▪ Journal of Competitiveness

Day-ahead electricity price forecasting considering the scaled integration of renewable energy: A fusion approach based on ICEEMDAN and iTransformer

Puliang Du, Miaoheng Yang, Fei Xie, Wei Hu, Jiyu Li

Abstract

As the global electricity market continues to evolve, day-ahead electricity price forecasting has become increasingly important for decision-making among various market entities. However, the continuous integration of high proportions of clean energy poses significant challenges for accurate day-ahead electricity price predictions. In response, we fully considered the coupling relationship between the output characteristics of renewable energy and the multidimensional features of electricity prices and proposed a day-ahead electricity price forecasting model based on improved complete ensemble empirical mode decomposition with adaptive noise (ICEEMDAN) and iTransformer. First, we decompose historical electricity price data using the ICEEMDAN method to obtain multidimensional time series data based on intrinsic mode functions (IMFs). Second, we leverage the attention mechanism in iTransformer to independently predict the multidimensional time-series data containing IMFs and renewable energy output, forming a forecasting model suited for the large-scale integration of renewable energy in the electricity market. Finally, using historical electricity price and renewable energy output data from Spain as a case study, we constructed a simulation model. The results demonstrate that the ICEEMDAN-iTransformer model effectively handles noise, nonlinearity, and non-smoothness in data following the integration of renewable energy, enabling more stable and accurate forecasting results.

Keywords: *Electricity price forecasting; Renewable energy; ICEEMDAN; iTransformer*

JEL Classification: C32, L11

Article history: Received: November 2024; Accepted: March 2025; Published: March 2025

1 INTRODUCTION

Electricity prices, serving as the cornerstone economic signals in power markets, exert significant influence on the operational decisions of generators, retailers, and end-users, while simultaneously shaping macroeconomic stability and power system equilibrium (Darandary et al., 2024; Enrich et al., 2024). However, the accelerating global deployment of renewable energy has introduced unprecedented challenges to conventional forecasting methodologies, primarily due to the inherent stochasticity and intermittency of renewable generation (Özen & Yildirim, 2021; Tan et al., 2023). This necessitates the development of adaptive forecasting frameworks capable of addressing the complex dynamics induced by high renewable penetration.

Current electricity price forecasting methodologies are typically categorized into day-ahead, intraday, and real-time predictions (Chen et al., 2025). Given the dominance of day-ahead markets in electricity trading, this paper focuses on day-ahead price forecasting as the critical domain for addressing renewable integration challenges. Recent advancements in this field have incorporated multidimensional factors such as price volatility analysis (Grothe et al., 2023), spatiotemporal correlation modeling (Meng et al., 2024), and underlying predictive indicators (Özen & Yildirim, 2021). Notwithstanding these developments, traditional single-model approaches have proven inadequate for capturing the multi-temporal and spatial characteristics of electricity prices, failing to meet the precision requirements of modern power markets (Guan et al., 2022; Luo et al., 2024). The proliferation of artificial intelligence has revolutionized electricity price prediction, yielding remarkable performance improvements (González-Tejero et al., 2023; Méndez-Suárez et al., 2023). For instance, Lehna et al. (2022) conducted a comprehensive comparative analysis of four forecasting frameworks in the German day-ahead market, demonstrating that hybrid CNN-LSTM architectures and expanded two-stage multivariate VAR models significantly enhance prediction accuracy.

Notwithstanding the ongoing evolution in electricity price forecasting research, there exists a demonstrable correlation between predictive model efficacy and input variable characteristics, underscoring the necessity for stringent feature selection protocols in forecasting frameworks (Guo et al., 2025). To address this gap, Pourdaryaei et al. (2024) introduced a novel hybrid forecasting framework integrating multi-head attention mechanisms with convolutional neural networks (CNNs). Concurrently, a novel feature selection framework was proposed, leveraging mutual information theory and neural networks to systematically filter input variables relevant to electricity price forecasting.

While advancements in prediction models and improved data quality have significantly enhanced forecasting accuracy, the rising penetration of renewable energy introduces new challenges arising from their inherent stochasticity and uncertainty, which may undermine predictive reliability. Existing research highlights that transmission system operator (TSO) forecasts of total electricity load and variable renewable generation (wind/solar) – critical inputs for market operations – are prone to systematic biases. Maciejowska et al. (2021) demonstrated that autoregressive modeling techniques can improve these forecasts, enabling preemptive market condition anticipation and potentially increasing revenue. This underscores the substantial impact of renewable generation data on day-ahead pricing dynamics. Concurrently, Davis and Brear (2024) revealed that short-term wind power forecasting errors correlate with increased unutilized energy hours and scarcity pricing events, leading to higher system costs and market prices. This unintended consequence exacerbates electricity price volatility, reduces wind turbine profitability, and marginally increases greenhouse gas emissions. The urgency to improve short-term wind forecasting intensifies when variable renewable generation exceeds 60% of annual electricity demand.

Addressing the challenges posed by high-penetration renewable energy systems requires strategic input variable selection and advanced feature extraction as critical methodological priorities. In response, Yao et al. (2020) developed a wind-dominated market forecasting framework that integrates multi-domain features with long short-term memory (LSTM) networks. This architecture leverages the LSTM’s temporal dependency modeling capabilities to improve forecasting accuracy in high-wind scenarios, particularly when incorporating wind power-to-load ratios and multi-timescale inputs. Building on this foundation, Yin et al. (2022) proposed a singular spectrum analysis (SSA)-LSTM-cross-entropy optimization (CSO) hybrid model to decompose historical price data into trend, periodic, and residual components. While effective for trend extraction, SSA fails to capture high-dimensional latent features. Xu et al. (2024) introduced a variable mode decomposition (VMD)-grey wolf optimization (GWO)-attention-LSTM framework, demonstrating that VMD contributes most significantly to performance. However, VMD’s sensitivity to noise often leads to modal distortion in tariff signals. To mitigate this, Chen et al. (2025) integrated local outlier factor (LOF) detection with VMD, improving decomposition robustness. More recently, Xu Yue et al. (2024) combined maximum information coefficient (MIC) with ensemble empirical mode decomposition (EEMD) and an enhanced informer model to uncover multidimensional latent features, achieving state-of-the-art precision. Liu et al. (2024) further advanced EEMD into improved complete ensemble empirical mode decomposition with adaptive noise (ICEEMDAN), enabling nonlinear and nonstationary pattern extraction from raw price series.

Notwithstanding the academic recognition garnered by the aforementioned studies, electricity price forecasting under high renewable penetration remains hindered by three fundamental challenges: the high-dimensional complexity of historical operational data, the cross-scenario generalizability limitations of forecasting models, and the technical barriers in extracting discriminative features from stochastic renewable generation dynamics. With the continuous advancement of artificial intelligence and computational capabilities, the transformer model has emerged as a promising solution due to its parallel processing efficiency, scalability, adaptability, and robust generalization capacity. Kaya et al. (2023) pioneered the application of a transformer encoder-decoder with self-attention (TEDSE) for electricity price forecasting. However, conventional transformer architectures exhibit limitations in multivariate long-term time series prediction, yielding suboptimal results (Huang et al., 2023; Nie et al., 2023). To address these limitations, Liu et al. (2024) introduced the i-Transformer, which revolutionizes architecture design through inverted structural optimization. This innovation preserves the model's inherent strengths while enhancing its ability to process multivariate long-term time series data (Zhao et al., 2017). This advancement provides novel opportunities for precise electricity price forecasting in large-scale renewable integration scenarios (Ramírez et al., 2024).

Consequently, this paper introduces an innovative day-ahead electricity price forecasting framework that integrates ICEEMDAN with the iTransformer. The contributions of this study are threefold: First, we address the nonlinear and nonstationary characteristics of electricity price data by applying ICEEMDAN to decompose raw time series into intrinsic mode functions (IMFs) (Colominas et al., 2014; Huang et al., 2022). These IMFs effectively capture the intrinsic dynamics of price fluctuations, including cyclical patterns and transient shocks (Yang et al., 2024; Ghimire et al., 2022). Second, we propose a multivariate correlation analysis using the iTransformer architecture to model cross-dimensional dependencies among decomposed IMFs and renewable generation data. This approach enables the identification of complex spatiotemporal patterns and long-range temporal correlations in electricity prices. Third, we develop a serialized prediction framework that forecasts multi-dimensional features of day-ahead prices by inputting individual IMFs and renewable energy generation data. Predictions are aggregated through a weighted summation method to produce the final electricity price forecast. To validate the model's efficacy, we conducted comprehensive simulations using the 2021 Spanish electricity price dataset, comparing our framework against state-of-the-art forecasting models.

The study consists of five sections. Following the introduction, we present the theoretical framework in Section II. We then discuss the data in Section III. Section IV presents the results of the study and assesses the generalizability and validity of the model. Conclusions come in Section V.

2 THEORETICAL FRAMEWORK

2.1 Theoretical research process

This section describes the relevant methods and theories of the proposed day-ahead electricity price forecasting framework in a sequential manner. The overall flowchart in this paper is shown in Figure 1. This method achieves accurate day-head electricity prices prediction through five key components: (a) Data preprocessing, (b) Model training and optimization, (c) Integration of the forecasting values, (d) Evaluation of prediction results, and (e) Model comparison and comparative experiments. (f) Model validation techniques. The specific workflow is as follows.

(a) Data preprocessing

The raw dataset from Valencia, Spain, with day-ahead electricity prices per hour and 11 renewable energy generation data in 2022, are collected and subjected to data cleaning and processing. An adaptive ICEEMDAN decomposition method is used to analyze the non-smooth signals generated by day-ahead electricity prices in the raw data, while residual noise and spurious modes are addressed. In our work, the original electricity price is decomposed into a residual series and 5 relatively stable IMFs, which are denoted by IMF1-IMF5 from high-frequency to low-frequency. To ensure that potentially useful information is not lost, a residual analysis is performed and found that its effect on the prediction results was negligible. Therefore, the residual component is chosen to ignore.

(b) Model training and optimization

Having discussed the data preprocessing steps, we now turn to the model training and optimization process. The residual series is eliminated because it does not contain useful information for forecasting. The processed electricity price data as target variate is fed into the iTransformer model through layer normalization (see Section 2.2.2 for a detailed introduction to iTransformer). Technically, we embed each renewable energy generation time series as variate tokens, adopt the attention for multivariate correlations, and employ the feed-forward network for series representations, which learn better series-global representations for time series forecasting. The model is evaluated based on the convergence of the loss function using the validation set, and the optimal model is selected.

(c) Integration of the forecasting values

Day-head electricity price is aggregated from the predicted values of IMF1-IMFk and can be defined as：

$$\hat{Y} = w_1 \times I\hat{M}F1 + w_2 \times I\hat{M}F2 + \ldots + w_k \times I\hat{M}Fk \tag{1}$$

After the ICEEMDAN decomposition, we obtain a series of $IMFs$ ($IMF1$, $IMF2$, ..., $IMFk$). When it is necessary to integrate these $IMFs$ for further analysis or processing, a simple and straightforward approach equal weighting is implemented, which means that each IMF is considered to have the same importance and impact in the integration process (Chatterjee & Chakraborty, 2024).

(d) Evaluation of prediction results

The selected model is applied to the test dataset, and its performance is assessed using three evaluation metrics and two-time efficiency values.

(e) Model comparison and comparative experiments

Our proposed model ICEEMDAN-iTransformer is used for multivariate characterization of renewable energy generation to predict day-ahead electricity price, and in order to better validate the advantages of the model, we choose the former-based model and their invert former-based model to compare with iTransformer. For example, Flowformer\Informer\Flashformer as a variant of transformer, mainly concerns the component adaptation, especially the attention module for the temporal dependency modeling and the complexity optimization on long sequences. Furthermore, Reformer pays more attention to the inherent processing of time series, which bring about consistently improved performance. Also, in comparison experiment I, the prediction of single-target day-ahead electricity prices by iTransformer demonstrates the advantages of the model. In comparison experiment II, taking into account the effect of multivariate renewable energy generation on the single target electricity price prediction and make projections, which improves the prediction results of iTransformer. Also in experiment II, we considered other features such as weather policies and analyzed them.

(f) Model validation techniques

Through interpretability analysis, cross-validation, and outlier sensitivity testing, the ICEEMDAN-iTransformer model excels in accuracy, probabilistic forecasting, and handling extreme price fluctuations, the model proves its reliability and generalizability, making it a powerful tool for electricity market participants and grid operators in managing price volatility and optimizing energy scheduling.

2.2 Theoretical Methodology

2.2.1. ICEEMDAN

Define y as the original variable to be decomposed, which in this study means the day-ahead electricity price. For the residual noise in modes, let $\bar{M}(\cdot)$ be the operator which produces the local mean of the signal that is applied and let $w^{(i)}$ be a realization of white Gaussian noise with zero mean and unit variance. For spurious modes, the use of $\hat{E}_k\left(w^{(i)}\right)$ to extract the kth mode instead of white noise directly is suggested. Note that $\langle\cdot\rangle$ refers to the averaging process. A flowchart of this new algorithm can be found in Figure 2.

