

Temperature and Humidity Prediction during Cross-Ecological Zone Transportation of Finished Grain Based on iTransformer and Wavelet Transform

Kun Wang^{1,a}, Fanghua Zhuo², Chao Yang^{2,b}, Muxiang Dai², Wenjing Fu²,
Lei Zhang¹, Bo Mao^{1,c,*}

¹Nanjing University of Finance & Economics, Nanjing, 210023, China

²Nan'an Branch of Fujian Grain Reserves Co., Ltd., Quanzhou, Fujian, 362341, China

^a867963893@qq.com, ^b645463665@qq.com, ^cbo.mao@nufe.edu.cn

*Corresponding author

Abstract: During the transportation of finished grain, particularly over long distances, dynamic fluctuations in temperature and humidity can affect the biochemical mechanisms of the grain, leading to increased energy loss and a higher risk of mold formation. However, existing prediction methods (e.g., LSTM, Informer) exhibit significant shortcomings in addressing these issues, especially in handling non-stationary time-series signals and cross-regional environmental coupling effects. To this end, this paper proposes a prediction method based on wavelet transform and the iTransformer model. This method innovatively constructs a dual-channel wavelet adaptive optimization module that dynamically selects the optimal wavelet basis functions for the temperature and humidity signals, respectively, to perform decomposition, achieving noise suppression and enhancement of key fluctuation features. Subsequently, the reconstructed features are input in parallel with the original data into the iTransformer model to improve prediction accuracy. Experimental results show that on the simulated Northeast-to-Yangtze-River-Delta transportation dataset, the proposed method achieves an MSE of 0.231 and an MAE of 0.233, which are significantly lower than those of the baseline and other prediction models. This research provides a new technical pathway for grain quality regulation in complex transportation environments.

Keywords: Grain Transportation; Prediction; Wavelet Transform; iTransformer

1. Introduction

Accurately predicting temperature and humidity is critical for ensuring grain quality during long-distance, cross-ecological zone transportation, a process that faces severe technical challenges. The time-series data from this context exhibits complex non-stationarity, characterized by the coupling of low-frequency trends driven by diurnal temperature variations and regional climate shifts, with high-frequency fluctuations caused by abrupt events such as loading and unloading operations. The risks of energy loss and mold formation in grain increase sharply when the temperature exceeds 25°C or the diurnal temperature difference is greater than 10°C.

Although existing time-series forecasting models, from classic LSTMs to advanced Transformer variants, have been widely applied, they share common limitations in addressing these specific challenges. These models struggle to effectively handle the non-stationary properties of the signals while simultaneously preserving long-term trends and accurately capturing the short-term, sharp fluctuations crucial for grain quality. Furthermore, the commonly used mixed-encoding approach often leads to feature interference between different variables.

To address this research gap, this paper proposes an innovative prediction method that fuses wavelet transform with the iTransformer model. The method employs a dynamic adaptive wavelet module to denoise the signals and enhance key fluctuation features to tackle non-stationarity. Subsequently, it leverages the unique independent variable encoding mechanism of iTransformer to model the multi-dimensional features without interference. The core contribution of this research lies in being the first to combine a dynamic wavelet basis selection strategy with the iTransformer architecture, providing an effective solution for the precise prediction problem in complex transportation environments and verifying its superiority through simulated experiments.

2. Related Work

In the domain of temperature and humidity forecasting for grain transportation, traditional statistical models like ARIMA^[1] struggle to fit non-linear dynamics, while classic deep learning models such as LSTM^[2] and GRU^[3] have limitations in capturing long-range dependencies, especially in accurately predicting local abrupt events within non-stationary signals.

Transformer-based models have recently become mainstream. Although the foundational Transformer model demonstrates powerful long-range modeling capabilities^[4], its variants still have specific shortcomings. For instance, Informer^[5] improves long-sequence forecasting efficiency with sparse attention but at the cost of sensitivity to high-frequency abrupt events. While PatchTST^[6] improves predictions for multi-scale series through patching, its mixed-encoding mechanism fails to resolve the issue of feature interference among multivariate inputs.

To address this feature interference, the recently proposed iTransformer^[7] model significantly enhances prediction accuracy for non-stationary multivariate signals via its innovative independent variable encoding mechanism. Concurrently, wavelet transform^[8], a technique from signal processing, is widely used for signal denoising and feature enhancement due to its excellent time-frequency analysis capabilities, and a few studies have attempted to combine it with deep learning models.^[9] However, these fusion methods typically rely on fixed wavelet basis functions, limiting their adaptability to complex signals.

