Journal of Electrical and Electronics Engineering (IOSR-JEEE)
e-ISSN: 2278-1676, p-ISSN: 2320-3331, Volume 21, Issue 2 Ser. 1 (Mar. – Apr. 2026), PP 10-18
www.iosrjournals.org

Short-Term Load Forecasting Using Modwt-Lstm: Multi-Scale Modeling With Temporal And Statistical Analysis

Do Quynh Huong, Do Phuong Linh
Hanoi University Of Science And Technology
Standard Chartered Bank Vietnam

Abstract:

Short-term load forecasting is a critical component in the operation and management of modern power systems, particularly in microgrid environments with high variability. This study proposes a hybrid forecasting framework based on the Maximal Overlap Discrete Wavelet Transform and Long Short-Term Memory networks (MODWT-LSTM) for hour-ahead electricity demand prediction using a Japan dataset. The MODWT is employed to decompose the original load signal into multiple frequency components, enabling the LSTM model to capture both high-frequency fluctuations and long-term temporal dependencies effectively. To validate the robustness of the proposed model, extensive experiments are conducted against multiple baseline models, including RNN, LSTM, GRU, CNN-LSTM, CNN-GRU, and DWT-LSTM. Performance is evaluated using MAE, RMSE, and MAPE metrics. Furthermore, the analysis is extended to different temporal scenarios, including weekday and special-day (holiday) conditions, to examine model adaptability under varying demand patterns. An ablation study is performed to assess the contribution of each component, and the Diebold–Mariano (DM) test is applied to verify the statistical significance of forecasting improvements. Experimental results demonstrate that the proposed MODWT-LSTM model consistently outperforms benchmark models across all evaluation scenarios, achieving superior accuracy and stability. The findings highlight the effectiveness of integrating wavelet-based decomposition with deep learning for enhancing short-term load forecasting performance in real-world energy systems.

Keywords: *Short-term load forecasting (STLF); MODWT (stationary wavelet, no-downsampling); LSTM; Hybrid wavelet–neural network; Day-ahead forecasting; Diebold–Mariano test; Ablation study*

Date of Submission: 14-04-2026 Date of Acceptance: 24-04-2026

I. Introduction

The increasing penetration of renewable energy sources and the growing complexity of modern power systems have made short-term load forecasting (STLF) an essential component for efficient energy management [1], [3]. Accurate hour-ahead load prediction plays a critical role in unit commitment, economic dispatch, and demand-side management, particularly in systems with high variability and uncertainty. In recent years, the rapid development of smart grids and microgrid technologies has further intensified the need for reliable and robust forecasting models [2], [4].

Traditional statistical approaches, such as autoregressive integrated moving average (ARIMA) and exponential smoothing methods, have been widely applied to load forecasting tasks [15]. However, these models often struggle to capture nonlinear patterns and complex temporal dependencies inherent in electricity demand data. To address these limitations, machine learning and deep learning techniques, including recurrent neural networks (RNN), long short-term memory (LSTM), and gated recurrent units (GRU), have been increasingly adopted due to their ability to model sequential data effectively [2], [4], [17].

Despite their advantages, standalone deep learning models still face challenges when dealing with highly non-stationary and multi-scale time series, such as electricity load data [11], [16]. In particular, load signals typically contain both high-frequency fluctuations and long-term trends, which are difficult to capture simultaneously using a single model. To overcome this issue, hybrid approaches that combine signal decomposition techniques with deep learning architectures have been proposed [5], [12]. Wavelet transform-based methods, such as discrete wavelet transform (DWT), have shown promising results by decomposing the original signal into multiple frequency components, allowing models to learn more structured representations [8], [14].

However, conventional DWT suffers from limitations such as lack of translation invariance and sensitivity to signal alignment. The Maximal Overlap Discrete Wavelet Transform (MODWT) addresses these issues by preserving alignment and providing redundant, shift-invariant decompositions, making it more suitable for time-series forecasting applications [8], [14]. Nevertheless, the integration of MODWT with deep learning models for short-term load forecasting remains relatively underexplored, particularly in the context of real-world datasets and comprehensive evaluation settings.

In addition, many existing studies focus solely on overall forecasting performance, without considering variations in demand patterns across different temporal contexts. In practice, electricity consumption exhibits distinct characteristics between weekdays and special days (e.g., holidays), which can significantly affect forecasting accuracy [6]. Furthermore, few studies incorporate rigorous statistical validation methods, such as the Diebold–Mariano (DM) test, to assess the significance of performance differences between competing models [16].