First, a set of white Gaussian noise $w^{(i)}$ are added to the original day-ahead electricity price sequence y, the new sequence $y^{(i)}$ constructed is as follows:

$$y^{(i)} = y + \beta_0 \hat{E}(w^{(i)}) \tag{2}$$

where constants $\beta_k = \varepsilon_k std(\vec{R}_k)$ are chosen to obtain a desired SNR between the added noise and the residue to which the noise is added. And get the first set of residuals as follows:

$$\vec{R}_1 = \left\langle \bar{M}\left(y^{(i)}\right)\right\rangle \tag{3}$$

Then, the first modal component $IMF1$ is calculated:

$$d_1 = y - \vec{R}_1 \tag{4}$$

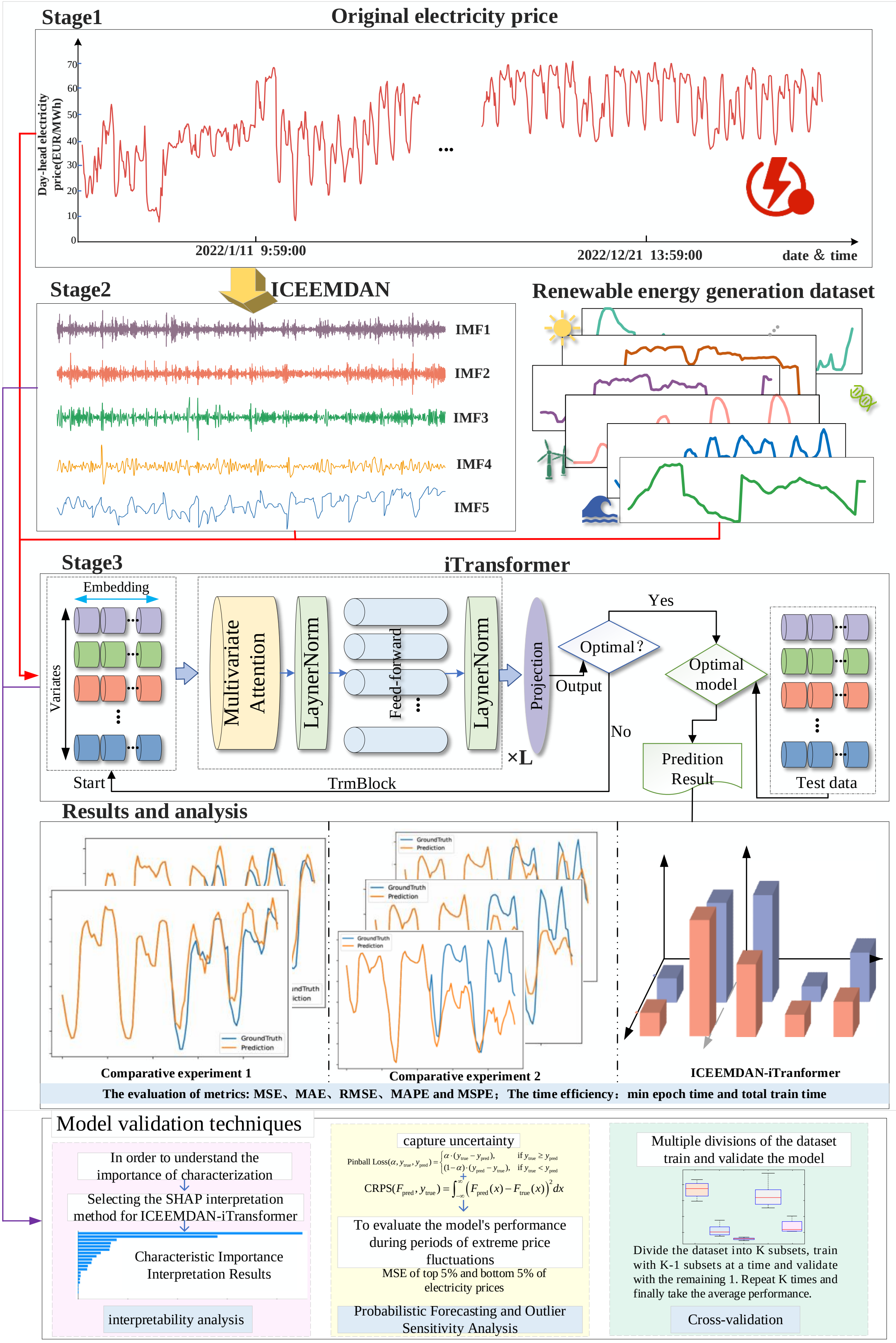


Fig. 1 – Workflow of theoretical research

continue to add white noise and compute the second set of residuals $\vec{R}_1 + \beta_1 \hat{E}(w^{(i)})$ using local mean decomposition. Calculate the second modal component $IMF2$ as follows:

$$d_2 = \vec{R}_1 - \vec{R}_2 = \vec{R}_1 - \left\langle \bar{M}(\vec{R}_1 + \beta_1 \hat{E}(w^{(i)})) \right\rangle \tag{5}$$

similarly, the kth residual $\vec{R}_k + \beta_k \hat{E}(w^{(i)})$ and modal components $IMFk$ are computed as follows:

$$d_k = \vec{R}_{k-1} - \vec{R}_k = \vec{R}_{k-1} - \left\langle \bar{M}(\vec{R}_{k-1} + \beta_{k-1} \hat{E}(w^{(i)})) \right\rangle \tag{6}$$

until the end of the computational decomposition, all modes and the number of residuals are obtained. Regarding the two problems mentioned above, the ICEEMDAN algorithm proceeds are as follows:

Algorithm 1 ICEEMDAN	
Require:	Signal $y(t)$, noise amplitude standard deviation sequence $\sigma(i)$, number of decomposition layers N, number of iterations M, intrinsic modal functions list $imfs$.
1:	▷ Initialize the list of remaining signals $res = y(t)$ and intrinsic modal functions list $imfs = []$.
2:	**for** i **in** $Range(1, N+1)$:
3:	▷ Initialize the noise signal for the first EMD decomposition: $noise = randn(size(y(t))) * \sigma(i)$ $signal\_with\_noise = res + noise$
4:	▷ Perform the EMD decomposition to get the first IMF: $imf1 = EMD(signal\_with\_noise)$
5:	**for** j **in** $Range(1, M+1)$:
6:	▷ Calculate the noise for the jth iteration: $noise\_j = randn(size(y(t))) * \sigma(i)$ $signal\_j = res + noise\_j$
7:	▷ EMD decomposition of the signal with added noise. $IMFs\_j = EMD(signal\_j)$
8:	▷ Calculate the mean IMF. $mean\_imf = mean(IMFs\_j[1])$
9:	▷ Use the average IMF to update the remaining signals $res = res - mean\_imf$
10:	▷ Add the resulting IMFs to the result list. $imfs.append(mean\_imf)$
11:	**Return** $imfs$

2.2.2 iTransformer

For day-ahead electricity price long-term series forecasting, we propose an elaborate inverted transformer, which adopts the encoder-only architecture of transformer (Vaswani et al., 2017), including the embedding, projection, and transformer blocks, as shown in Figure 3.

In multivariate time series forecasting, given historical observations $X=\{x_1,\dots,x_T\}\in R^{T\times N}$ with T time steps and N variates, we predict the future S time steps $Y=\{x_{T+1},\dots,x_{T+S}\}\in R^{S\times N}$ .For convenience, we denote $X_{t,:}$ as the simultaneously recorded time points at the step t, and $X_{:,n}$ as the whole time series of each variate indexed by n. The process of predicting future series of each specific variate $\hat{Y}_{:,n}$ based on the lookback series $X_{:,n}$ is simply formulated as follows:

$$h_n^0 = Embedding(X_{:,n}) \tag{7}$$

$$H^{l+1} = TrmBlock(H^l), l=0,\dots,L-1 \tag{8}$$

$$\hat{Y}_{:,n} = \Pr ojection(h_n^L) \tag{9}$$

where $H=\{h_1,\dots,h_N\}\in R^{N\times D}$ contains N embedded tokens of dimension D and the superscript denotes the layer index. $Embedding: R^T \mapsto R^D$ and $\Pr ojection: R^D \mapsto R^S$ are both implemented by multi-layer perceptron (MLP).

We organize a stack of L blocks composed of the layer normalization, feed-forward network, and self-attention modules (Hiew et al., 2023). In our inverted version, the normalization is applied to the series representation of individual variate as Eq. (9).

$$LayerNorm(H) = \left\{ \frac{h_n - Mean(h_n)}{\sqrt{Var(h_n)}} \middle| n=1,\dots,N \right\} \tag{10}$$

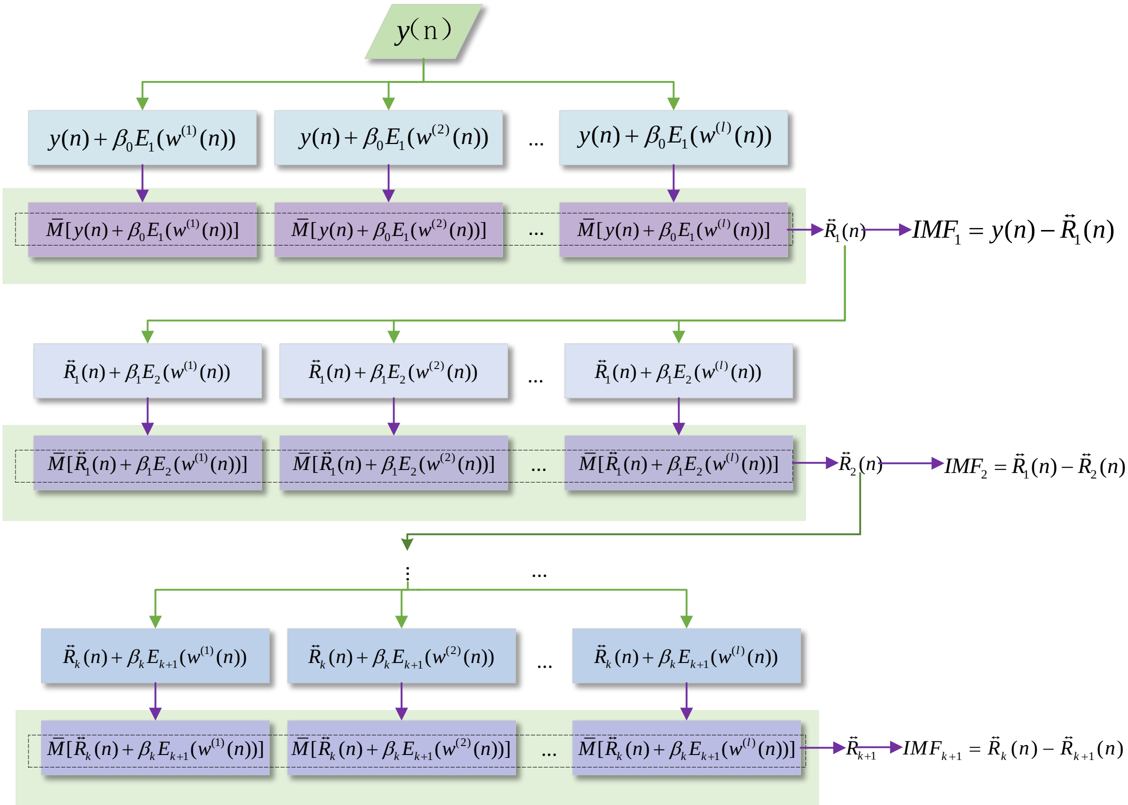


Fig. 2 – Flowchart describing the improved version of CEEMDAN

In the inverted version, FFN is leveraged on the series representation of each variate token, they can extract complicated representations to describe a time series. With comprehensively extracted representations of each time series $H = \{h_1, \ldots, h_N\} \in R^{N \times D}$, the inverted model regards the whole series of one variate as an independent process. The self-attention module adopts linear projections to get queries, keys, and values $Q, K, V \in R^{N \times d_k}$,where d_k is the projected dimension. With denotation of $q_i, k_j \in R^{d_k}$ as the specific query and key of one (variate) token, we notice that each entry of the pre-Softmax scores is formulated as:

$$A_{i,j} = (QK^{\mathrm{T}} / \sqrt{dk})_{i,j} \propto q_i^{\mathrm{T}} k_j \tag{11}$$

The whole score map $A \in R^{N \times N}$ exhibits the multivariate correlations between paired variate tokens, highly correlated variate will be more weighted for the next representation interaction with values V. The algorithm for iTransformer is as follows:

Algorithm 2 iTransformer -Overall Architecture.

Require: Input lookback time series $X \in R^{T \times N}$; Input Length T;predicted length S;variates number N;token dimension D; iTransformer block number L.

1:	$X = X.transpose$	▷ $X \in R^{N \times T}$
2:	▷Multi-layer Perceptron works on the last dimension to embed series into variate tokens.	
3:	$H^0 = MLP(X)$	▷ $H^0 \in R^{N \times D}$
4:	**For** l **in** $\{1, \ldots, L\}$:	▷Run through iTransformer blocks.
5:	▷ Self-attention layer is applied on variate tokens.	
6:	$H^{l-1} = LayerNorm(H^{l-1} + Self-Attn(H^{l-1}))$	▷ $H^{l-1} \in R^{N \times D}$
7:	▷ Feed-forward network is utilized for series representations, broadcasting to each token.	
8:	$H^l = LayerNorm(H^{l-1} + Feed-Forward(H^{l-1}))$	▷ $H^l \in R^{N \times D}$
9:	▷ LayerNorm is adopted on series representations to reduce variates discrepancies.	
10:	**End for**	
11:	$\hat{Y} = MLP(H^L)$	▷Project tokens back to predicted series, $\hat{Y} \in R^{N \times S}$
12:	$\hat{Y} = \hat{Y}.transpose$	▷ $\hat{Y} \in R^{S \times N}$
13:	**Return** $\hat{Y}$	▷ Return the prediction result $\hat{Y}$

2.3. Evaluation metrics

2.3.1 Residual test

In time series analysis, the ADF (augmented Dickey-Fuller) test and PACF (partial autocorrelation function) are two commonly used tools to test the smoothness and analyze the autocorrelation of a time series, respectively, and they are particularly important when analyzing residuals. The ADF test is based on the following regression model:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{i=1}^{p} \phi_i \Delta y_{t-i} + \check{n}_t \tag{12}$$

where y_t is the value of the time series at time t; $\Delta y_t = y_t - y_{t-1}$ is the first-order difference of the time series; α is the constant term; βt is the time-trend term; γ is the coefficient of the lag term, which is used to test for a unit root; ϕ_i is the coefficient of the difference lag term; and $\check{n}_t$ is the error term. The PACF test is based on the following regression model:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_k y_{t-k} + \check{n}_t \tag{13}$$

where ϕ_k is the partial autocorrelation coefficient of the kth lag term. The combined use of ADF and PACF allows a comprehensive assessment of whether the model' s residuals are consistent with the white noise assumption (i.e., the residuals are smooth and free of autocorrelation), and thus the validity of the model can be judged.

