean Square Error (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \tag{5}$$

Mean Absolute Percentage Error (MAPE)

$$\text{MAPE} = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{6}$$

Coefficient of determination(R^2)

ISSN 2616-5902 Vol. 7, Issue 7: 168-175, DOI: 10.25236/AJBM.2025.070722

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2} \tag{7}$$

The above four indicators can comprehensively evaluate the performance of the prediction model from different perspectives, namely, absolute error, relative error and model fitting ability. In the subsequent experimental analysis chapter, the prediction performance of the Informer model and other comparison models will be compared based on these indicators.

4. Experimental results and analysis

4.1 Experimental setup description

In order to verify the effectiveness of the proposed Informer model in climate finance and industrial economic forecasting tasks, this paper constructs a unified experimental framework to compare and evaluate multiple mainstream models and quantify the forecasting performance based on multiple indicators. In terms of hardware environment, the experiment was conducted on a computing platform equipped with NVIDIA GeForce RTX 3080 GPU, the operating system is Windows 11, the deep learning framework used is PyTorch 2.0, and the Python version is 3.10. GPU acceleration is enabled during model training to improve the efficiency of large-scale sequence processing.

In terms of data processing, based on the data set constructed in Chapter 3, the sliding window method is used to generate input samples. The window length is set to L = 60 (i.e., input 60 days of historical data), and the number of prediction steps is T = 7 (predicting the target variable for the next 7 days). The data is divided into 70% for the training set, 15% for the validation set, and 15% for the test set. All data have been standardized before division to ensure the consistency of input scale.

In terms of model setting, this paper selects LSTM, Transformer, LightGBM and Informer for comparative analysis: All models are trained and tested under the same data conditions. The model parameters are tuned through the validation set. The loss function uniformly adopts the mean square error (MSE). The optimizer is Adam, the initial learning rate is set to 1e-4, the batch size is set to 32, the maximum number of training rounds is 100, and the EarlyStopping tolerance is set to 10 rounds to prevent overfitting. The overall configuration of the model is shown in Table 3.

Table 3: Experimental Configuration of Different Models

Model	Architecture	Input Length (L)	Prediction Length (T)	Optimizer	Loss Function
LSTM	2 layers, 128 hidden units	60	7	Adam	MSE
Transformer	2 encoder layers, 4 heads	60	7	Adam	MSE
LightGBM	100 trees, max_depth=7	60	7	-	MSE
Informer	2 encoder + decoder, ProbSparse attention	60	7	Adam	MSE

4.2 Model prediction performance analysis

To further verify the effectiveness of the Informer model in climate finance and industrial economic forecasting tasks, this paper compares and analyzes the forecasting performance of Informer with three other models (LSTM, Transformer, LightGBM) on the test set. The comparison dimensions cover point forecast accuracy, error distribution characteristics, and multi-step forecast stability. The evaluation indicators use four standard indicators: MAE, RMSE, MAPE, and R^2.

Secondly, to demonstrate the prediction effect of the Informer model on specific samples, this paper selects typical time periods to visually compare the true value and the predicted value, and the results are shown in Figure 2. It can be observed from the figure that Informer can follow the real trend well at multiple peaks and turning points. The predicted curve fits the real curve closely, and the fluctuation rhythm is consistent, reflecting good trend fitting ability and dynamic response ability.

In addition, to further compare the distribution characteristics of the prediction errors of each model, the MAE box plot on the test set is drawn, as shown in Figure 3. The results show that the Informer model has a lower median error on multiple test samples, a narrower error distribution range, and better stability and robustness. However, the Transformer and LSTM models have a larger fluctuation range,

ISSN 2616-5902 Vol. 7, Issue 7: 168-175, DOI: 10.25236/AJBM.2025.070722

and the errors of some samples are significantly higher.

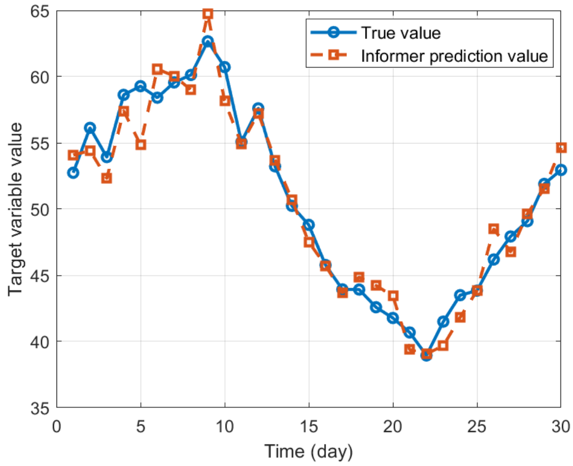

Figure 2: Comparison of the actual value and the predicted value of the Informer model