To address these research gaps, this study proposes a hybrid MODWT-LSTM framework for hour-ahead load forecasting using a real-world Japan electricity dataset. The proposed model leverages MODWT to decompose the load signal into multiple components, which are then fed into an LSTM-based architecture to capture both short-term dynamics and long-term dependencies. Extensive experiments are conducted to evaluate the performance of the proposed model against a range of baseline methods, including RNN, LSTM, GRU, CNN-LSTM, CNN-GRU, and DWT-LSTM [4], [7], [12].

Moreover, this study introduces a comprehensive evaluation strategy that includes analysis under different temporal conditions (weekday and special-day scenarios), ablation studies to investigate the contribution of model components, and statistical validation using the Diebold–Mariano test [16]. The main contributions of this paper can be summarized as follows::

(1) A novel hybrid forecasting framework combining MODWT and LSTM is proposed to effectively capture multi-scale characteristics of electricity load data.
(2) A comprehensive experimental evaluation is conducted, including multiple baseline comparisons, temporal scenario analysis, and ablation studies.
(3) Statistical significance of forecasting improvements is rigorously validated using the Diebold–Mariano test.
(4) The proposed approach is validated on a real-world Japan dataset, demonstrating superior performance and robustness under diverse conditions.

The remainder of this paper is organized as follows. Section II presents the methodology, including data preprocessing, MODWT decomposition, and model architecture. Section III describes the experimental setup and results. Section IV discusses the findings, and Section V concludes the paper.

II. Methodology

This section presents the proposed MODWT-LSTM framework, including the problem formulation, MODWT-based signal decomposition, and the dual-branch LSTM architecture.

A. Problem Formulation

Let $\{x_t\}_{t=1}^{T}$ denote the hourly electricity load time series. Given a look-back window of length L, the forecasting task is to predict the load at the next time step based on the past observations:

$$x_{t-L+1:t} = [x_{t-L+1}, \dots, x_t] \rightarrow \widehat{x_{t+1}} \qquad (1)$$

In this study, the look-back window is set to $L = 24$, corresponding to one-day historical information for hour-ahead forecasting.

B. Motivation and Design Rationale

Electricity load data exhibits strong multi-scale characteristics, including long-term trends (daily/weekly cycles), medium-term variations (peak transitions), and short-term fluctuations (noise) .

While LSTM models are effective in capturing long-term temporal dependencies [17], they often struggle to explicitly separate patterns across different frequency scales. In contrast, wavelet transforms decompose the signal into low-frequency (trend) and high-frequency (detail) components, enabling more structured feature learning [14].

In particular, the Maximal Overlap Discrete Wavelet Transform (MODWT) is adopted in this study due to its key advantages:

- No downsampling, preserving the original temporal resolution;
- Shift-invariance, avoiding phase misalignment;
- Alignment with the original time grid, which is crucial for hour-ahead forecasting.

Based on these observations, a dual-branch architecture is proposed, where:

1. a raw LSTM branch learns nonlinear dynamics directly from the original signal, and
2. a MODWT-based branch extracts multi-scale representations before temporal modeling.

These two representations are complementary and are fused at the feature level, enabling the model to jointly exploit temporal dynamics and multi-scale structures for improved forecasting accuracy.

C. MODWT-Based Multi-Scale Decomposition

1) Wavelet Filters

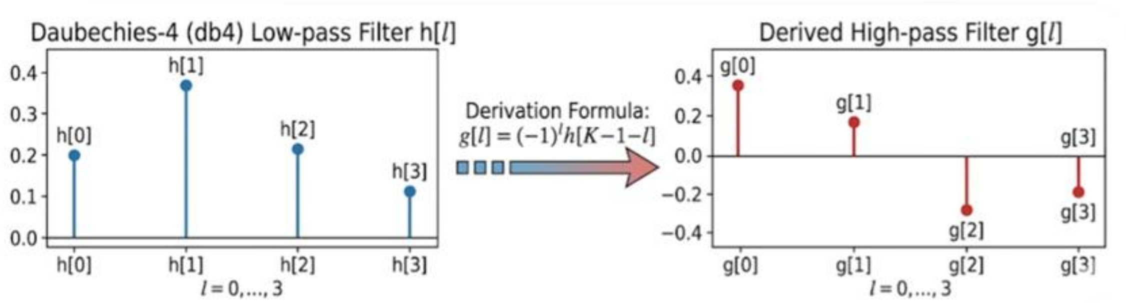

Fig. 1. Construction of db4 wavelet filters in MODWT.