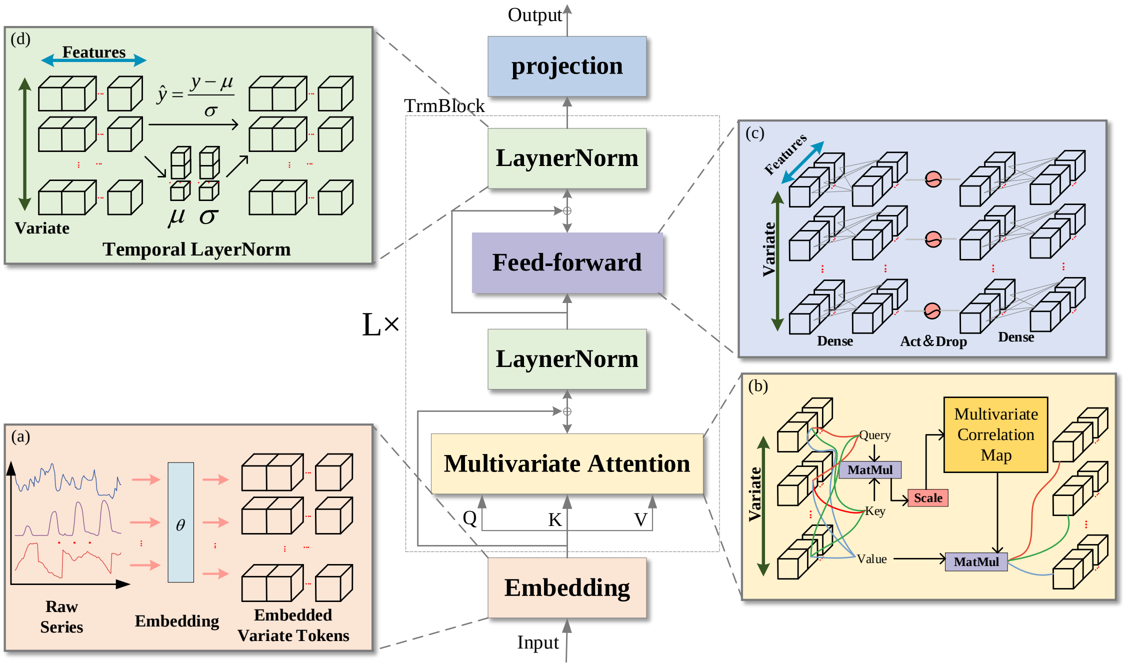


Fig. 3 – Overall structure of iTransformer

2.3.2 Predictive performance

In order to compare the predictive performance of the models, five different evaluation metrics are chosen in our work: mean square error (MSE), mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE), and mean squared percentage error (MSPE), and which both characterize the error between the predicted and actual values, with smaller values representing better model predictions. Given the true value $\vartheta_{d,h}$ and the predicted value $\hat{\vartheta}_{d,h}$, $d = 1, 2, \cdots, N_d$ represents days and $h = 1, 2, \cdots, 24$ represents hours. The formulas for each are shown below:

$$MSE = \frac{1}{24N_d} \sum_{d=1}^{N_d} \sum_{h=1}^{24} (\vartheta_{d,h} - \hat{\vartheta}_{d,h})^2 \tag{14}$$

also known as the L2 paradigm loss, it is averaged by summing the true value minus predicted value and then squaring it.

$$MAE = \frac{1}{24N_d}\sum_{d=1}^{N_d}\sum_{h=1}^{24}\left|\vartheta_{d,h} - \hat{\vartheta}_{d,h}\right| \tag{15}$$

also known as the L1 paradigm loss, this is the average of the absolute differences between the predicted and observed values.

$$RMSE = \sqrt{\frac{1}{24N_d}\sum_{d=1}^{N_d}\sum_{h=1}^{24}(\vartheta_{d,h} - \hat{\vartheta}_{d,h})^2} = \sqrt{MSE} \tag{16}$$

MSE becomes RMSE after the root sign, and the magnitude of the error result is on the same level as the original data, which makes it easier to describe the results of our prediction. The differences between RMSE and MAE are: for poorly predicted values, RMSE gives higher penalties than MAE; RMSE is smooth and differentiable, which makes it easier to perform mathematical operations, such as finding the gradient; and MAE is more robust to anomalous values. robustness. Therefore, the model is evaluated by combining the respective characteristics of MSE, MAE and RMSE metrics. MSPE and MAPE can indicate relative error preferences:

$$MAPE = \frac{1}{24N_d}\sum_{d=1}^{N_d}\sum_{h=1}^{24}\left|\frac{\vartheta_{d,h} - \hat{\vartheta}_{d,h}}{\vartheta_{d,h}}\right| \tag{17}$$

For each sample, the absolute error is divided by the target value and the MAPE can be considered a weighted version of the MAE. However, if the outlier has a very small value, the MAPE will be very biased in its favor because this outlier will have the highest weight.

$$MSPE = \frac{1}{24N_d}\sum_{d=1}^{N_d}\sum_{h=1}^{24}\left(\frac{\vartheta_{d,h} - \hat{\vartheta}_{d,h}}{\vartheta_{d,h}}\right)^2 \tag{18}$$

MSPE is considered as a weighted version of MSE with samples whose weights are inversely proportional to their true target squares. Based on the characteristics of MAPE, MSPE and data, it is only used for model evaluation for univariate tariff forecasting.

2.3.3 Probabilistic forecasting

The ICEEMDAN-iTransformer model to generate probabilistic forecasts by predicting multiple quantiles (e.g., 10th, 50th, and 90th percentiles) of the electricity price distribution. This allows us to construct confidence intervals and evaluate the model's ability to capture uncertainty. The quantile predictions were obtained by modifying the loss function to minimize the pinball loss, which is defined as

$$\text{Pinball Loss}(\alpha, y_{\text{true}}, y_{\text{pred}}) = \begin{cases} \alpha \cdot (y_{\text{true}} - y_{\text{pred}}), & \text{if } y_{\text{true}} \geq y_{\text{pred}} \\ (1-\alpha) \cdot (y_{\text{pred}} - y_{\text{true}}), & \text{if } y_{\text{true}} < y_{\text{pred}} \end{cases} \quad (19)$$

where α is the quantile level, y_{true} is the true value, and y_{pred} is the predicted quantile. The continuous ranked probability score (CRPS) was also calculated to measure the difference between the predicted and observed cumulative distribution functions:

$$\text{CRPS}(F_{\text{pred}}, y_{\text{true}}) = \int_{-\infty}^{\infty} \left(F_{\text{pred}}(x) - F_{\text{true}}(x) \right)^2 dx \quad (20)$$

where $F_{\text{pred}}(x)$ is the predicted cumulative distribution function and $F_{\text{true}}(x)$ is the true cumulative distribution function.

3 DATA ANALYSIS

This section familiarizes the reader with the data utilized, in particular the input characteristics and forecast targets. The European electricity market consists of derivatives, spot and balancing segments, the most important being the spot market, in particular the day-ahead auction (Zhu et al., 2024). It is held once a day at noon, and all products of the next day are traded in a uniform price auction. Since all hours of the next day are traded at the same time, all hours are based on the same set of information. In Europe, where electricity price levels vary considerably from country to country, the Spanish electricity market is dominated by renewable energy sources, particularly wind power, solar PV and solar thermal generation increasing year-on-year (Karahan et al., 2024). In addition, the government has announced plans to retire coal-fired power plants by 2025, oil-fired power plants by 2030, and nuclear power plants by the end of 2035 (Bonilla et al., 2022). With the retirement of these energy sources, the demand for renewable energy is expected to increase significantly to compensate for the retired generation resources while meeting the country's growing electricity demand and is likely to dominate the Spanish electricity market over the forecast period (Ciarreta et al., 2020). Therefore, considering the impact of renewable energy on electricity prices, we have selected the day-ahead electricity prices for the Spanish electricity market for forecasting.

On an hourly basis, the day-ahead spot electricity price data of the Spanish electricity market for the time period 2022/1/1 0:59 to 2022/12/31 23:59 is selected as our target variable, as well as the data for 11 renewable energy generation, including generation biomass, generation geothermal and another 9 renewable energy generation datasets, etc. The dataset spans one year and includes 8760 hourly data points for each of the 12 variables. Detailed day-ahead spot electricity price and renewable energy generation datasets information separately are provided in Appendix A and Appendix B. Since the spot electricity price in Spain before the day is affected by daylight saving time, which resulted in missing or discontinuous data, the missing values are interpolated using forward fill or backward fill based on the distributional characteristics of the data and the expert knowledge to estimate the most probable values of the missing values, in order to ensure the accuracy and consistency of the data. Then, the dataset is divided into training, validation and test sets with 8:1:1, 7008, 876 and 876 entries, respectively, to perform uni/multivariate day-ahead electricity price prediction in the long term.

Table 1 presents the descriptive statistics of electricity prices and renewable energy generation. The mean value of electricity price is 41.6325EUR/MWh; the standard deviation is 15.168, which means that the electricity price data fluctuates about 15EUR/MWh above and below the mean value, which is not a particularly large range of fluctuation compared to the mean value; the minimum electricity price is 2.3EUR/MWh, and the maximum price is 88.44 EUR/MWh; the kurtosis and skewness are -0.12509 and -0.1606, indicating that the distribution of electricity price is slightly left-skewed, i.e, low electricity price appears slightly more often than high electricity price, and the kurtosis of the electricity price distribution is slightly flat compared with the normal distribution, and the trend of electricity price is shown in Figure 4. And, from these data, it can be seen that different types of electricity generation have different characteristics: biomass, waste and other renewable energy sources are more stable, and onshore wind and photovoltaic power generation with a standard deviation of 5242.79 and 1402.747 have a great volatility, respectively; see Figure 5.

Tab. 1 – Descriptive statistics of day-head electricity price and renewable energy generation.

index	mean	std	min	25%	50%	75%	max	Skew	Kurt
price	41.06	15.17	2.3	32.12	41.69	50.05	88.44	-0.13	-0.16
biomass	366.13	76.73	0	327	352	391	592	0.52	0.65
geothermal	0	0	0	0	0	0	0	0	0
pumped storage	527.52	828	0	0	101	740	4162	1.95	3.28
run-of-river	994	445	0	607	925	1359	1939	0.41	-1.05
water reservoir	3302	2230	0	1269	3041	4947	9389	0.46	-0.84
marine	0	0	0	0	0	0	0	0	0
other	81.84	10.85	0	75	83	89	115	-0.31	0.74
solar	1403	1667	0	67	614	2428	5792	1.07	-0.28
waste	259.10	49.95	0	220	265	302	350	-0.65	0.01
offshore	0	0	0	0	0	0	0	0	0
onshore	5243	3176	0	2764	4609	7104	16284	0.91	0.40

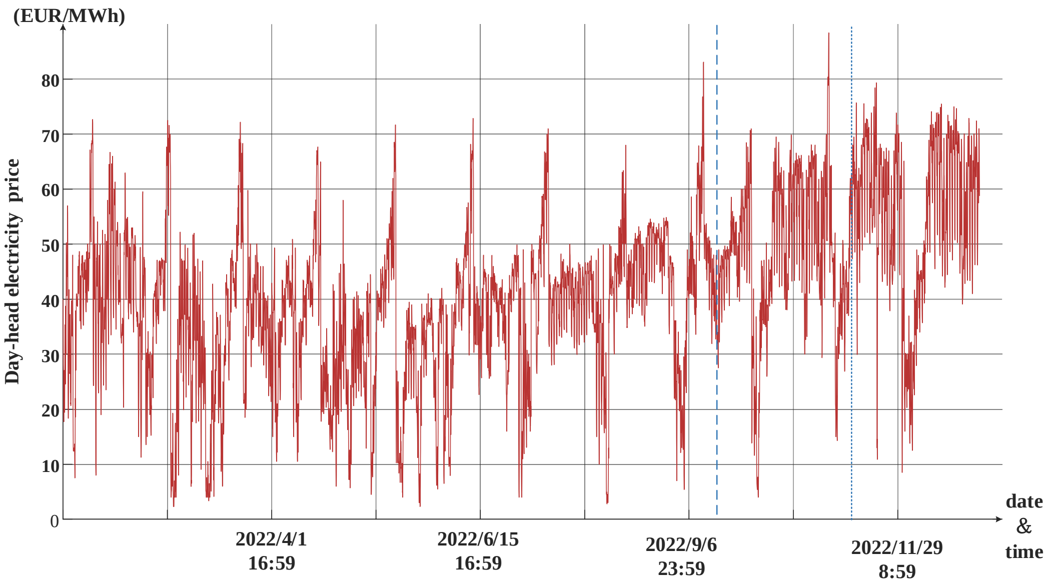


Fig. 4 – Spanish day-ahead electricity price data for 2022 year

Based on this analysis of the European electricity market and the Spanish electricity market, as well as the analysis of day-ahead electricity prices and renewable energy sources, the aim is to construct an ICEEMDAN-iTransformer model capable of accurately predicting day-ahead spot electricity prices. The model will take into account the impact of variations in generation from various renewable energy sources and daily fluctuations in electricity consumption on the day-ahead spot electricity price. It is expected to provide electricity market participants with valuable information on electricity price forecasts and help them better formulate energy purchasing, sales and storage strategies to cope with market uncertainty and volatility.