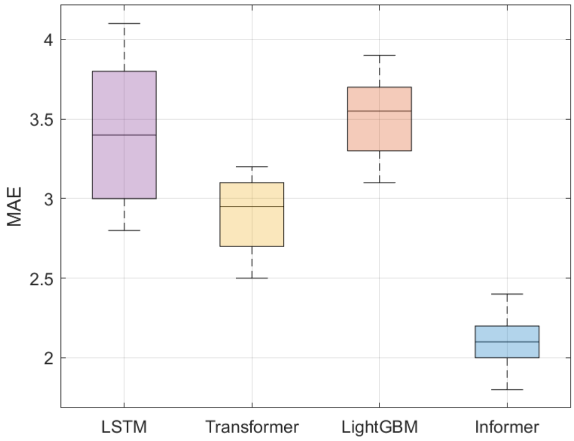

Figure 3: Box plot of prediction error distribution of each model

4.3 The impact of different feature variables on prediction performance

In the climate finance and industrial economic forecasting tasks, different input features have different impacts on the model prediction accuracy. In order to further explore the importance of various variables in the Informer model and their contribution to the final prediction results, this paper designed several sets of feature combination experiments and compared the changes in the model's prediction performance under different variable combinations.

First, based on the original input variables, we remove certain types of variables (such as meteorological, carbon finance, financial market, etc.) in turn, construct different input feature combinations, and use the Informer model to train and test under the same experimental conditions. The results are shown in Figure 4. As can be seen from the figure, when meteorological variables (such as temperature and precipitation) are removed, the performance of the model in terms of MAPE and R^2 indicators decreases significantly, indicating that climate information plays a key role in the prediction results; and when carbon price-related variables are removed, the model's prediction accuracy for industrial output value is also weakened, indicating that carbon market signals are forward-looking in changes in industrial activities.

Secondly, in order to further quantify the influence of each feature variable, this paper adopts a variable importance analysis method based on attention weights, combined with the self-attention weights of the encoder layer in the Informer model, and statistically calculates the average weights of different variables in different time steps. The analysis results are shown in Figure 5. It can be observed that temperature, carbon price and energy indicators (such as crude oil prices) have higher weights in most time windows, indicating that these variables play an important role in the model's understanding of the time series structure.

ISSN 2616-5902 Vol. 7, Issue 7: 168-175, DOI: 10.25236/AJBM.2025.070722

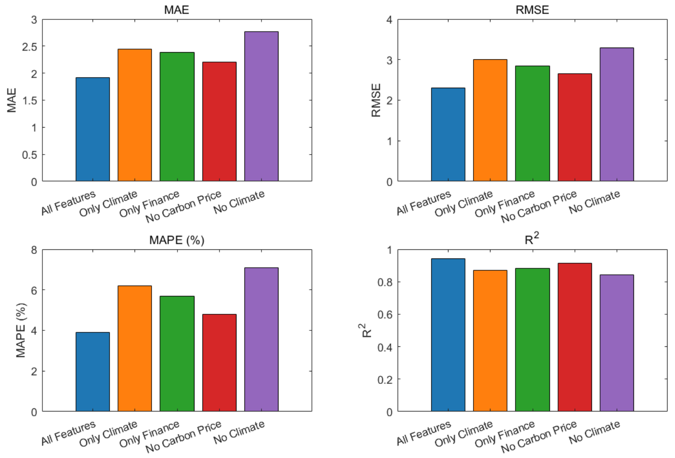

Figure 4: Comparison of the prediction performance of the Informer model under different variable combinations

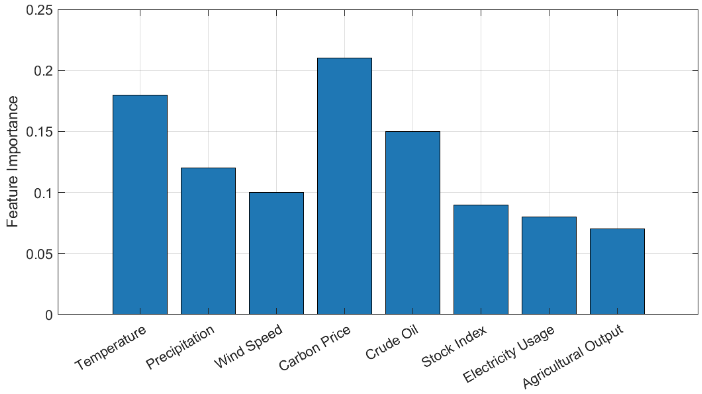

Figure 5: Histogram of average attention weights for different input variables

5. Conclusion

Aiming at the problem of financial and industrial fluctuations driven by climate change, this paper proposes a multi-step time series forecasting method based on the Informer model, which integrates multi-source heterogeneous data input and improves the forecasting accuracy and system modeling capabilities. By constructing a unified data framework including meteorological variables, carbon financial indicators and industrial output value, the deep attention mechanism is used to realize high-dimensional feature time series modeling, effectively capturing long-term dependencies and nonlinear change patterns. Experimental results show that the constructed Informer forecasting system is superior to traditional models in multiple evaluation indicators, especially in dealing with multi-step forecasts and dynamic interactions of variables.

From the perspective of modeling methods, Informer significantly reduces the time complexity of long sequence modeling by introducing the ProbSparse attention mechanism. Its Encoder-Decoder structure is stable in multi-step prediction and is suitable for computationally intensive scenarios. With the help of time position embedding and feature encoding strategies, the model can flexibly handle data streams of different frequencies and granularities, taking into account both prediction accuracy and generalization capabilities, and has good transferability. Studies have shown that the fusion information modeling strategy can significantly improve the time series perception ability of the neural network model, and provide a structural optimization idea for sequence prediction tasks in complex systems.