The MODWT is implemented using the Daubechies-4 (db4) wavelet. Let $h[l]$ and $g[l]$ denote the low-pass (scaling) and high-pass (wavelet) filters, respectively. The high-pass filter is defined as:

$$g[l] = (-1)^{l} h[K-1-l], \quad l = 0, \dots, K-1 \qquad (2)$$

where K is the filter length.

2) Dilated Convolution without Downsampling

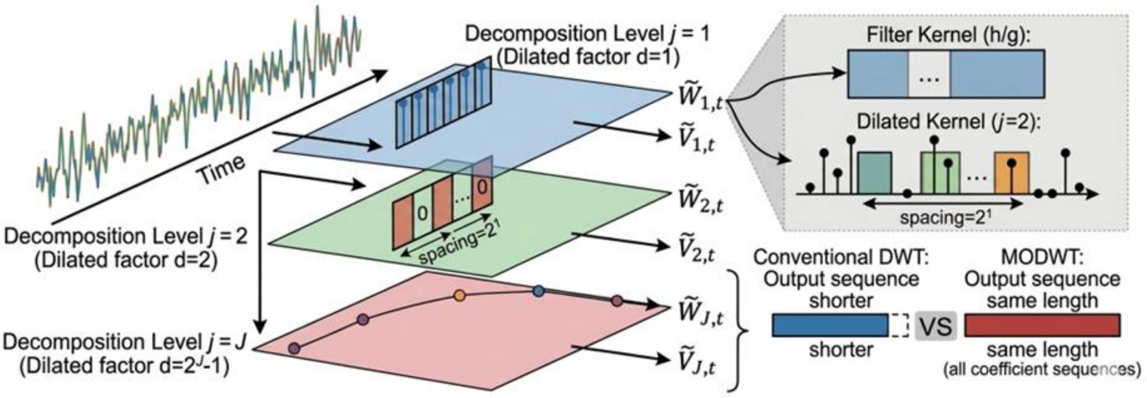

Fig. 2. Illustration of MODWT-based multi-scale decomposition using dilated convolution.

At decomposition level $j = 1, \dots, J$, MODWT applies dilated convolution with dilation factor 2^{j-1}:

$$\widetilde{W}_{j,t} = \sum_{l=0}^{K-1} g[l]\, x_{t-l\cdot 2^{j-1}} \quad (3)$$

$$\widetilde{V}_{j,t} = \sum_{l=0}^{K-1} h[l]\, x_{t-l\cdot 2^{j-1}} \quad (4)$$

Here, $\widetilde{W}j, t$ and $\widetilde{V}j, t$ represent the detail and approximation coefficients at level j, respectively.

Unlike the conventional DWT, MODWT does not perform downsampling, ensuring that all coefficient sequences have the same length as the original signal.

3) Multi-Scale Feature Construction

Starting from $A_t^{(0)} = x_t$, the decomposition is computed recursively using dilated convolutions with "same" padding to preserve sequence length.

The final multi-scale feature vector at each time step is constructed by channel-wise concatenation:

$$z_t = \left[A_t^{(J)}, D_t^{(J)}, \dots, D_t^{(1)}\right] \in R^{J+1} \qquad (5)$$

For a univariate input $F = 1$ and $J = 2$, the resulting feature tensor has shape $[B, T, 3]$.

D. Dual-Branch MODWT-LSTM Architecture

The proposed model adopts a dual-branch structure to jointly exploit raw and multi-scale representations.