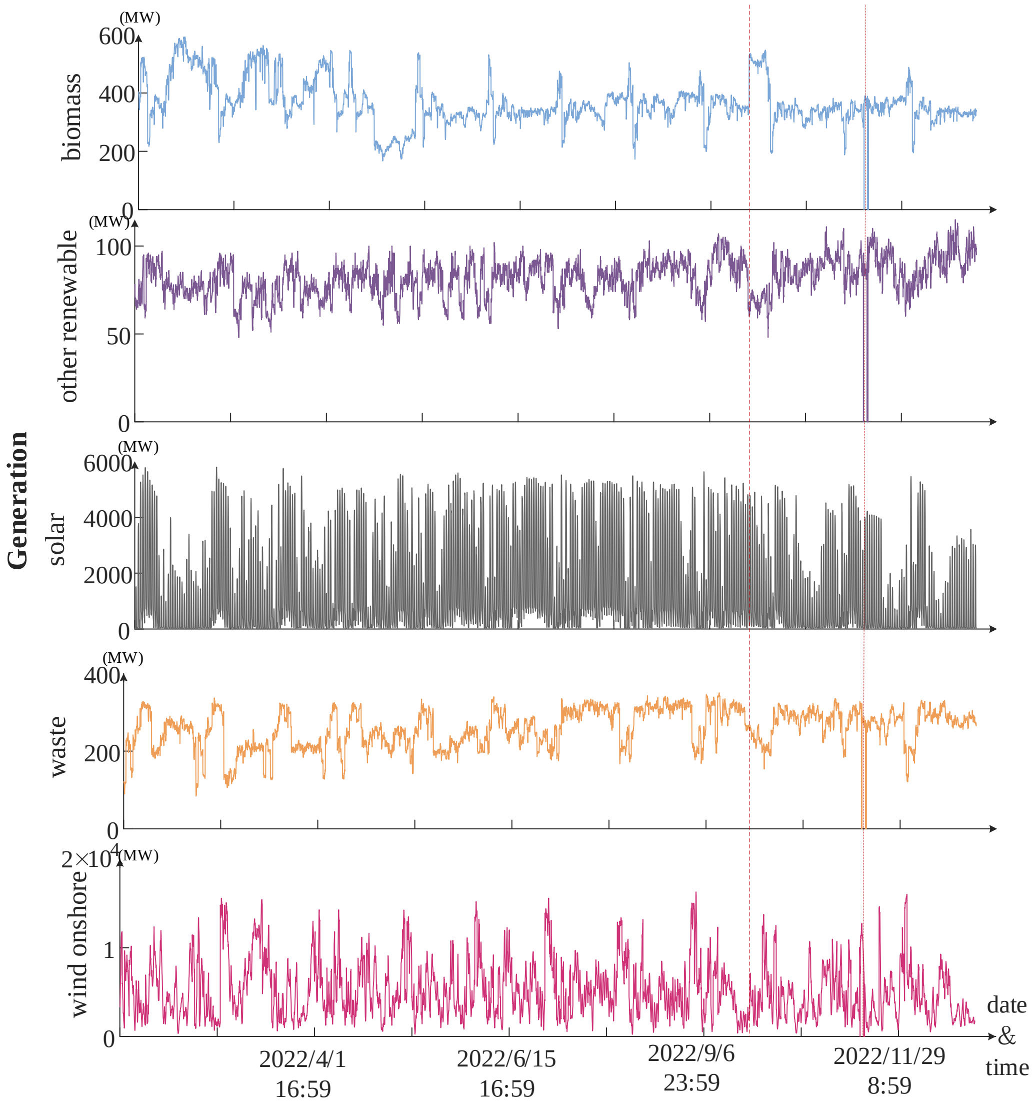


Fig. 5 – Five renewable energy generation data for Spain in 2022

4 CASE STUDY

Forecasting in energy markets is recognized as one of the most impactful areas where machine/deep learning contributes to transitioning to a renewable-based electrical infrastructure. In addition to the evaluation metrics and time efficiency analysis, it is crucial to validate the performance of the proposed models more comprehensively. Cross-validation is a widely used technique in machine learning to assess how well a model will generalize to an independent dataset. In this study, we apply cross-validation to further enhance the reliability of our model evaluation results and provide more robust evidence for the effectiveness of the ICEEMDAN-iTransformer model. In the process of forecasting long-term day-ahead electricity prices, the section 4.1 describes the time efficiency of the model including the minimum epoch time and total time, as well as the parameters of the iTransformer model, the environment, hardware configuration and platform required for this work. In section 4.2, which starts univariate day-head electricity price long-term forecasting by using the proposed iTransformer model and compares it to the benchmark models based on the formers and the invert-formers, and in section 4.3, the multivariate electricity price prediction is carried out considering the impact of renewable energy on electricity prices, and the same comparison between the models is carried out. Section 4.4 utilizes the ICEEMDAN decomposition method to preprocess the electricity prices in order to improve the prediction speed and prediction accuracy of the iTransformer model. In order to improve the robustness and reliability of the research results, interpretability, cross-validation and uncertainty quantification techniques were used in Section 4.5 to verify the ICEEMDAN-iTransformer model. Finally, in section 4.6, the policy implications and implementation recommendations of the electricity price and forecasting model are presented.

4.1. Predictive model parameters and time efficiency

Usually, models with higher complexity have longer response times. However, long-term electricity price time-series forecasting requires high real-time performance, meaning the model must quickly process input data and generate predictions. Therefore, comparing the responsiveness of the complex model through the time efficiency values is crucial for predicting the electricity price, and we implement the minimum epoch time and the total training time to evaluate the model responsiveness (Bian et al., 2021). The total training time is the total time required for the whole training process from the beginning to the end, and the minimum epoch time is the shortest time required to complete an epoch, and they depend on several factors, including the size of the model, the amount of training data, the hardware environment (e.g., the performance and the number of CPUs and GPUs), the optimization algorithm, and the speed of data loading. Overall, total training time and minimum epoch time are two important values for evaluating the training efficiency of machine learning models. The total training time reflects the time-consumption of the whole training process, while the minimum epoch time focuses more on the efficiency of individual epochs. Among them, the iTransformer model parameter settings are shown in the Table 2., and the operating environments for the work experiments are CPU 16 2.10GHz Intel(R) Xeon(R) Platinum 8352V, GPU RTX4090(24GB)*1, 30 GB of RAM, 120 GB of memory, Windows 10 OS and PyCharm 2023.1 platform, interpreter virtual environments PyTorch 2.0.0, Python 3.8 (ubuntu20.04) and Cuda 11.8.

Tab. 2 – The iTransformer model's parameters, description, and default values

Parameters	Description	Default
d_model	The principal dimension of the model, which usually determines the size of embeddings and transformations within the model	512
n_heads	Number of heads used in multi-head attention mechanisms	8
e_layers	Layers of Encoder	16
d_layers	Number of layers of decoder	8
d_ff	Dimension of the feedforward neural network	2048
moving_avg	Window size of the moving average	67
factor	Attention factor	1
dropout	Discard rate, which randomly "turns off" a portion of the nodes in the network to prevent overfitting	0.3
activation	Type of activation function	softmax
batch_size	Batch size of training input data	16

Furthermore, to mitigate the risk of overfitting, the model incorporates a dropout layer with a rate of 0.3 in the architecture. This helps prevent the model from over-relying on specific features and enhances its generalization ability. Additionally, we employed early stopping during the training process to monitor the validation loss, ensuring that the model stops training when the validation performance no longer improves. These strategies effectively reduce the risk of overfitting, especially given the relatively short period of the training dataset.

4.2. Comparative experiment 1：Electricity price forecasting

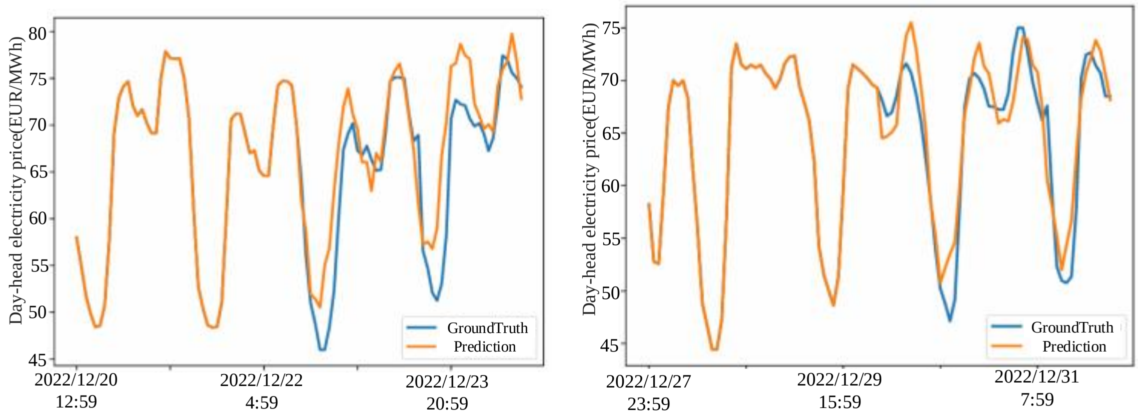
In experiment 1, before the beginning of the comparison experiment, the univariate day-ahead electricity price is predicted by iTransformer on the test set of 878 data samples, and the univariate is denoted by S. From the results of the prediction, 603-699 and 811-878 are selected as a demonstration, as shown in Figure 6. The evaluation metrics are compared with those of the transformer, informer, reformer, flowformer, and flashformer models, and their invertible models: iInformer, iReformer, iFlowformer, and iFlashformer. The results of the univariate day-ahead electricity price prediction are presented in Table 3, which compares the performance of different models. The results of the regression evaluation metrics of which $MSE_S^{iT} = 15.16403$, $MAE_S^{iT} = 2.66547$, $RMSE_S^{iT} = 3.89410$, $MAPE_S^{iT} = 0.05807$, and $MSPE_S^{iT} = 0.01111$, time efficiency $T*_S^{iT} = 19.08380s$ ($T*$ stands for the total training completion time, as follows) and $MINepo_S^{iT} = 4.65607s$ ($MINepo$ stands for the single minimum epoch training completion time, as follows). Compared to the second-best performance $RMSE_S^{T} = 4.02761$ improved by 3.3149% and the worst performance $RMSE_S^{In} = 15.63953$ improved by 75.1009%. For the remaining four regression evaluation metrics MSE_S, MAE_S, $MSPE_S$, and $MAPE_S$, the iTransformer' metrics are also obtained Optimum compared to the remaining nine models. When training large data samples, by comparing the training total time and the minimum epoch training time, iTransformer also embodies the best time performance. In addition, invertible models perform better than the general model to verify the accuracy advantage of former models' regression prediction，such as $MAE_S^{R} \prec MAE_S^{iR}$, $MSPE_S^{In} \prec MSPE_S^{iIn}$, etc, and the transformer model can predict the electricity price better than the benchmark model. Conclusions obtained from the experiments in this subsection are as follows.

The iTransformer model performs well on the univariate day-ahead electricity price long-term prediction task, not merely outperforming other former variants and its inverted model in terms of prediction accuracy, but demonstrating significant advantages in training efficiency. The better performance of inverted models (such as iTransformer) compared to the benchmark former model in predicting time series data such as electricity prices demonstrates the effectiveness of invert operations or architectural adjustments in improving model performance.

Tab. 3 – The evaluation metrics of the univariate day-ahead electricity price long-term prediction

Experiment1	MSE	MAE	RMSE	MAPE	MSPE	min epoch time	train total time
iTransformer	**15.16403**	**2.66547**	**3.89410**	**0.05807**	**0.01111**	4.65607	19.08380
Transformer	16.22160	2.70541	4.02761	0.05934	0.01161	16.21223	81.90100
iInformer	16.22187	2.66739	4.02764	0.05812	0.01162	5.39546	23.39110
Informer	244.5949	12.3916	15.6396	0.3252	0.3980	9.26983	67.8623
iReformer	16.22188	2.66743	4.02764	0.05812	0.01162	7.05380	65.29040
Reformer	25.11915	3.58263	5.01190	0.0873	0.0275	21.3851	86.4040
iFlowformer	16.22444	2.67316	4.02796	0.0583	0.0116	7.06277	29.1705
Flowformer	200.7352	11.5294	14.1681	0.2857	0.2780	16.12163	81.68250

iFlashformer	169.5713	10.58634	13.02195	0.25856	0.20943	5.98980	24.69680
Flashformer	105.9423	8.30927	10.29283	0.19888	0.12010	17.34529	88.07880



a) iTransformer (603-699)   b) iTransformer (782-878)

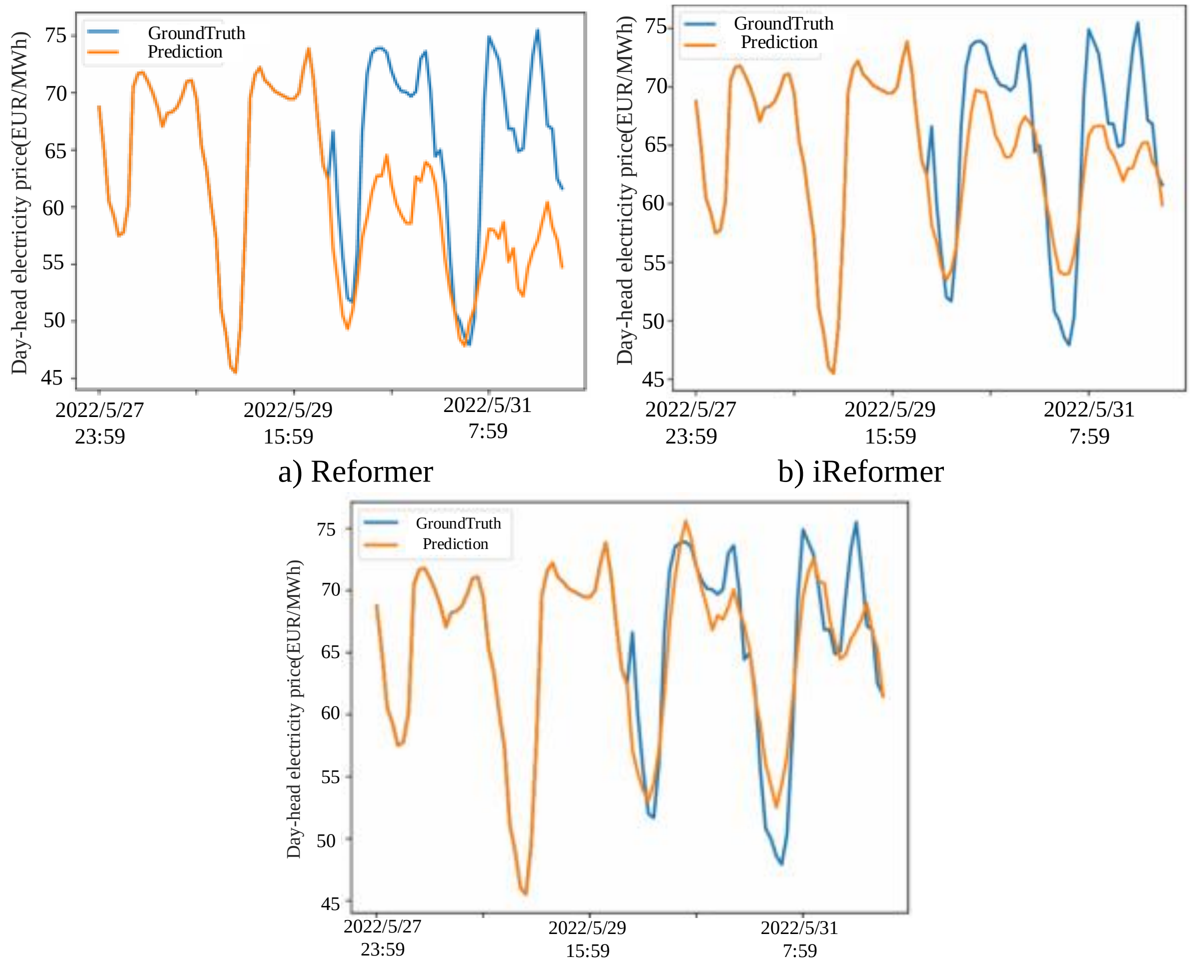
Fig. 6 – Partly dated iTransformer prediction and groundtruth curves