In summary, existing methods either struggle to capture critical local fluctuations, suffer from feature interference, or lack an adaptive signal processing capability. This research aims to fill these gaps by fusing a dynamic wavelet selection strategy with the independent encoding advantages of the iTransformer model.

3. Methodology

3.1. Overall Framework

The core of our proposed prediction framework lies in the integration of dynamic wavelet feature enhancement with the iTransformer model. The overall process is illustrated in Figure 1: First, we independently apply a single-level wavelet decomposition to the input temperature and humidity signals. During this process, a dynamic basis selection strategy (detailed in Section 3.2.2) is employed to choose the optimal wavelet basis for each signal. The resulting low-frequency components are preserved directly to maintain long-term trends, while the high-frequency components undergo soft-thresholding denoising and a 30x feature enhancement to amplify key fluctuation characteristics.

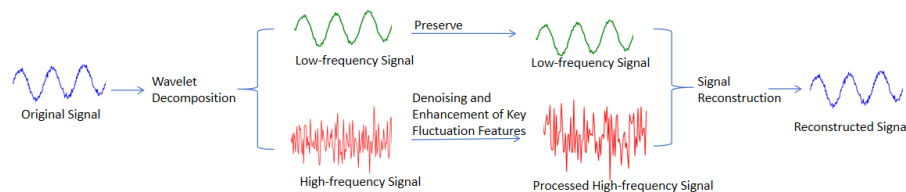


Figure 1: The process of signal decomposition and reconstruction.

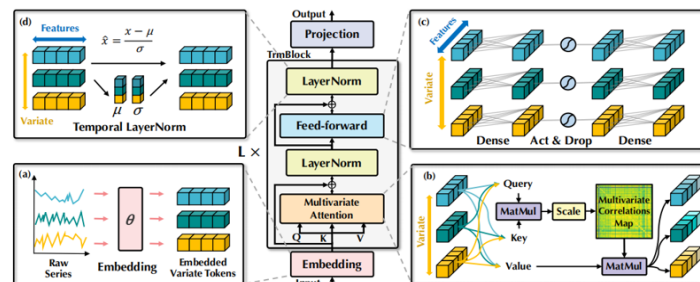


Figure 2: Internal execution details of iTransformer.

Subsequently, the processed high- and low-frequency components are reconstructed into an enhanced time series. Finally, this reconstructed series is concatenated with the original signals to form a four-dimensional input tensor (original temperature, original humidity, reconstructed temperature, reconstructed humidity), which is then fed into the iTransformer model for the final prediction. This dual-channel input approach both preserves original information and provides enhanced features. The processing of this concatenated data within the iTransformer is shown in Figure 2. This model architecture, reproduced from reference[7], performs a collaborative analysis on the four-dimensional features through its independent variable encoding mechanism.

3.2. Wavelet Feature Extraction and Enhancement

The wavelet processing module in this paper is designed to denoise the original signals and enhance key fluctuation features. We independently apply a single-level Discrete Wavelet Transform (DWT) to the temperature and humidity signals, decomposing them into low-frequency approximation coefficients (cA_1) and high-frequency detail coefficients (cD_1). A single decomposition level was chosen to effectively separate high- and low-frequency information while avoiding the boundary effects that can be introduced by multi-level decomposition.

To improve the precision of the signal decomposition, we propose a dynamic wavelet basis selection strategy. From a candidate set of 28 functions (covering Haar, Daubechies, Symlets, and Coiflets families), this strategy selects the optimal basis function for each signal by minimizing the Root Mean Square Error (RMSE) between the original and reconstructed signals, as shown in Equation (1). Experimental results identified Symlets3 (sym3) as the optimal basis for the temperature signal and Coiflets2 (coif2) for the humidity signal.

$$\text{Wavelet}_{\text{optimal}} = \arg \min_{\text{wavelet} \in \Omega} \text{RMSE}(\mathbf{x}_{\text{original}}, \mathbf{x}_{\text{reconstructed}}) \quad (1)$$

The high-frequency detail coefficients cD_1 obtained from the decomposition are first denoised using a soft-thresholding method. The threshold T is dynamically determined by Equation (2), where the scaling factor λ is set to 0.02 (optimized via grid search).

$$T = \lambda \cdot \max(|cD_1|), \quad \lambda = 0.02 \quad (2)$$

The denoised high-frequency coefficients, cD_{denoised} , are calculated using Equation (3).