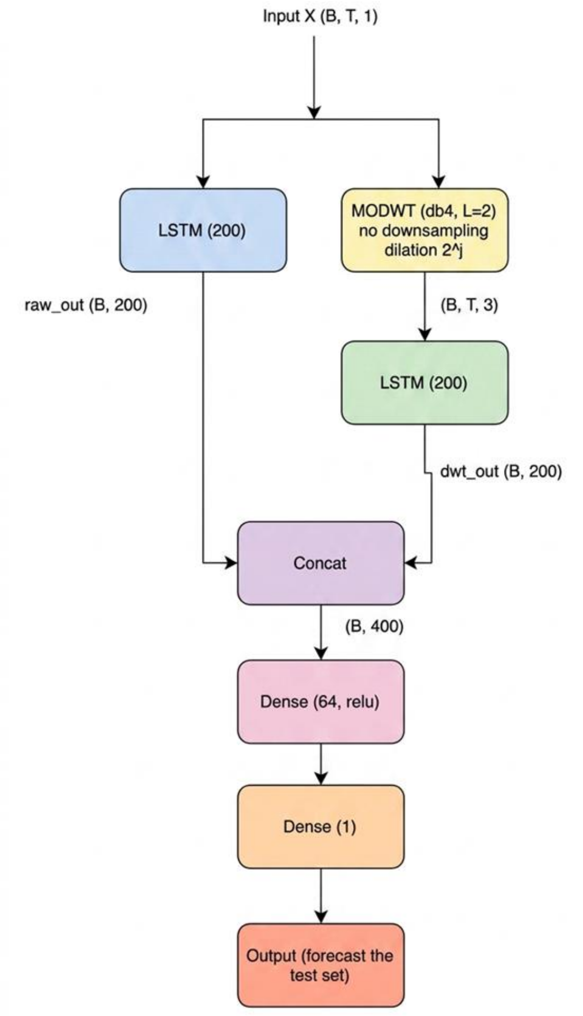

Fig. 3. Architecture of the proposed MODWT-LSTM model

1) Raw Signal Branch

The original input sequence $X \in R^{B\times T\times 1}$ is processed by an LSTM layer:
$h_{raw} = LSTM_{200}(X) \in R^{B\times 200}$ (6)
This branch captures nonlinear temporal dependencies directly from the raw signal.

2) MODWT Branch

The input sequence is first transformed using MODWT to obtain multi-scale features:
$Z = MODWT(X) \in R^{B\times T\times (J+1)}$ (7)
These features are then processed by another LSTM layer:
$hmodwt = LSTM200(Z) \in R^{B\times 200}$ (8)

3) Feature Fusion and Prediction Head

The outputs of the two branches are concatenated:
$h = Concat(hraw, hmodwt) \in R^{B\times 400}$ (9)
The fused representation is then passed through a fully connected head:
$\hat{x}t + 1 = Dense1\left(ReLU\left(Dense_{64}(h)\right)\right)$ (10)

This design corresponds to a late-fusion strategy, allowing the model to learn complementary representations before final prediction.

E. Implementation Details

The MODWT operation is implemented using dilated one-dimensional convolutions with “same” padding to preserve temporal alignment and ensure that all coefficient sequences have the same length as the input signal. The wavelet filters are fixed (non-trainable), providing a stable and interpretable multi-scale decomposition.

All models are trained under a unified experimental setting to ensure fair comparison. As summarized in Table 1, the dataset consists of hourly electricity demand data, with a look-back window of 24 hours and a one-hour forecasting horizon.

The training configuration is detailed in Table 2. All models are optimized using the Adam optimizer with Mean Squared Error (MSE) loss. To improve training stability and prevent overfitting, early stopping with a patience of 60 epochs and a ReduceLROnPlateau scheduler (factor = 0.98, patience = 4, minimum learning rate = 1e-4) are applied.

Table 3 summarizes the model-specific hyperparameters. For the proposed MODWT-LSTM model, both the raw and MODWT branches use LSTM layers with 200 hidden units. The MODWT decomposition is performed using the db4 wavelet with two decomposition levels. The extracted features are fused via concatenation and passed through a fully connected head consisting of a Dense layer with 64 units followed by the output layer.

For baseline models, consistent configurations are adopted to ensure comparability. In particular, RNN, LSTM, and GRU models use 64 hidden units, while CNN-based hybrid models employ convolutional feature extraction followed by recurrent layers. All models use a similar prediction head consisting of a dropout layer and a fully connected output layer.

Table 1. Dataset & Experimental Setup

Item	Value
Sampling frequency	1 hour
Target	electricity_demand
Input window (look-back)	24 hours
Forecast horizon	1 hour (hour-ahead)
Total output windows	69,014 windows
Train / Val / Test (windows)	55,211 / 6,901 / 6,902
Approx. days (Train / Val / Test)	2300.5 / 287.5 / 287.6 days
Scaler	Min-Max scaling
Missing values	dropna() (rows with NaN removed)
Framework	Keras/TensorFlow

Table 2. Summary of Hyperparameters

For all methods:

Optimizer	Adam (lr=0.001 default)
LR schedule	ReduceLROnPlateau (0.98, 4, min 1e-4)
Early stopping	Patience 60
Max epochs	200
Batch size	48
Look-back	24
Horizon	1
Loss	MSE