4.3. Comparative experiment 2：Multivariate Renewable Energy Forecasts for Long-Term Day-Ahead Electricity Prices

Based on the long-term univariate day-ahead electricity price forecasting in Section 4.2, the iTransformer model demonstrates superior prediction performance. However, the model prediction regression indicator evaluation metrics are still high, such as $2.6 \prec MAE_S^{iT}$, in order to improve the accuracy of electricity price prediction and reduce the error between the real value and the predicted value. To account for the impact of renewable energy on electricity prices, we incorporate datasets including offshore wind power generation, solar generation, and nine other renewable energy generation types to predict long-term day-ahead electricity prices. The values of the evaluation metrics are shown in the Table 4. Among them, the $MAE_{MS}^{iT} = 0.68159$ is reduced by 1.983876 compared with the univariate prediction value, and the prediction accuracy is improved by 74.4287%; similarly, the $RMSE_{MS}^{iT} = 2.16687$ is reduced by 1.72723 compared with the univariate prediction value $RMSE_S^{iT}$, and the prediction accuracy is improved by 44.355%, which means that multivariate prediction of long-term day-ahead electricity price by adding renewable energies can improve prediction accuracy compared with the single day-ahead electricity price. At the same time, the performance of each evaluation metrics of iTransformer model is still optimal compared with other models. The prediction of reformer, iReformer and iTransformer test dataset 469-565 are selected as shown in Figure 7. This respectively represents a), b), and c), and it can be seen that reformer predicts poorly, and the invert of Reformer improves the prediction curve and real value curve. The iTransformer model demonstrates significantly better performance, showing an excellent fit between its prediction curve and the true value curve while substantially reducing prediction errors compared to the iReformer model. Therefore, the conclusions in Section 4.2 are still applicable to renewable energy multivariate prediction of day-ahead electricity price, reflecting the versatility of iTransformer in different scenarios. The key conclusions are summarized below:

Incorporating renewable energy generation data as an input to the forecasting model can significantly improve the forecasting accuracy of day-ahead electricity prices in the long term.

Meanwhile, the iTransformer model is still proven to be the optimal model choice in renewable energy multivariate forecasting due to its excellent forecasting performance.



a) Reformer b) iReformer

c) iTransformer

Fig. 7 – Partly dated Reformer, iReformer, iTransformer prediction and groundtruth curves

Tab. 4 – The evaluation metrics of multivariate features forecasts electricity prices

Experiment2	MSE	MAE	RMSE	min epoch time	train total time
iTransformer	**4.695327**	**0.681594**	**2.16687**	14.54956913	59.70454788
Consider other features	4.669511	0.758503	2.134172	45.17404	581.5983
Transformer	9.163141	1.227417	3.027068	21.68234706	87.8282671
iInformer	4.695855	0.683136	2.166992	15.5395484	62.58625722
Informer	75.03659	3.870082	8.662366	16.46343279	66.52321124
iReformer	4.695373	0.681657	2.166881	15.86065292	63.99518967
Reformer	5.335418	0.910943	2.309852	16.92648888	68.05429554
iFlowformer	4.695336	0.6817	2.166872	15.88871312	64.60833597
Flowformer	6.440564	1.046804	2.537827	25.24079251	102.6827655
iFlashformer	4.695335	0.681698	2.166872	15.87379789	63.82011175
Flashformer	6.284764	1.039362	2.506943	31.73850846	129.0352755

We further incorporate meteorological factors, carbon policy data (Zeng et al., 2023), social activity-derived load patterns, and renewable generation features to assess their combined effects on iTransformer's price forecasting accuracy. The experiments in Table 4 show that the model prediction performance is improved after fusing these features, but the improvement is limited. The MSE of the model is 4.669, the MAE is 0.758, and the RMSE is 2.134, but the minimum training round time reaches 45.174, and the total training time is as high as 581.598. Due to the significant increase in training time and the insignificant improvement in prediction performance, we decided to eliminate these features by considering the prediction cost and computational efficiency, and then carry out the subsequent data decomposition and model optimization work.

4.4. The ICEEMDAN-based decomposition electricity prices prediction

Section 4.4 incorporates renewable energy generation data into the iTransformer model for multivariate long-term forecasting of day-ahead electricity prices, which significantly reduces forecasting errors. However, the day-ahead electricity price is not only affected by renewable energy, but also by a series of other complex factors, including the supply and demand situation in the electricity market, seasonal patterns, tariff regulation mechanism, environmental and climatic factors, and cost factors, etc., each of which may have different degrees of influence on electricity price, and there may be interactions and correlations among these factors, making it impossible to cover all the relevant factors for predicting the day-ahead electricity price in the forecast. Therefore, section 4.5 proposes to decompose the raw time series data into a number of components or trends in order to better understand the intrinsic structure and dynamics of the data.

In addition, we also performed an ADF test on the residuals of the ICEEMDAN decomposition as shown in Figure. 8, which shows an ADF statistic of -7.275 with a p-value of 1.55e-10. This result indicates that the residuals are smooth and do not contain important trend or period information. Therefore, the residual component was chosen to be ignored and its effect on the prediction results is negligible.

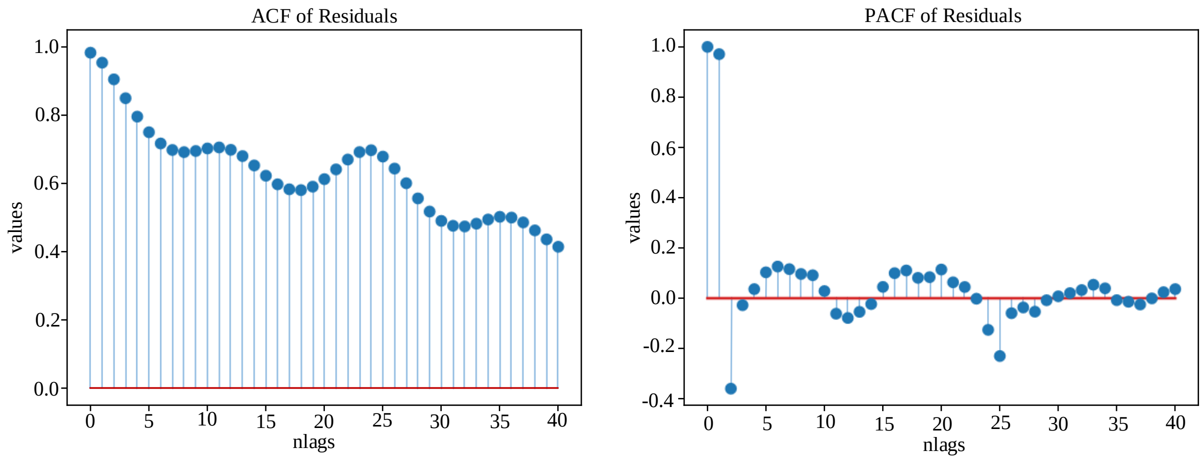


Fig. 8 – Residual test

The ICEEMDAN decomposition algorithm adaptively decomposes the day-ahead electricity price data, with a signal-to-noise ratio of 0.2 and a maximum of 100 iterations to reduce data volatility and prediction complexity, and the decomposition results are shown in Figure 9. Sets of intrinsic mode functions (IMF) are obtained, and the frequencies of each modal component are relatively stable, and there is no obvious modal aliasing. The components gradually decrease from the high-frequency IMF1 to the low-frequency IMF5, and the data tend to stabilize gradually. In addition, high-frequency data represent short-term fluctuation trends, and lower frequency data represent longer fluctuation trends, which clearly identifies the long-term trends and cyclical changes of the data.

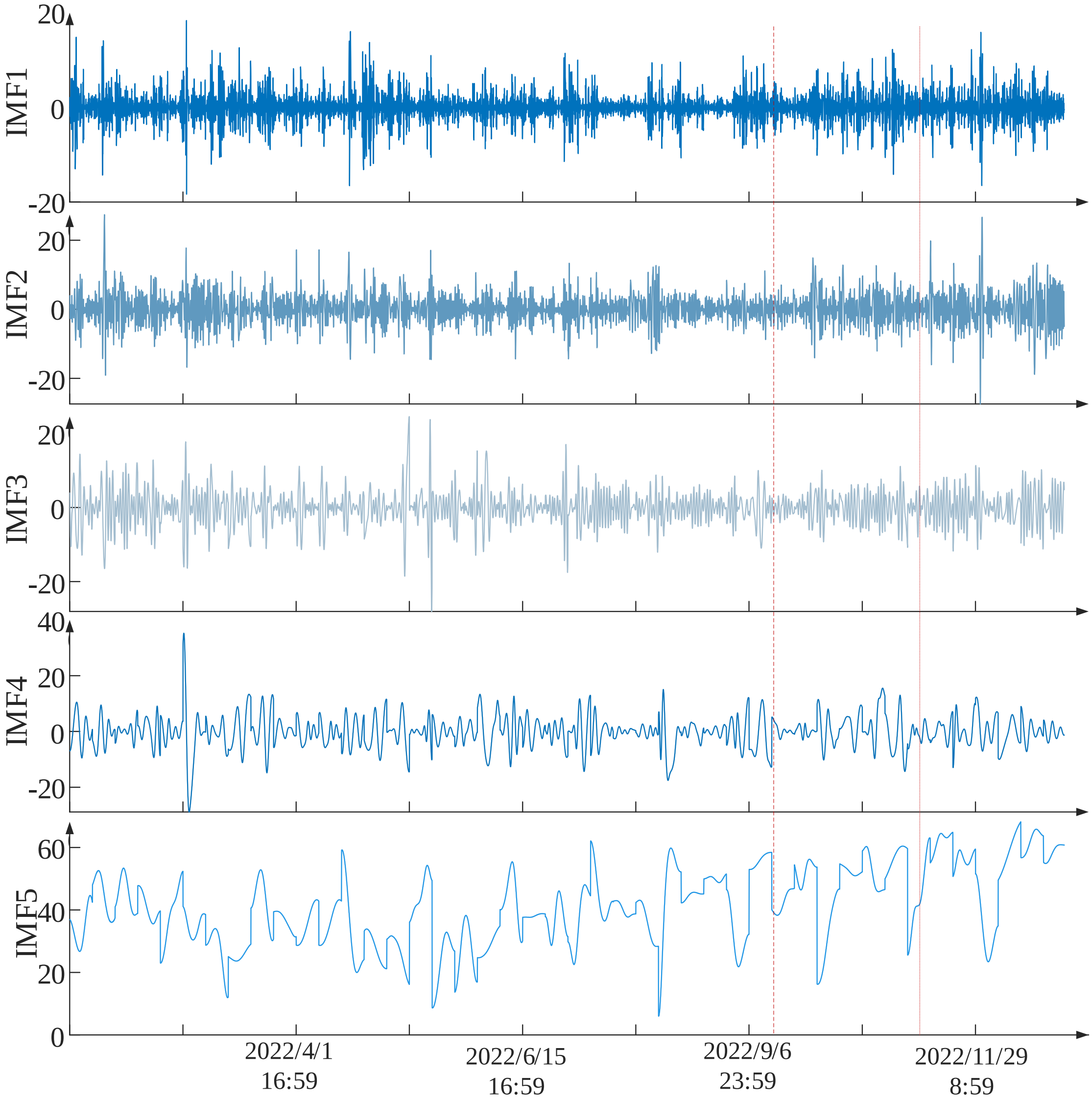


Fig. 9 – The ICEEMDAN algorithmic decomposition of electricity prices

After obtaining the components, the mean aggregation reconstruction is used to obtain the predicted electricity price under the day-head electricity price dataset, and Table 5 shows the performance of the iTransformer model for multivariate long-term prediction of five eigenmodal components based on the regression evaluation metrics. The comparative results between iTransformer and ICEEMDAN-iTransformer are illustrated in Figure 10. The evaluation value $MAE_{MIMF1}^{Ice} = 2.03284$ of the prediction metrics for the component $IMF1$ is larger than that $MAE_{MIMF2}^{Ice} = 1.79558$ for the component $IMF2$ and the MAE_{MIMF}^{Ice} decreases sequentially in the order of the number from IMF1 to IMF5, the smallest component $IMF5$ has an evaluation value of $MAE_{MIMF5}^{Ice} = 0.15415$, and the other two metrics MSE_{MIMF}^{Ice} and $RMSE_{MIMF}^{Ice}$ also follow the law of sequential decrease pattern, which indicates that the lower the frequency of the intrinsic modal component, the smoother the data, the better the prediction and the lower the error. In order to verify the prediction effect of ICEEMDAN-iTransformer, the decomposed modal prediction indicator evaluation values are equally weighted through equal weighting method, and the MAE_M^{Ice-iT}, MSE_M^{Ice-iT}, and $RMSE_M^{Ice-iT}$ are 3.21430, 0.99947, and 1.57464, compared to the undecomposed model metrics evaluation value $RMSE_M^{iT} = 2.16688$. The decomposed prediction error is reduced by 0.592238, and the prediction accuracy is improved by 27.331%. Although the MAE is slightly higher, ICEEMDAN-iTransformer performs better in terms of MSE, RMSE, and minimum training epoch time, which implies that it fits the data better overall and the fluctuation range of the predictions is smaller. Moreover, MAE is more sensitive to outliers, while MSE/RMSE gives more weight to larger errors. As a result, ICEEMDAN-iTransformer may perform slightly worse than iTransformer in dealing with extreme values or outliers, but in most cases it provides more stable predictions. In combination with iTransformer, ICEEMDAN-iTransformer may be more advantageous in dealing with noise, nonlinearity, and non-smoothness in the data, i.e., it is more robust in dealing with complex data. Despite the current slightly higher MAE metrics, ICEEMDAN-iTransformer has the potential to be further optimized. By adjusting the parameters or combining other techniques, the MAE can be further reduced while maintaining the advantages of other metrics.