$$cD_{\text{denoised}} = \text{sign}(cD_1) \cdot \max(|cD_1| - T, 0) \quad (3)$$

Subsequently, to emphasize the impact of abrupt events like loading and unloading, we apply a 30x linear enhancement to the denoised high-frequency coefficients to obtain cD_{enhanced} (Equation 4).

$$cD_{\text{enhanced}} = \alpha \cdot cD_{\text{denoised}}, \quad \alpha = 30 \quad (4)$$

Finally, the preserved original low-frequency approximation coefficients, cA_1 , are combined with the enhanced high-frequency components, cD_{enhanced} , and reconstructed via an inverse wavelet transform to produce the final enhanced signal, $\mathbf{x}_{\text{reconstructed}}$ (Equation 5). This is then merged with the original signal to form the multimodal input features (Equation 6).

$$\mathbf{x}_{\text{reconstructed}} = W^{-1}([cA_1, cD_{\text{enhanced}}], \text{wavelet}) \quad (5)$$

$$\text{Input} = [\mathbf{x}_{\text{temp}}, \mathbf{x}_{\text{humi}}, \mathbf{x}_{\text{temp_rec}}, \mathbf{x}_{\text{humi_rec}}] \quad (6)$$

4. Experiments and Results

4.1. Experimental Setup

Dataset and Preprocessing: The dataset used in this study was generated by simulating the temperature and humidity of an autumn transportation route from the Northeast to the Yangtze River Delta in an artificial climate chamber, with a sampling frequency of every two minutes. After cleaning and Min-Max normalization, the data was split into training, validation, and test sets with a 7:1:2 ratio.

Baselines and Parameters: To validate performance, we compared our method against several baselines: Transformer^[4], Informer^[5], DLinear^[10], TimeMixer^[11], PatchTST^[6], and our base model,

iTransformer^[7]. For fairness, all models were trained with a unified parameter configuration: an input sequence length of 48 and an output of 12; a hidden dimension (d_{model}) of 512 with 8 attention heads (n_{heads}); and 2 encoder layers with 1 decoder layer. The feed-forward network (FFN) dimension was 2048 with a dropout rate of 0.1. Models were optimized using Adam with the GeLU activation function, a batch size of 32, and an early stopping mechanism with a patience of 10. Detailed hardware configurations are shown in Table 1.

Table 1: Experimental environment configuration.

Environment Information	Specific Configuration
Operating System	Windows 10
Development Language	Python 3.8.20
Deep Learning Framework	PyTorch 2.0.1 + CUDA 11.8
CPU	Intel(R) Core(TM) i7-8700@3.20GHz
GPU	NVIDIA GeForce RTX 2070
Memory (RAM)	16GB

Evaluation Metrics: We evaluate model performance using two complementary metrics. Mean Squared Error (MSE), defined in Equation (7), is more sensitive to large errors, making it suitable for assessing prediction accuracy on abrupt events. Mean Absolute Error (MAE), defined in Equation (8), focuses more on overall stability and is used to measure performance in forecasting long-term trends.

In these equations, y_i is the true value, $\hat{y}_i$ is the predicted value, and N is the total number of samples.

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (7)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (8)$$

4.2. Experimental Results and Analysis

The predictive performance of our proposed method and the baselines was evaluated on the simulated dataset, with the results summarized in Table 2. Our method achieves state-of-the-art performance, reaching an MSE of 0.231 and an MAE of 0.233. This represents a significant improvement over the baseline iTransformer (3.0% and 1.7% reductions in MSE and MAE, respectively) and substantially outperforms all other models.

Table 2: Comparison results of the predictive performance of different models

Model	MSE	MAE
Transformer	0.673	0.626
Informer	1.162	0.803
DLinear	0.309	0.313
TimeMixer	0.254	0.251
PatchTST	0.240	0.236
iTransformer	0.238	0.237
Wavelet transform + iTransformer	0.231	0.233

Further analysis reveals clear performance tiers. Traditional models like Transformer and Informer perform poorly ($MSE > 0.6$), struggling with the non-stationary signals due to their mixed-encoding schemes and insensitivity to abrupt events. While more recent models like DLinear, TimeMixer, and PatchTST show considerable improvement through innovative architectures, they still lack the sensitivity to fully capture sharp local fluctuations, resulting in larger local errors.