Table 3. Model-specific hyperparameters

Method	Hyperparameter	Value
MODWT-LSTM (ours)	Hidden units (raw branch)	200 (LSTM)
	Hidden units (MODWT branch)	200 (LSTM)
	MODWT levels / wavelet	2 / db4 (fixed, no trainable params)
	Fusion	Concatenate + Dense(64) → Dense(1)
	Optimizer	Adam
	Initial learn rate	0.001
	LR schedule	ReduceLROnPlateau (factor 0.98, patience 4, min LR 1e-4)
	Gradient threshold	clipnorm = 20.0
RNN	Hidden units	64 (SimpleRNN)
	Head	Dropout(0.2) → Dense(1)
LSTM	Hidden units	64 (LSTM)
	Head	Dropout(0.2) → Dense(1)
GRU	Hidden units	64 (GRU)
	Head	Dropout(0.2) → Dense(1)
CNN-GRU	CNN branch	Conv1D(64, kernel=3, ReLU) → MaxPool1D(2) → Flatten
	Recurrent branch	GRU-128
	Fusion	Concatenate → Dense(64) → Dense(1)

CNN-LSTM	CNN branch	Conv1D(64, kernel=3, ReLU) → MaxPool1D(2) → Flatten
	Recurrent branch	LSTM-128
	Fusion	Concatenate → Dense(64, ReLU) → Dense(1)

III. Results And Discussion

Metrics

To evaluate the performance of the proposed model, several widely used baseline models in short-term load forecasting are considered for comparison, including standalone CNN and LSTM models, as well as hybrid architectures such as CNN-GRU and CNN-LSTM.

The model performance is assessed using three common error metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE), defined as follows:

$$\mathbf{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - y_i^*| \quad (11)$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - y_i^*)^2} \quad (12)$$

$$\text{MAPE} = \frac{100\%}{n}\sum_{i=1}^{n} \left|\frac{y_i - y_i^*}{y_i}\right| \quad (13)$$

Results

1) Overall Performance on the Test Set

Table 4 presents the forecasting performance of all models on the full test set. The results show that the proposed MODWT-LSTM model significantly outperforms all baseline methods across all evaluation metrics.

In particular, the proposed model achieves the lowest MAE, RMSE, and MAPE, with a substantial margin compared to the best-performing baseline (CNN-LSTM). The MAPE is reduced from 3.84% (CNN-LSTM) to 2.25%, indicating a significant improvement in prediction accuracy. Similar trends are observed for MAE and RMSE, confirming the robustness of the proposed approach.

These results demonstrate that incorporating multi-scale representations via MODWT effectively enhances the model's ability to capture complex temporal dynamics in electricity load data.

Table 4. Performance of the implemented methods on the full test set (hour-ahead)

Method	MAE	RMSE	MAPE (%)
RNN	19.32	31.99	5.37
LSTM	18.25	29.41	5.27
GRU	15.81	27.56	4.52
CNN-GRU	15.01	26.17	4.19
CNN-LSTM	13.91	24.38	3.84
MODWT-LSTM (ours)	7.95	18.05	2.25

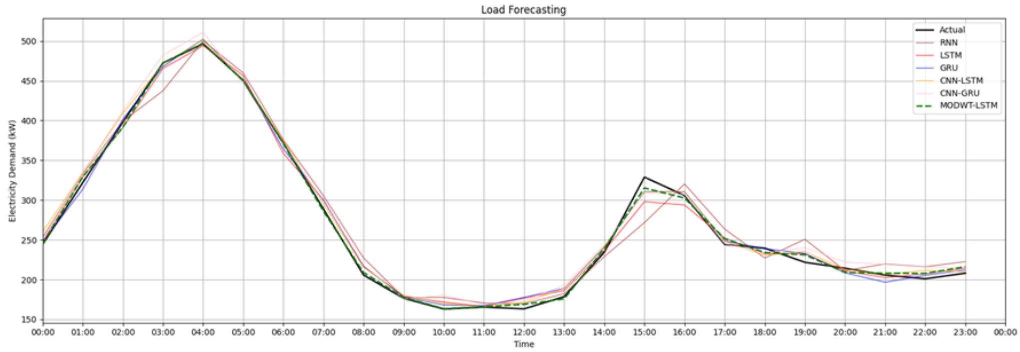

Fig.4. Hour-ahead load forecasting results on the full test set.

2) Performance under Different Temporal Conditions

To further evaluate model robustness, Table 5 reports performance across different calendar categories, including weekdays, special days, and the best-performing day.