Tab. 5 – The evaluation metrics of ICEEMDAN-based decomposition electricity prices prediction.

Experiment3		MSE	MAE	RMSE	min epoch time	train total time
iTransformer	IMF1	8.21625	2.032836	2.8664	15.1222	62.76353
	IMF2	5.292674	1.795578	2.300581	13.37018	55.21547
	IMF3	1.407574	0.779975	1.186412	13.87818	87.74128
	IMF4	0.586017	0.234817	0.765518	12.72086	81.29473
	IMF5	**0.568975**	**0.154148**	**0.754304**	14.80903	60.33127
ICEEMDAN- iTransformer		**3.214298**	0.999471	**1.574643**	13.98009	69.46926
iTransformer		4.695327	**0.681594**	2.166881	14.54957	59.70455

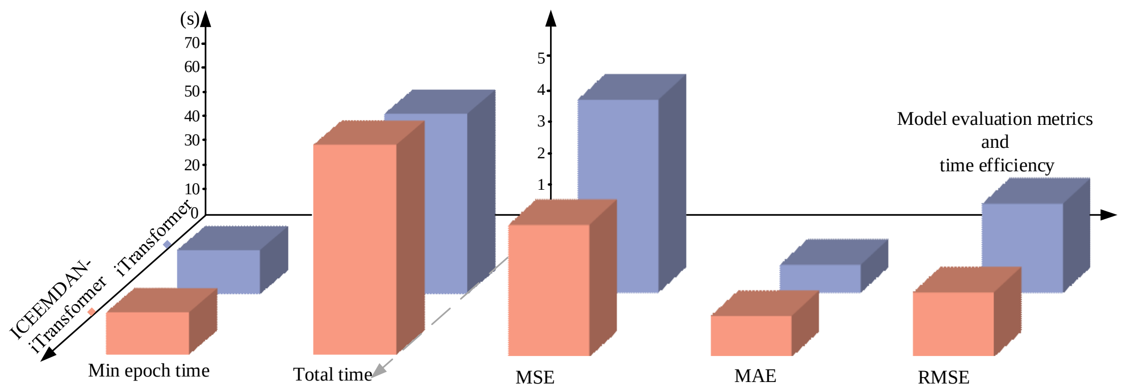


Fig. 10 – The evaluation comparative results between iTransformer and ICEEMDAN-iTransformer

4.5 Model Validation Techniques

4.5.1 Interpretability

The interpretability of the model is a key aspect that requires more attention. While the ICEEMDAN-iTransformer model significantly improves forecast accuracy, this section fully explains its predictions from the perspective of market participants. Figure 11 illustrates the average absolute SHAP values for renewable energy generation for each input feature, which represent the average degree to which these features influence the model's output. For example, the relatively high average absolute SHAP values for onshore wind generation and solar generation indicate that they have a significant impact on the final model output and dominate the final forecast results.

In addition to these renewable energy output characteristics, the average absolute SHAP values for hydropower-related characteristics (e.g., reservoir hydropower, run-of-river hydropower, and pumped-storage consumption) are in the medium range, indicating that they also have some influence on the model outputs, and are one of the important factors influencing the prediction results, but to a slightly lesser extent than the former two. As for the other energy generation features, their average absolute SHAP values are low and have relatively little impact on the final model output.

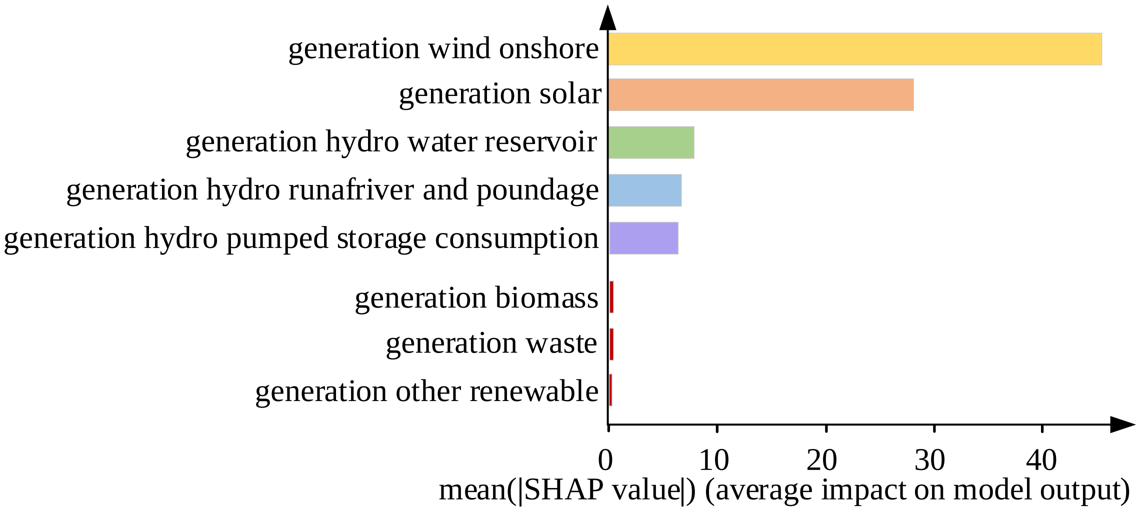


Fig. 11 – Average degree of influence of input features on SHAP output from iTransformer

To elucidate the logical relationships between the features in the model's explanatory framework - in particular, their positive or negative impact on electricity price forecasts - we refer to Figure 12. Typically, high wind generation negatively affects the price of electricity (resulting in lower prices), which is due to the low cost and abundant supply of wind power, while low wind power generation may lead to higher prices. Similarly, high solar generation typically has a negative impact on electricity prices, especially during periods of abundant sunlight, while low solar generation may push up prices. By analyzing these SHAP values, market participants can have a more intuitive understanding of the model's forecasting logic, thereby increasing confidence in the forecast results.

This visual presentation provides a more intuitive way to understand how different input characteristics affect the forecast results. More detailed analyses based on such visualizations (e.g., further exploring how fluctuations in wind and solar power generation affect forecasts) can be valuable to policymakers, market operators, and industry practitioners.

4.5.2 Probabilistic Forecasting and Outlier Sensitivity Analysis

The results of probabilistic forecasting are presented in Table 6 and Figure 13. The ICEEMDAN-iTransformer model consistently achieves the lowest pinball loss across all quantiles (α=0.1 to α=0.9), demonstrating superior performance in probabilistic forecasting.

For instance, at α=0.1, the pinball loss is 1.3235, significantly lower than other models (e.g., informer: 2.6726, iInformer: 1.9944). The iTransformer model also performs well, with pinball loss values slightly higher than ICEEMDAN-iTransformer but still better than the baseline transformer and informer models. This suggests that the inverted transformer architecture enhances probabilistic forecasting accuracy. The informer and transformer models exhibit higher pinball loss values, indicating less accurate probabilistic forecasts, particularly at extreme quantiles (α=0.1 and α=0.9). These results indicate that the ICEEMDAN-iTransformer model effectively captures uncertainty in electricity prices and provides reliable confidence intervals.

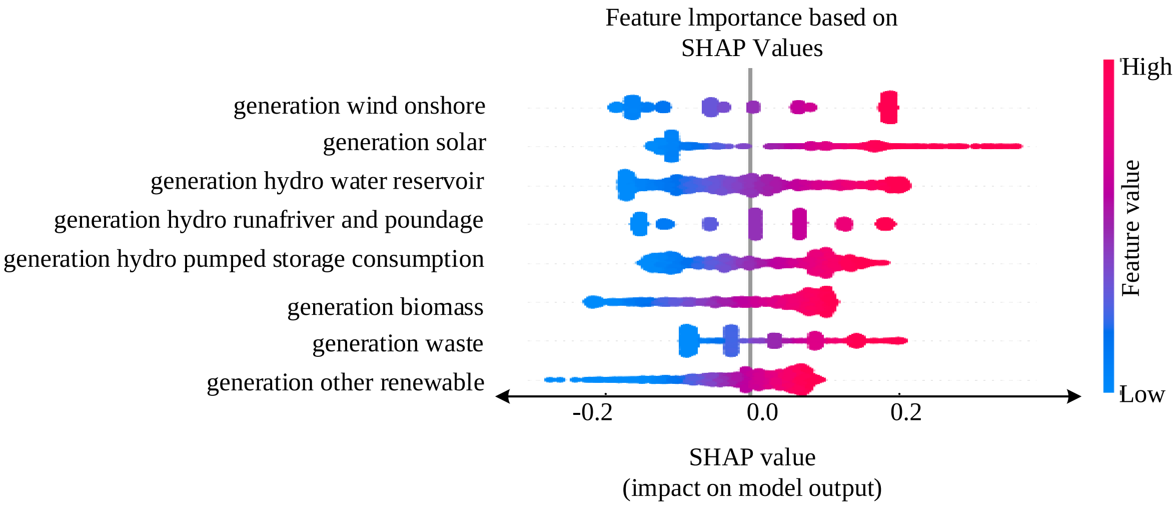


Fig. 12 – Feature Importance of input features based on SHAP values from iTransformer

The ICEEMDAN-iTransformer model achieves a CRPS of 8.6278, slightly higher than the iInformer (8.6112) but lower than the transformer (8.6681) and informer (8.6680). This indicates that ICEEMDAN-iTransformer provides a better fit to the observed cumulative distribution function compared to the baseline models. The iInformer model has the lowest CRPS, suggesting it performs slightly better in terms of overall probabilistic forecasting accuracy. However, the difference is minimal, and ICEEMDAN-iTransformer remains competitive.

Tab. 6 – Comparative performance of models in probabilistic forecasting and outlier sensitivity analysis

Model	Pinball Loss (α=0.1)	Pinball Loss (α=0.3)	Pinball Loss (α=0.5)	Pinball Loss (α=0.7)	Pinball Loss (α=0.9)	CRPS	MSE (Top 5%)	MSE (Bottom 5%)
Informer	2.6726	2.6873	2.7020	2.7167	2.7313	8.6680	17.7662	0.1960
iInformer	1.9944	1.9542	1.9141	1.8740	1.8339	8.6112	19.0153	0.0063
ICEEMDAN-iTransformer	1.3235	1.3260	1.3286	1.3312	1.3338	8.6278	10.8407	0.0858
Transformer	2.1531	2.2231	2.2932	2.3632	2.4332	8.6681	18.1176	0.1130
iTransformer	1.9741	1.9331	1.8920	1.8510	1.8099	8.6311	14.8767	0.1198

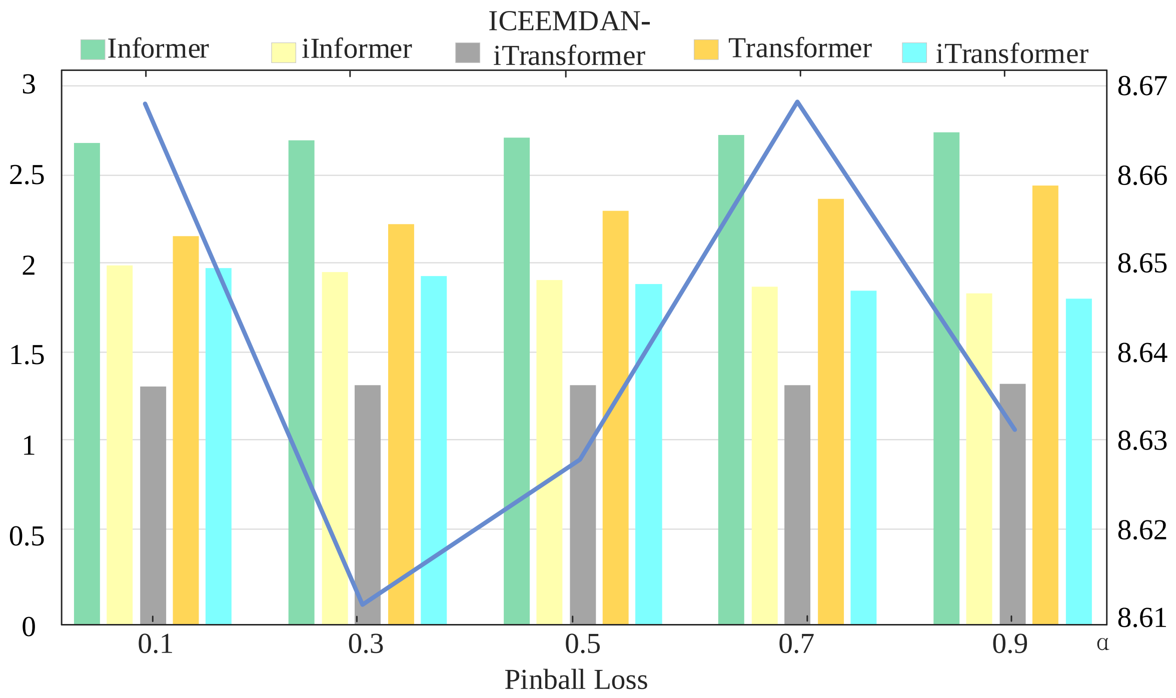


Fig. 13 – Pinball loss and CRPS values for different models

The results of outlier sensitivity are presented in Table 6 and Figure 14. The ICEEMDAN-iTransformer model achieves the lowest MSE (10.8407) for the top 5% of price values, indicating its robustness in handling extreme price spikes. This is a significant improvement over the informer (17.7662) and transformer (18.1176) models. The iInformer model performs slightly worse than ICEEMDAN-iTransformer, with an MSE of 19.0153, suggesting that the ICEEMDAN decomposition further enhances the model's ability to handle extreme values.