The superiority of our method stems from the synergistic effect of its two core components. The baseline iTransformer already demonstrates strong performance ($MSE=0.238$), leveraging its independent variable encoding. Our key advantage lies in the wavelet pre-processing, which enhances the critical fluctuation details in the input features. This allows the iTransformer to more accurately model both the abrupt changes and long-term trends inherent in the transportation environment.

To more intuitively compare the prediction effects of the baseline iTransformer model and the method proposed in this paper, we provide visual comparison charts of the true and predicted values for a subset of sample points, as shown in Figures 3, 4, 5, and 6. Figures 3 and 4 show the prediction effects of the

baseline iTransformer model, while Figures 5 and 6 show the effects of the proposed Wavelet Transform + iTransformer model. Overall, both models can accurately capture the general trends of the temperature and humidity signals, and their overall predictive performances are relatively close.

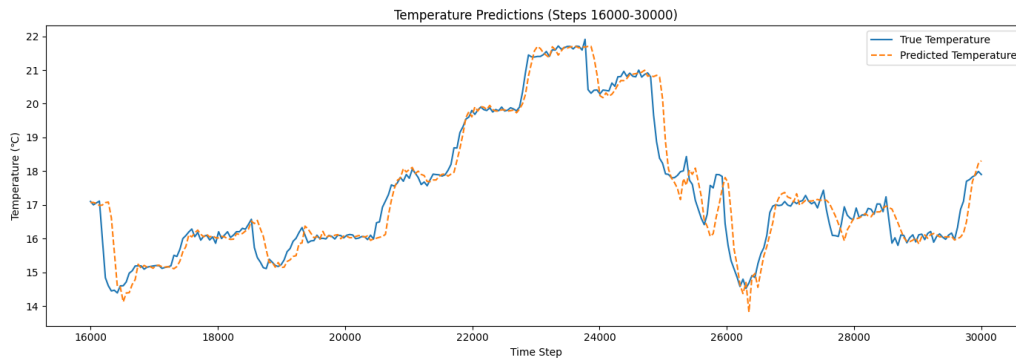


Figure 3: Temperature comparison chart predicted by iTransformer.

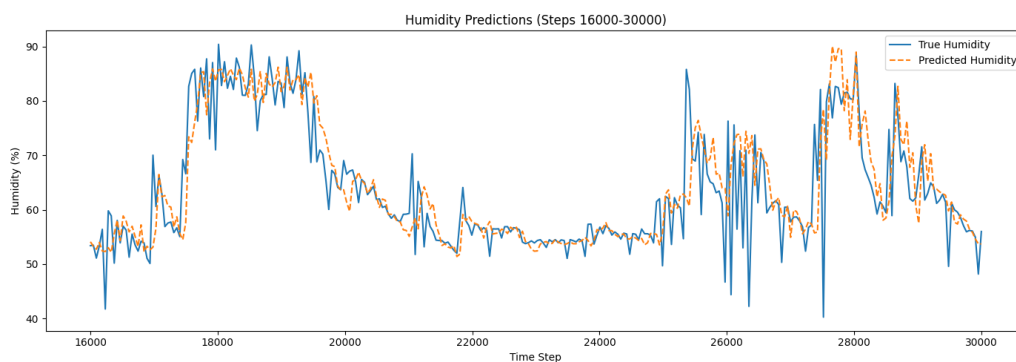


Figure 4: Humidity comparison chart predicted by iTransformer.

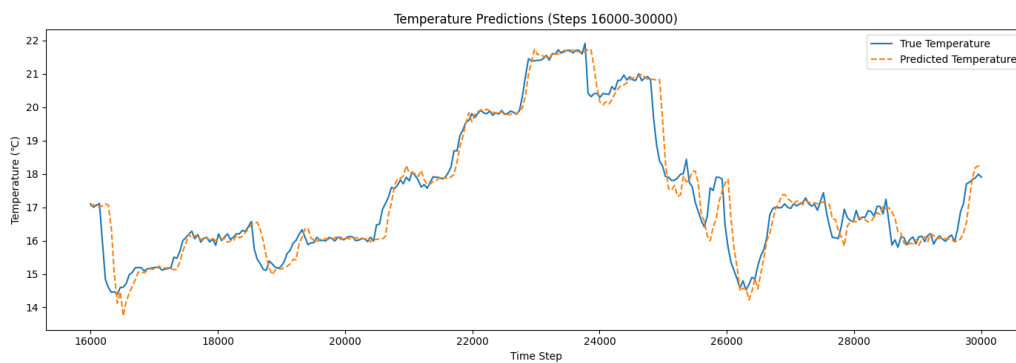


Figure 5: Temperature comparison chart predicted by wavelet transform + iTransformer.