On weekdays, where load patterns are relatively stable, all models perform reasonably well. Nevertheless, the proposed MODWT-LSTM model achieves the lowest error across all metrics, indicating superior modeling of regular temporal patterns.

On special days, which involve irregular consumption behaviors such as holidays, the forecasting task becomes more challenging. As expected, all models exhibit increased errors. However, the proposed model maintains a clear advantage, achieving significantly lower MAPE compared to the baselines. This suggests that the multi-scale representation helps improve generalization under non-stationary conditions.

Table 5. Performance of the implemented methods on calendar categories (hour-ahead)

Method	Weekdays			Special days			Best day		
	MAE	RMSE	MAPE (%)	MAE	RMSE	MAPE (%)	MAE	RMSE	MAPE (%)
RNN	18.95	29.54	5.23	21.41	32.63	6.19	22.37	26.12	3.91
LSTM	17.91	27.41	5.16	18.52	29.15	5.23	21.63	25.91	3.46
GRU	15.45	25.10	4.38	17.83	29.38	5.09	20.98	25.04	3.02
CNN-GRU	14.69	23.77	4.08	16.10	26.29	4.60	15.40	19.77	1.89
CNN-LSTM	13.62	22.68	3.73	15.22	26.32	4.32	17.52	21.99	1.72
MODWT-LSTM (ours)	6.77	15.09	2.02	11.08	22.24	2.90	3.23	4.37	0.34

Special days: The set above includes fixed national holidays, "Happy Monday" holidays (observed on the 2nd or 3rd Monday), the vernal/autumnal equinoxes, New Year's holidays (Jan 1–3), Obon (Aug 13–16), and extra Golden Week days (Apr 30, May 2, May 6).

Figs. 5–7 provide detailed comparisons for weekdays, special days, and the best-performing day.

For weekdays, the proposed model closely tracks the actual load profile across multiple days, with reduced deviation during both peak and off-peak periods. For special days, although prediction errors increase due to irregular patterns, the proposed model still captures the overall structure more accurately than competing methods.

Notably, on the best-performing day, the MODWT-LSTM model achieves near-perfect alignment with the ground truth, resulting in extremely low forecasting error. This demonstrates the model's ability to fully exploit structured temporal patterns when they are present.

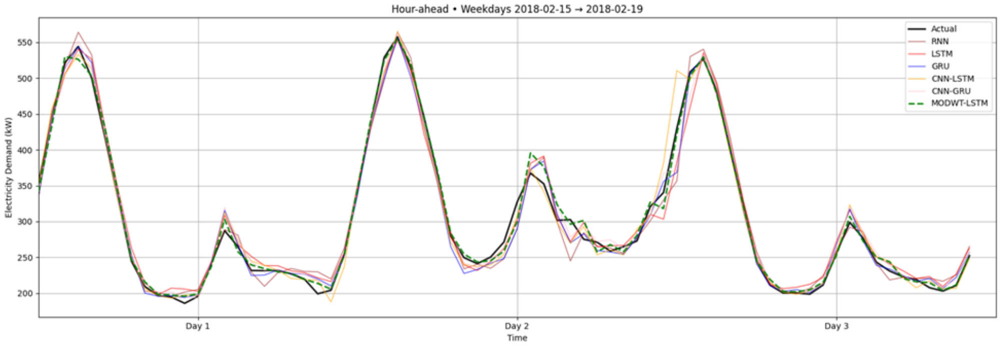

Fig. 5. Hour-ahead load forecasting results on weekdays.

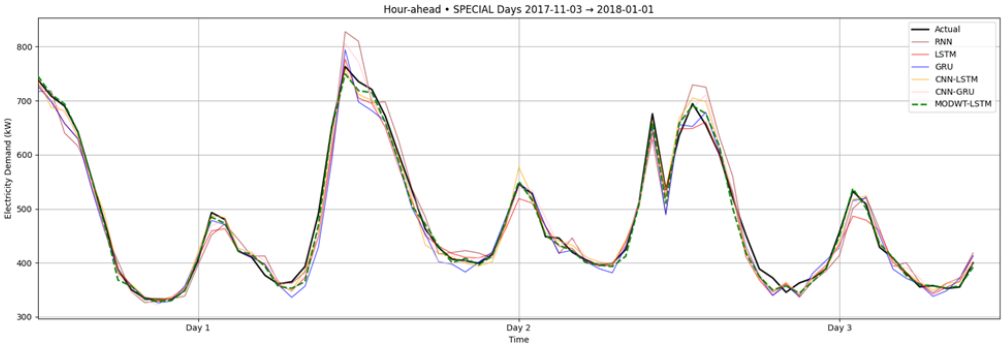

Fig. 6. Hour-ahead load forecasting results on special days.