The iInformer model achieves the lowest MSE (0.0063) for the bottom 5% of price values, indicating excellent performance in predicting extreme price drops. However, the ICEEMDAN-iTransformer model also performs well, with an MSE of 0.0858, significantly better than the transformer (0.1130) and informer (0.1960) models. The iTransformer model has a slightly higher MSE (0.1198) for the bottom 5% outliers, suggesting that the ICEEMDAN decomposition helps improve the model's sensitivity to extreme low prices. This demonstrates the model's robustness in handling extreme price fluctuations, which are common in electricity markets due to factors such as sudden changes in renewable energy generation or demand spikes.

Overall, the ICEEMDAN-iTransformer model demonstrates the best performance in probabilistic forecasting, with the lowest pinball loss across all quantiles and the lowest MSE for the top 5% outliers. This indicates that the model is highly effective in capturing uncertainty and handling extreme price fluctuations. The iInformer model performs well in terms of CRPS and MSE for the bottom 5% outliers but is slightly outperformed by ICEEMDAN-iTransformer in handling extreme price spikes. The baseline informer and transformer models show higher errors across all metrics, highlighting the limitations of traditional transformer architectures in probabilistic forecasting and outlier sensitivity.

The ICEEMDAN-iTransformer model stands out as the most robust and accurate model for probabilistic forecasting and handling extreme price fluctuations in electricity markets. Its ability to capture uncertainty and adapt to extreme values makes it a strong candidate for real-world applications in electricity price forecasting. The iInformer model also shows promise, particularly in handling extreme low prices, but the ICEEMDAN-iTransformer model offers a more balanced and comprehensive performance across all metrics.

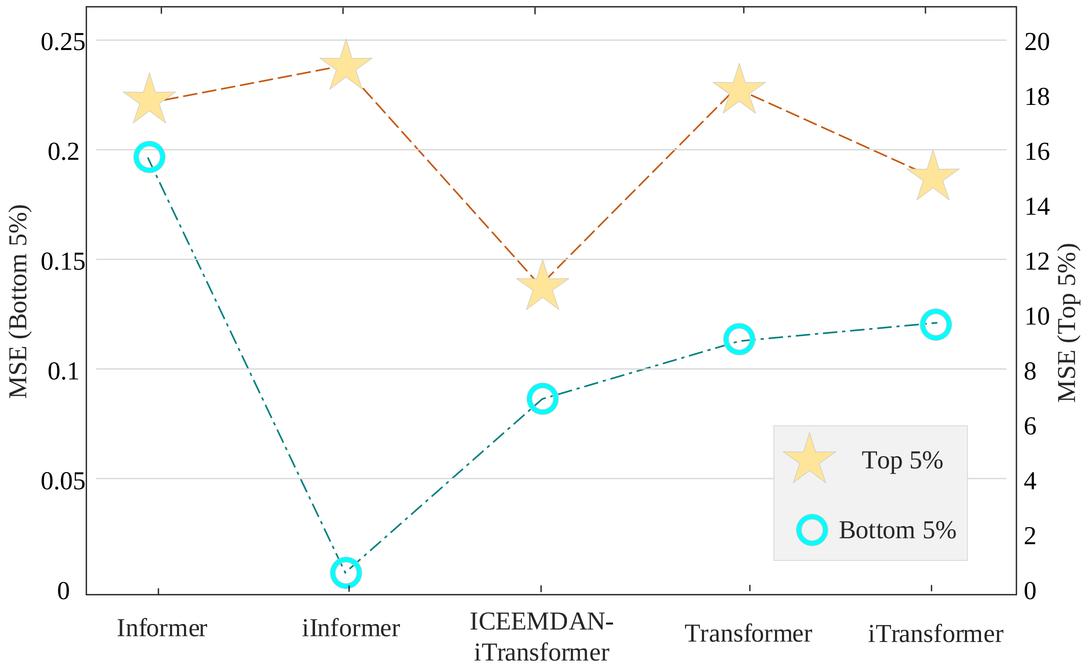


Fig. 14 – MSE for top and bottom 5% outliers for different models

4.5.3 Cross-Validation

To evaluate the generalization ability of the models, we conducted k-fold cross-validation with k=5. The dataset was divided into five folds, and each model was trained on four folds while validated on the remaining fold. This process was repeated five times, ensuring that each fold was used exactly once as the validation set. The average performance metrics across all folds are reported in Table 7, and the main observations are illustrated in Figure 15.

The ICEEMDAN-iTransformer model achieves the lowest average MSE (2.698) across all five folds, indicating the highest overall accuracy in predicting day-ahead electricity prices. This performance is significantly better than the baseline models, such as the informer (5.464) and transformer (4.578). The iInformer model also performs well, with an average MSE of 3.788, which is lower than the baseline transformer and informer models but higher than ICEEMDAN-iTransformer. This suggests that the inverted transformer architecture improves forecasting accuracy, but the ICEEMDAN decomposition further enhances performance. The iTransformer model, which does not include the ICEEMDAN decomposition, achieves an average MSE of 3.662, slightly better than the iInformer but still outperformed by ICEEMDAN-iTransformer.

The ICEEMDAN-iTransformer model exhibits the lowest standard deviation (0.066), indicating the most consistent performance across different folds. This suggests that the model is robust and generalizes well to unseen data. The iInformer and iTransformer models also show relatively low standard deviations (0.131 and 0.095, respectively), indicating stable performance across folds. However, their standard deviations are higher than ICEEMDAN-iTransformer, suggesting slightly less consistency. The informer and transformer models have higher standard deviations (0.112 and 0.156, respectively), indicating more variability in their performance across different folds. This suggests that these models are less robust and may overfit to specific subsets of the data.

Tab. 7 – Cross-validation performance of models in day-ahead electricity price forecasting

Model	Metrics	Fold1	Fold2	Fold3	Fold4	Fold5	Average	Standard deviation
Informer	MSE	5.403	5.349	5.526	5.654	5.387	5.464	0.112
	MAE	2.653	2.622	2.682	2.704	2.632	2.659	0.034
	MAPE	19.719	18.518	18.454	19.294	17.886	18.774	0.728
	MSPE	21.564	16.935	14.801	19.953	18.743	18.399	2.629
iInformer	MSE	3.828	3.565	3.942	3.730	3.874	3.788	0.131
	MAE	2.460	2.380	2.504	2.430	2.484	2.452	0.049
	MAPE	12.208	10.581	12.468	11.188	11.569	11.603	0.764
	MSPE	8.828	3.831	6.057	4.391	5.143	5.650	1.963
ICEEMDAN-iTransformer	MSE	2.621	2.676	2.754	2.797	2.643	2.698	0.066
	MAE	1.920	1.939	1.964	1.987	1.928	1.947	0.028
	MAPE	8.970	8.757	9.145	8.682	8.215	8.754	0.352
	MSPE	3.522	3.056	3.109	2.841	2.439	2.993	0.396
Transformer	MSE	4.586	4.431	4.593	4.855	4.423	4.578	0.156
	MAE	2.531	2.436	2.508	2.565	2.462	2.500	0.052
	MAPE	17.933	16.490	15.845	16.582	15.221	16.414	1.011
	MSPE	23.567	16.664	12.705	15.939	14.194	16.614	4.181
iTransformer	MSE	3.784	3.527	3.707	3.716	3.577	3.662	0.095
	MAE	2.302	2.206	2.271	2.295	2.245	2.264	0.039
	MAPE	13.177	11.933	12.124	11.241	11.281	11.951	0.788
	MSPE	10.141	7.856	5.432	5.918	5.355	6.940	2.057

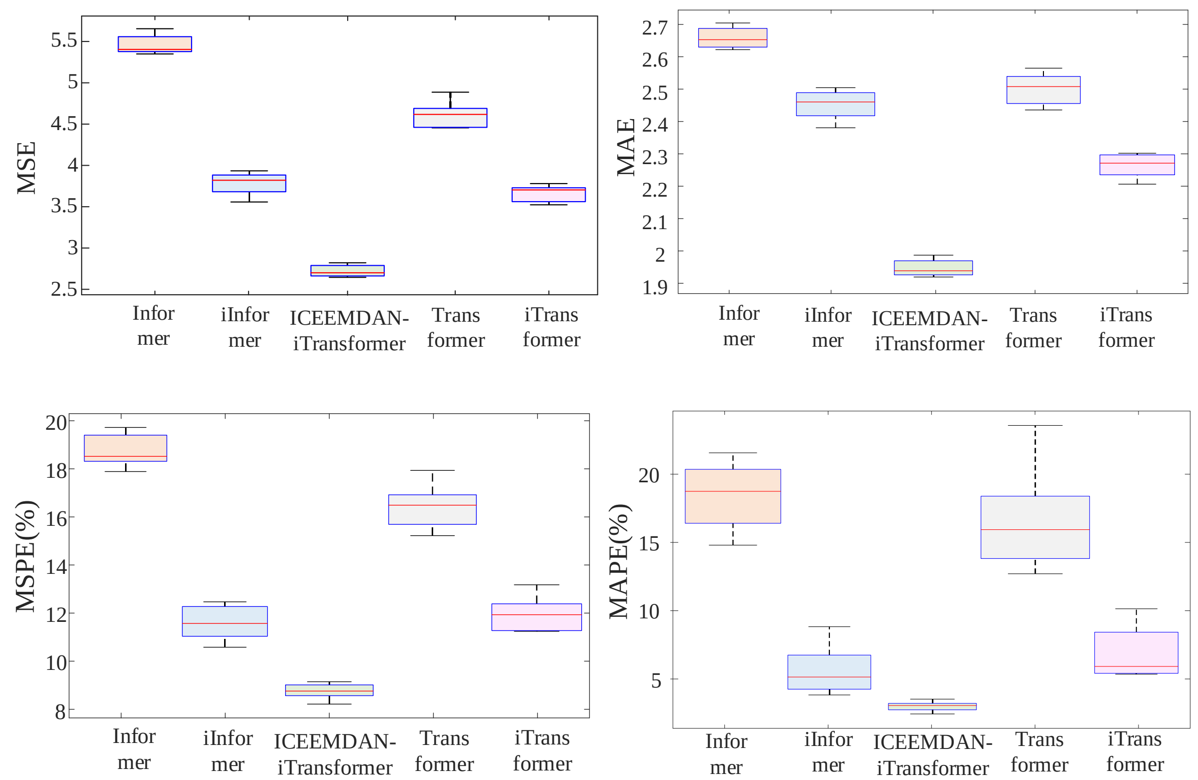


Fig. 15 – Boxplot of Cross-Validation for Day-Ahead Electricity Price Forecasting Models

The ICEEMDAN-iTransformer model consistently achieves the lowest MSE in each fold, with values ranging from 2.621 (fold1) to 2.797 (fold4). This demonstrates its ability to maintain high accuracy across different subsets of the data. The iInformer and iTransformer models show competitive performance, with MSE values generally lower than the baseline models but higher than ICEEMDAN-iTransformer. For example, the iInformer achieves an MSE of 3.565 in fold2, which is its best performance, but it still underperforms compared to ICEEMDAN-iTransformer. The informer and transformer models exhibit higher MSE values across all folds, with the transformer showing particularly poor performance in fold4 (MSE = 4.855). This highlights the limitations of traditional transformer architectures in handling the complexities of electricity price forecasting.

The ICEEMDAN-iTransformer model demonstrates the best overall performance in cross-validation, with the lowest average MSE and the most consistent results across folds. This indicates that the model is robust and generalizes well to unseen data. The iInformer and iTransformer models also show strong performance, especially in terms of consistency, but they are outperformed by ICEEMDAN-iTransformer in terms of accuracy. The baseline informer and transformer models show higher errors and larger fold changes, indicating that they are less suitable for day-ahead tariff forecasting in volatile markets. The results demonstrate that the ICEEMDAN-iTransformer model maintains consistent performance across different subsets of the data, indicating its robustness to variations in the training and validation sets.

4.6 Policy Implications and Implementation Suggestions

The ICEEMDAN-iTransformer model offers significant implications for both enterprises and governments, particularly in the context of Spain’s day-ahead electricity market, which operates through time-of-day bidding and accounts for approximately 86% of the total spot market volume. By leveraging the model’s accurate day-ahead electricity price forecasts, stakeholders can make more informed decisions to optimize resource allocation, improve market efficiency, and promote sustainable energy practices.

(1) Government Applications

Governments can use the model’s predictions to optimize the allocation of resources, ensuring a more efficient and fair electricity market. This includes planning energy development strategies that prioritize the utilization of renewable energy resources while facilitating the transition and upgrading of traditional energy infrastructure.

(2) Enterprise Applications

Enterprises can use the model’s forecasts to better control costs, formulate production plans, and assess risks. For example, predicting price spikes allows businesses to adjust energy consumption or purchase electricity in advance to minimize expenses.