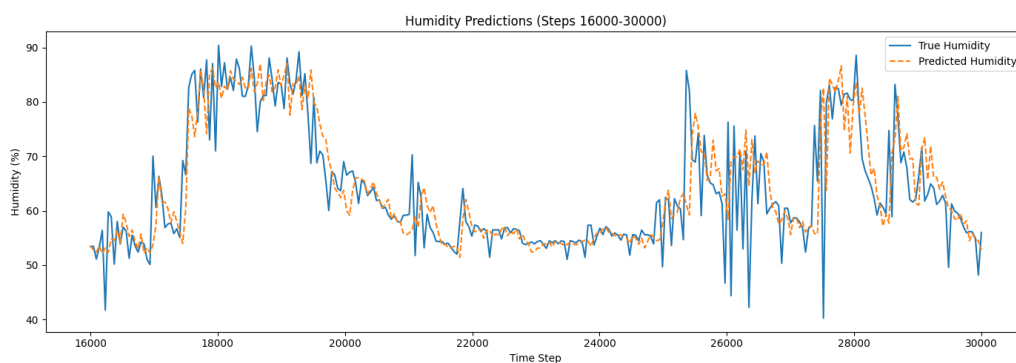


Figure 6: Humidity comparison chart predicted by wavelet transform + iTransformer.

However, a more detailed observation of the figures reveals that the Wavelet Transform + iTransformer model demonstrates higher sensitivity and better predictive performance when forecasting local sharp fluctuations or abrupt events in the temperature and humidity signals. Taking the abrupt fluctuation region around the 26,000th sample point of the temperature signal as an example, the baseline iTransformer model exhibits a noticeable error in its prediction of this fluctuation. In contrast, the Wavelet Transform + iTransformer model is able to more accurately follow the trend of the actual temperature changes, as shown in Figures 7.

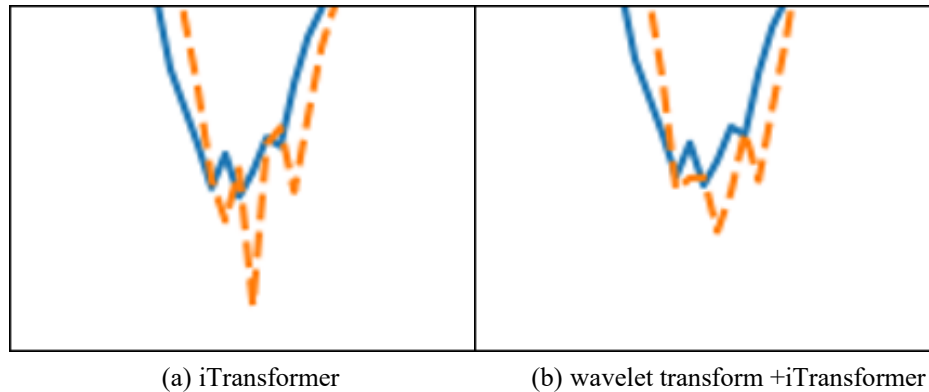


Figure 7: Local temperature comparison.

Furthermore, a similar situation is observed in the prediction of the humidity signal. Taking the region around the 28,000th sample point of the humidity signal as an example, the prediction curve of the baseline iTransformer model exhibits considerable volatility and deviates significantly from the true values in some periods. In contrast, the prediction curve of the Wavelet Transform + iTransformer model as a whole aligns more closely with the fluctuation trend of the actual signal, as shown in Figures 8.

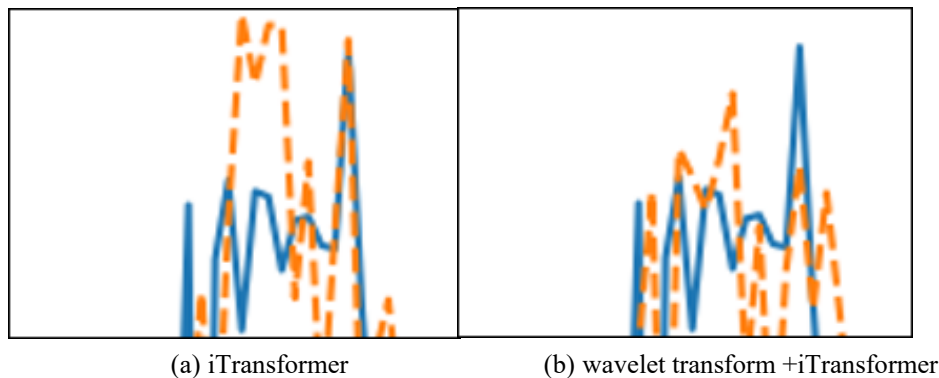


Figure 8: Local humidity comparison.