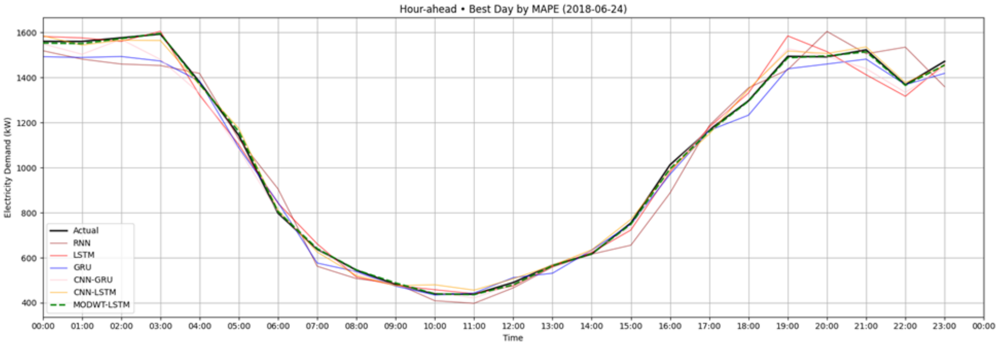

Fig. 7. Hour-ahead load forecasting results on the best-performing day (lowest MAPE).

3) Ablation Study

To evaluate the contribution of each component in the proposed architecture, an ablation study is conducted, as summarized in Table 6.

The results show that the Raw-LSTM baseline achieves a MAPE of 4.64%, while the MODWT-only model performs worse, with a MAPE of 5.00%. This indicates that applying MODWT alone, without jointly modeling the raw signal, is insufficient for accurate forecasting.

In contrast, the proposed dual-branch MODWT-LSTM model (Concat L2) significantly improves performance, reducing the MAPE to 3.38%. Similar improvements are observed in MAE and RMSE.

These findings confirm that multi-scale features extracted by MODWT are not effective in isolation but become highly beneficial when combined with time-domain representations. The results highlight the complementary nature of frequency-domain and temporal features in load forecasting.

Table 6. Ablation study of MODWT–LSTM variants on the test set

Variant	MAE	RMSE	MAPE (%)
Raw-LSTM	15.195	25.941	4.64
MODWT-Only L2	17.042	27.070	5.00
Concat L2 (ours)	11.632	21.901	3.38

4) Statistical Significance Analysis

To further validate the performance improvement, the Diebold–Mariano (DM) test is conducted to compare the proposed MODWT-LSTM model with the strongest baseline, CNN-LSTM.

As shown in Table 7, the DM statistic is -5.710 with a two-sided p-value of 0.0000, indicating a statistically significant difference in predictive accuracy.

A negative DM value implies that the proposed model achieves lower forecasting error than the baseline. Moreover, the large magnitude of the DM statistic, together with the near-zero p-value, suggests that the observed improvement is highly unlikely to be due to random variation.

Therefore, the null hypothesis of equal predictive performance is rejected, confirming that the proposed MODWT-LSTM model significantly outperforms the CNN-LSTM model in hour-ahead load forecasting under squared-error loss.

Table 7. Diebold–Mariano (DM) test comparing MODWT–LSTM against the strongest baseline

Best baseline	CNN-LSTM
Loss	SE
h	1
DM*	-5.710
df	10352
Two-sided p	0.0000
One-sided p (A<B)	0.0000

IV. Conclusion

This paper proposes a novel MODWT-LSTM framework for short-term load forecasting, which integrates maximal overlap discrete wavelet transform (MODWT) with a dual-branch LSTM architecture. By combining raw time-domain signals with multi-scale frequency-domain representations, the proposed model effectively captures both long-term dependencies and short-term fluctuations in electricity demand.

Experimental results on a real-world dataset demonstrate that the proposed approach consistently outperforms conventional RNN-based and CNN-based models across multiple evaluation metrics. In particular, significant improvements are observed in MAE, RMSE, and MAPE on the full test set, as well as under different temporal conditions, including weekdays and special days. The ablation study further confirms that the integration of MODWT and raw signal modeling is essential for achieving superior performance, while the Diebold–Mariano test verifies that the improvements are statistically significant.