Future research can further combine advanced structures such as graph neural networks (GNNs) and causal attention to deeply model the dynamic relationship between climate variables and industrial economy. At the same time, it is possible to consider introducing new distributed computing frameworks such as federated learning and edge computing to achieve real-time deployment and adaptive updates of models in multi-node and multi-region systems. This direction will show broader computing value in applications such as smart energy scheduling, green financial monitoring, and urban carbon neutrality.

References

[1] Y. Wang, Z. Wang, and Y. Luo, "A hybrid carbon price forecasting model combining time series clustering and data augmentation," Energy, vol. 308, p. 132929, Nov. 2024, doi: 10.1016/j. energy. 2024. 132929.
[2] M. Mohsin and F. Jamaani, "A novel deep-learning technique for forecasting oil price volatility using historical prices of five precious metals in context of green financing – a comparison of deep learning, machine learning, and statistical models," Resour. Policy, vol. 86, p. 104216, Oct. 2023, doi: 10.1016/j.resourpol.2023.104216.
[3] M. Ji, J. Du, P. Du, T. Niu, and J. Wang, "A novel probabilistic carbon price prediction model: integrating the transformer framework with mixed-frequency modeling at different quartiles," Appl. Energy, vol. 391, p. 125951, Aug. 2025, doi: 10.1016/j.apenergy.2025.125951.
[4] A. Ma, F. Shen, Z. Li, X. Peng, and W. Zhong, "An adjustable robust optimization model under dynamic informer-based framework for industrial renewable energy systems," Process Saf. Environ. Prot., vol. 197, p. 107062, May 2025, doi: 10.1016/j.psep.2025.107062.
[5] X. Zhang, J. Wang, and X. He, "An optimal multi-scale ensemble transformer for carbon emission allowance price prediction based on time series patching and two-stage stabilization," Energy, vol. 328, p. 136457, Aug. 2025, doi: 10.1016/j.energy.2025.136457.
[6] X. Cui and D. Niu, "Carbon price point–interval forecasting based on two-layer decomposition and deep learning combined model using weight assignment," J. Cleaner Prod., vol. 483, p. 144124, Dec. 2024, doi: 10.1016/j.jclepro.2024.144124.
[7] B. Jia, H. Wu, and K. Guo, "Chaos theory meets deep learning: a new approach to time series forecasting," Expert Syst. Appl., vol. 255, p. 124533, Dec. 2024, doi: 10.1016/j.eswa.2024.124533.
[8] A. Zeb, J. Linnosmaa, M. Seppi, and O. Saarela, "Developing deep learning surrogate models for digital twins in mineral processing – a case study on data-driven multivariate multistep forecasting," Miner. Eng., vol. 216, p. 108867, Sep. 2024, doi: 10.1016/j.mineng.2024.108867.
[9] M. J. Jiménez-Navarro, M. Lovrić, S. Kecorius, E. K. Nyarko, and M. Martínez-Ballesteros, "Explainable deep learning on multi-target time series forecasting: an air pollution use case," Results Eng., vol. 24, p. 103290, Dec. 2024, doi: 10.1016/j.rineng.2024.103290.
[10] A. Nagdiya, V. Kapoor, and V. Tokekar, "Hybrid deep learning architecture combining attention mechanisms and recurrent neural networks for enhanced time series analysis," Int. Conf. Mach. Learn. Data Eng., vol. 258, pp. 3315–3323, Jan. 2025, doi: 10.1016/j.procs.2025.04.589.
[11] Z. Chen, M. Ma, T. Li, H. Wang, and C. Li, "Long sequence time-series forecasting with deep learning: a survey," Inf. Fusion, vol. 97, p. 101819, Sep. 2023, doi: 10.1016/j.inffus.2023.101819.
[12] J. Chen, A. Guan, J. Du, and A. Ayush, "Multivariate time series prediction with multi-feature analysis," Expert Syst. Appl., vol. 268, p. 126302, Apr. 2025, doi: 10.1016/j.eswa.2024.126302.
[13] D. Yang, B. Gao, S. Wang, and H. Xiang, "Robustness test for fouling state identification of homogeneous pressure electrodes based on confidence ellipsoids," IEICE Electron. Express, 2025, doi: 10.1587/elex.22.20240283.
[14] F. M. D'Arcangelo, T. Kruse, M. Pisu, and M. Tomasi, "The effect of climate policies on firm financing and investment through the banking channel," J. Environ. Manage, vol. 387, p. 125866, Jul. 2025, doi: 10.1016/j.jenvman.2025.125866.
[15] C. Zhao, K. Dong, R. Nepal, and K. K. Zander, "The panacea of heatwaves: can climate finance mitigate heatwave welfare costs?," International Journal of Disaster Risk Reduction, vol. 117, p. 105197, Feb. 2025, doi: 10.1016/j.ijdrr.2025.105197.