In summary, based on both the overall performance metrics and the visual comparison of the prediction curves, the proposed "Wavelet Feature Enhancement + iTransformer Model" not only significantly outperforms traditional time-series forecasting models in overall performance but also demonstrates a more pronounced predictive advantage compared to using the advanced Transformer variant alone. This confirms the innovativeness and effectiveness of the method presented in this paper.

5. Conclusion

This paper introduces a novel prediction model that fuses dynamic wavelet transform with iTransformer, effectively solving the challenge of non-stationary temperature and humidity forecasting in cross-ecological zone grain transportation. By employing an adaptive wavelet module to denoise signals and enhance key fluctuation features, combined with the independent variable encoding mechanism of iTransformer, our model achieves a precise capture of complex time-series signals. On the simulated Northeast-to-Yangtze-River-Delta transportation dataset, our method achieves State-of-the-Art (SOTA) performance (MSE=0.231, MAE=0.233), significantly outperforming a range of mainstream baselines, including Transformer and Informer. This research confirms the innovativeness and effectiveness of the proposed fusion framework, establishing a new technical pathway for precise grain

quality regulation in complex environments and offering a valuable reference for integrating signal processing with advanced Transformer architectures in time-series forecasting.

While the proposed method has demonstrated excellent performance, future work can be extended in the following directions: (1) Optimizing the dynamic wavelet basis function selection strategy to further enhance the model's adaptability; (2) Incorporating multi-source heterogeneous data, such as grain type, packaging materials, and weather factors, to explore a more comprehensive set of influencing factors; and (3) Conducting on-site validation in real-world transportation scenarios to further test and improve the model's generalization and practical value.

Acknowledgements

This research was supported by the National Key R&D Program of China (No. 2024YFF1105504) and the Postgraduate Research & Practice Innovation Program of Jiangsu Province (No. KYCX25_2127).

References

- [1] Ho S L, Xie M. The use of ARIMA models for reliability forecasting and analysis[J]. *Computers & industrial engineering*, 1998, 35(1-2): 213-216.
- [2] Yu Y, Si X, Hu C, et al. A review of recurrent neural networks: LSTM cells and network architectures[J]. *Neural computation*, 2019, 31(7): 1235-1270.
- [3] Dey R, Salem F M. Gate-variants of gated recurrent unit (GRU) neural networks[C]//2017 IEEE 60th international midwest symposium on circuits and systems (MWSCAS). IEEE, 2017: 1597-1600.
- [4] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. *Advances in neural information processing systems*, 2017, 30.
- [5] Zhou H, Zhang S, Peng J, et al. Informer: Beyond efficient transformer for long sequence time-series forecasting[C]//*Proceedings of the AAAI conference on artificial intelligence*. 2021, 35(12): 11106-11115.
- [6] Nie Y, Nguyen N H, Sinthong P, et al. A time series is worth 64 words: Long-term forecasting with transformers[J]. *arXiv preprint arXiv:2211.14730*, 2022.
- [7] Liu Y, Hu T, Zhang H, et al. itransformer: Inverted transformers are effective for time series forecasting[J]. *arXiv preprint arXiv:2310.06625*, 2023.
- [8] Zhang D. Wavelet transform[M]//*Fundamentals of image data mining: Analysis, Features, Classification and Retrieval*. Cham: Springer International Publishing, 2019: 35-44.
- [9] Wu J, Wang Z. A hybrid model for water quality prediction based on an artificial neural network, wavelet transform, and long short-term memory[J]. *Water*, 2022, 14(4): 610.
- [10] Zeng A, Chen M, Zhang L, et al. DLinear: Linear-based models for long-term time series forecasting[C]//*Proceedings of the AAAI conference on artificial intelligence*. 2023, 37(9): 11121-11128.
- [11] Wang S, Wu H, Shi X, et al. Timemixer: Decomposable multiscale mixing for time series forecasting[J]. *arXiv preprint arXiv:2405.14616*, 2024.