Overall, the results highlight the effectiveness of multi-scale decomposition combined with deep learning for load forecasting tasks. The proposed MODWT-LSTM model provides a robust and interpretable framework that can be extended to other time-series forecasting problems.

Future work will focus on incorporating additional exogenous variables, such as weather and economic indicators, as well as exploring advanced architectures (e.g., attention mechanisms or transformer-based models) to further enhance forecasting accuracy.

References

[1]. Y. Liu, X. Zhang, And H. Wang, "Short-Term Load Forecasting Model Based On Lstm And Optimization Algorithms," Applied Sciences, Vol. 14, No. 14, P. 5966, 2024.

[2]. Q. Dong, Y. Chen, And Z. Li, "Short-Term Electricity Load Forecasting By Deep Learning: A Survey," Arxiv Preprint Arxiv:2408.16202, 2024.

[3]. H. Li, J. Wu, And Y. Sun, "Short-Term Power Load Forecasting For Integrated Energy Systems Using Hybrid Deep Learning," Frontiers In Energy Research, Vol. 12, 2024.

[4]. I. Westphal, M. Schmidt, And T. Müller, "Short-Term Load Forecasting With Deep Learning Methods," Applied Thermal Engineering, Vol. 240, 2025.

[5]. Y. Mu, Q. Li, And X. Zhou, "A Short-Term Load Forecasting Method Considering Multiple Factors," Expert Systems With Applications, Vol. 245, 2025.

[6]. P. Li, H. Zhao, And J. Wang, "Short-Term Electricity Load Forecasting With External Features And Gpt-Based Refinement," Sustainable Energy, Grids And Networks, Vol. 40, 2025.

[7]. Z. Yang, L. Chen, And Y. Wang, "Hybrid Cnn-Bilstm Model For Very Short-Term Load Forecasting," Energy Sources, Part A, 2025.

[8]. K. Li, X. Zhang, And Y. Liu, "Short-Term Load Forecasting Using Vmd-Based Decomposition," Energies, Vol. 18, No. 23, P. 6097, 2025.

[9]. Y. Liu, J. Zhang, And H. Xu, "Transformer-Based Short-Term Load Forecasting Model," Electronics, Vol. 14, No. 2, P. 230, 2025.

[10]. L. Zhu, X. Chen, And Y. Li, "Spatial–Temporal Dynamic Graph Transformer For Load Forecasting," Journal Of Computational Design And Engineering, Vol. 12, No. 2, Pp. 92–105, 2025.

[11]. H. J. Bae, S. Kim, And J. Lee, "Short-Term Load Forecasting With Clustering And Dimensionality Reduction," Scientific Reports, Vol. 15, 2025.

[12]. A. Jahani, M. Shafie-Khah, And J. P. S. Catalão, "Hybrid Spm-Lstm Model For Microgrid Load Forecasting," Sustainable Cities And Society, Vol. 95, 2023.

[13]. Q. Huang, S. Li, And M. Shahidehpour, "Improved Cnn-Based Probabilistic Load Forecasting Method," Energy, Vol. 209, 2020.

[14]. J. Chen, Y. Wang, And Z. Li, "Hybrid Vmd-Lstm Model For Short-Term Load Forecasting," Applied Sciences, Vol. 14, No. 14, 2024.

[15]. K. Kondaiah And P. S. Reddy, "A Review On Short-Term Load Forecasting Models," Iet Generation, Transmission & Distribution, Vol. 16, No. 10, Pp. 2100–2115, 2022.

[16]. F. Rodrigues, C. Cardeira, And J. M. F. Calado, "A Systematic Review Of Short-Term Load Forecasting Methods," Preprints, 2023.

[17]. Q. Dong, Y. Chen, And Z. Li, "Deep Learning-Based Short-Term Load Forecasting: Methods And Challenges," Expert Systems With Applications, Vol. 250, 2025.

[18]. J. Kim, H. Park, And S. Lee, "Gru-Based Very Short-Term Load Forecasting With Attention Mechanism," Energies, Vol. 18, No. 13, P. 3229, 2025.

[19]. I. Beloev, D. Ivanov, And K. Georgiev, "Short-Term Load Forecasting Using Xgboost," Energies, Vol. 18, No. 19, P. 5144, 2025.

[20]. B. Çadırcı, M. Demir, And A. Korkmaz, "Short-Term Load Forecasting Considering Seasonality Effects," 2025.