(3) Integration into Trading and Grid Management

The model can be integrated into electricity trading platforms through a two-step process. First, a data transformation layer standardizes input data (e.g., historical prices and renewable energy generation) for compatibility with the ICEEMDAN decomposition. Second, an API connects the model to trading algorithms, enabling real-time decision-making. For instance, traders can adjust bidding strategies based on predicted price spikes or drops. Grid operators can use the model’s forecasts to optimize power generation scheduling. For example, if the model predicts high prices during peak hours and abundant renewable energy supply in the afternoon, operators can prioritize renewable generation, reducing reliance on fossil fuels. The model also helps predict power shortages or surpluses, allowing operators to take preventive measures such as importing power or implementing demand-response programs.

By implementing these strategies, the ICEEMDAN-iTransformer model can enhance decision-making in electricity trading, grid management, and policy development, while its adaptability ensures relevance across diverse energy markets. This contributes to a more efficient, sustainable, and resilient energy ecosystem.

5 CONCLUSION

In this paper, we proposed a novel day-ahead electricity price forecasting model, ICEEMDAN-iTransformer, which integrates improved complete ensemble empirical mode decomposition with adaptive noise (ICEEMDAN) and the iTransformer model. The model was designed to address the challenges posed by the increasing integration of renewable energy into the electricity market, which introduces significant uncertainty and volatility in electricity prices. By decomposing historical electricity price data using ICEEMDAN, we were able to extract intrinsic mode functions (IMFs) that capture the nonlinear and non-stationary characteristics of the data. The iTransformer model, with its innovative inverted architecture, was then employed to independently predict the multidimensional time series data, including both IMFs and renewable energy generation data.

A set of day-ahead electricity prices from the Valencia electricity market in Spain are used in the predictive analysis and led to some important conclusions.

(1) Our experimental results demonstrated that the ICEEMDAN-iTransformer model outperforms several benchmark models as well as their inverted variants. The model exhibited superior performance in terms of prediction accuracy, handling of noise, nonlinearity, and non-smoothness in the data, and robustness to extreme price fluctuations. Specifically, the ICEEMDAN-iTransformer model achieved the lowest mean squared error (MSE), mean absolute error (MAE), and root mean square error (RMSE) in both univariate and multivariate forecasting scenarios. Additionally, the model demonstrated excellent performance in probabilistic forecasting, as evidenced by its low pinball loss and continuous ranked probability score (CRPS), indicating its ability to capture uncertainty and provide reliable confidence intervals.

(2) The incorporation of renewable energy generation data as an input to the forecasting model significantly improved the accuracy of day-ahead electricity price predictions. Furthermore, the ICEEMDAN decomposition process enhanced the model's ability to handle complex data structures, leading to more stable and accurate forecasts. The model's robustness was further validated through cross-validation and outlier sensitivity analysis, which showed consistent performance across different subsets of the data and during periods of extreme price fluctuations.

The proposed model is a robust tool for day-ahead electricity price forecasting and can be extended to other complex time series forecasting tasks. However, in the future, the following aspects of this study still need to be improved and studied in depth: (i) More advanced model validation techniques, including uncertainty quantification methods (e.g., Bayesian neural networks or Monte Carlo dropout), should be further explored to more fully assess the predictive power and potential risks of models. (ii) Expand the model to consider the interrelationships between different electricity markets (e.g., regional, national, and international markets). As energy trading becomes more globalized, price movements in one market can influence others.

References

1. Bian, S., et al. (2021). Machine learning-based real-time monitoring system for smart connected worker to improve energy efficiency. *Journal of Manufacturing Systems*, *61*, 66-76. https://doi.org/10.1016/j.jmsy.2021.08.009
2. Bonilla, J., Blanco, J., Zarza, E., & Alarcón-Padilla, D. C. (2022). Feasibility and practical limits of full decarbonization of the electricity market with renewable energy: Application to the Spanish power sector. *Energy*, *239*(E), 122437. https://doi.org/10.1016/j.energy.2021.122437
3. Chatterjee, S., & Chakraborty, S. (2024). A study on the effects of objective weighting methods on TOPSIS-based parametric optimization of non-traditional machining processes. *Decision Analytics Journal*, *11*, 100451. https://doi.org/10.1016/j.dajour.2024.100451
4. Che, P., Zhang, C., Liu, Y., & Zhang, Y. (2025). An integrated learning and optimization approach to optimal dynamic retail electricity pricing of residential and industrial consumers. *Applied Energy*, *382*, 125234. https://doi.org/10.1016/j.apenergy.2024.125234
5. Chen, Z., et al. (2025). Outlier-adaptive-based non-crossing quantiles method for day-ahead electricity price forecasting. *Applied Energy*, *382*, 125328. https://doi.org/10.1016/j.apenergy.2025.125328
6. Ciarreta, A., Pizarro-Irizar, C., & Zarraga, A. (2020). Renewable energy regulation and structural breaks: An empirical analysis of Spanish electricity price volatility. *Energy Economics*, *88*, 104749. https://doi.org/10.1016/j.eneco.2020.104749
7. Colominas, M. A., Schlotthauer, G., & Torres, M. E. (2014). Improved complete ensemble EMD: A suitable tool for biomedical signal processing. *Biomedical Signal Processing and Control*, *14*. https://doi.org/10.1016/j.bspc.2014.06.009
8. Darandary, A., Mikayilov, J. I., & Soummane, S. (2024). Impacts of electricity price reform on Saudi regional fuel consumption and CO2 emissions. *Energy Economics*, *131*, 107400. https://doi.org/10.1016/j.eneco.2024.107400
9. Davis, D., & Brear, M. J. (2024). Impact of short-term wind forecast accuracy on the performance of decarbonising wholesale electricity markets. *Energy Economics*, 130, 107304. https://doi.org/10.1016/j.eneco.2024.107304
10. Enrich, J., Li, R., Mizrahi, A., & Reguant, M. (2024). Measuring the impact of time-of-use pricing on electricity consumption: Evidence from Spain. *Journal of Environmental Economics and Management*, *123*, 102901. https://doi.org/10.1016/j.jeem.2023.102901
11. Ghimire, S., Deo, R. C., Casillas-Pérez, D., & Salcedo-Sanz, S. (2022). Improved complete ensemble empirical mode decomposition with adaptive noise deep residual model for short-term multi-step solar radiation prediction. *Renewable Energy*, *190*, 408-424. https://doi.org/10.2139/ssrn.3995358
12. González-Tejero, C. B., Ancillo, A., Gavrila, S. G. & Blanco, A. G. (2023). Uncovering the complexities of intellectual property management in the era of AI: Insights from a bibliometric analysis. *Journal of Competitiveness*, *15*(4), 69-86. https://doi.org/10.7441/joc.2023.04.05

13. Grothe, O., Kächele, F., & Krüger, F. (2023). From point forecasts to multivariate probabilistic forecasts: The Schaake shuffle for day-ahead electricity price forecasting. *Energy Economics*, *120*, 106602. https://doi.org/10.1016/j.eneco.2023.106602
14. Guan, H., et al. (2022). Financial fraud identification of the companies based on the logistic regression model. *Journal of Competitiveness*, *14*(4), 155-171. https://doi.org/10.7441/joc.2022.04.09
15. Guo, K., et al. (2025). Can climate factors improve the forecasting of electricity price volatility? Evidence from Australia. *Energy*, *315*, 134332. https://doi.org/10.1016/j.energy.2024.134332
16. Hiew, S. Y., et al. (2023). Prediction of ultimate conditions and stress–strain behaviour of steel-confined ultra-high-performance concrete using sequential deep feed-forward neural network modelling strategy. *Engineering Structures*, *277*, 115447. https://doi.org/10.1016/j.engstruct.2022.115447
17. Huang, W., Gao, T., Hao, Y., & Wang, X. (2023). Transformer-based forecasting for intraday trading in the Shanghai crude oil market: Analyzing open-high-low-close prices. *Energy Economics*, *127*, 107106. https://doi.org/10.1016/j.eneco.2023.107106
18. Huang, Y., et al. (2022). Application of a hybrid model based on the prophet model, ICEEMDAN and multi-model optimization error correction in metal price prediction. *Resources Policy*, *79*, 102969. https://doi.org/10.1016/j.resourpol.2022.102969
19. Karahan, C. C., Odabaşi, A., & Tiryaki, C. S. (2024). Wired together: Integration and efficiency in European electricity markets. *Energy Economics*, *133*, 107505. https://doi.org/10.2139/ssrn.4407959
20. Kaya, M., Karan, M. B., & Telatar, E. (2023). Electricity price estimation using deep learning approaches: An empirical study on Turkish markets in normal and Covid-19 periods. *Expert Systems with Applications*, *224*, 120026. https://doi.org/10.1016/j.eswa.2023.120026
21. Lehna, M., Scheller, F., & Herwartz, H. (2022). Forecasting day-ahead electricity prices: A comparison of time series and neural network models taking external regressors into account. *Energy Economics*, *106*, 105742. https://doi.org/10.1016/j.eneco.2021.105742
22. Liu, H., et al. (2024). Interpretable two-layer day-ahead electricity price forecast based on calibration window combination and coupled market characteristics. *Proceedings of the CSEE*, *04*(44), 1272-1286. https://doi.org/10.13334/j.0258-8013.pcsee.221453
23. Liu, Y., et al. (2024). ITransformer: Inverted transformers are effective for time series forecasting. Cornell University arXiv. https://doi.org/10.48550/arXiv.2310.06625
24. Luo, Y., et al. (2024). Evaluating the impact of AI research on industry productivity: A dynamic qualitative comparative analysis approach. *Journal of Competitiveness*, *16*(3), 187-203. https://doi.org/10.7441/joc.2024.03.09
25. Maciejowska, K., Nitka, W., & Weron, T. (2021). Enhancing load, wind and solar generation for day-ahead forecasting of electricity prices. *Energy Economics*, *99*, 105273. https://doi.org/10.1016/j.eneco.2021.105273
26. Méndez-Suárez, M., Simón-Moya, V. & Muñoz-de Prat, J. (2023). Do current regulations prevent unethical AI practices? *Journal of Competitiveness*, *15*(3), 207-222. https://doi.org/10.7441/joc.2023.03.11

27. Meng, A., Zhu, J., Yan, B., & Yin, H. (2024). Day-ahead electricity price prediction in multi-price zones based on multi-view fusion spatio-temporal graph neural network. *Applied Energy*, *369*, 123553. https://doi.org/10.1016/j.apenergy.2024.123553
28. Nie, Y., Nguyen, N. H., Sinthong, P., & Kalagnanam, J. (2023). A time series is worth 64 words: Long-term forecasting with transformers. Cornell University arXiv. https://doi.org/10.48550/arXiv.2211.14730
29. Özen, K., & Yildirim, D. (2021). Application of bagging in day-ahead electricity price forecasting and factor augmentation. *Energy Economics*, *103*, 105573. https://doi.org/10.1016/j.eneco.2021.105573
30. Pourdaryaei, A., et al. (2024). A new framework for electricity price forecasting via multi-head self-attention and CNN-based techniques in the competitive electricity market. *Expert Systems with Applications*, *235*, 121207. https://doi.org/10.1016/j.eswa.2023.121207
31. Ramírez, I., Aizpurua, J. I., Lasa, I., & Río, L. (2024). Probabilistic feature selection for improved asset lifetime estimation in renewables. Application to transformers in photovoltaic power plants. *Engineering Applications of Artificial Intelligence*, *131*, 107841. https://doi.org/10.1016/j.engappai.2023.107841
32. Tan, Y. Q., Shen, Y. X., Yu, X. Y., & Lu, X. (2023). Day-ahead electricity price forecasting employing a novel hybrid frame of deep learning methods: A case study in NSW Australia. *Electric Power Systems Research*, *220*, 109300. https://doi.org/10.2139/ssrn.4156560
33. Vaswani, A., et al. (2017). Attention is all you need. Cornell University arXiv. https://doi.org/10.48550/arXiv.1706.03762
34. Xu, Y., et al. (2024). VMD-ATT-LSTM electricity price prediction based on grey wolf optimization algorithm in electricity markets considering renewable energy. *Renewable Energy*, *236*, 121408. https://doi.org/10.1016/j.renene.2024.121408
35. Xu, Y., Li, Q., & Cui, H. (2024). Short-term multi-step price prediction for the electricity market with a high proportion of clean energy and energy storage based on MIC-EEMD-improved informer. *Power System Technology*, *03*(48), 949-958. https://doi.org/10.13335/j.1000-3673.pst.2023.1671
36. Yang, H., Wu, Q., & Li, G. (2024). A multi-stage forecasting system for daily ocean tidal energy based on secondary decomposition, optimized gate recurrent unit and error correction. *Journal of Cleaner Production*, *449*, 141303. https://doi.org/10.1016/j.jclepro.2024.141303
37. Yao, Z., Zhang, L., Zou, B., & Gu, S. (2020). Electricity price prediction for electricity market with high proportion of wind power. *Automation of Electric Power Systems*, *12*(44), 49-55. https://doi.org/10.3390/en15041345
38. Yin, H., et al. (2022). Day-ahead electricity price forecasting of electricity market with high proportion of new energy based on LSTM-CSO model. *Power System Technology*, *02*(46), 472-480. https://doi.org/10.13335/j.1000-3673.pst.2021.1056
39. Zeng, J., Ribeiro-Soriano, D., Ren, J. & Sansaloni, E. M. (2023). Does China's low-carbon city pilots policy stimulate green development? Evidence from a multiple mediating effect model. *Journal of Competitiveness*, *15*(4), 179-213. https://doi.org/10.7441/joc.2023.04.10
40. Zhao, Z., et al. (2017). LSTM network: A deep learning approach for short-term traffic forecast. *IET Intelligent Transport Systems*, *11*(2), 68–75. https://doi.org/10.1049/iet-its.2016.0208

41. Zhu, Y., et al. (2024). A clearing method for cascade hydropower spot market considering coupling relations of hydraulic power, electric power and electricity price. *Journal of Cleaner Production*, *440*, 140758. https://doi.org/10.3390/en15041345

Contact information

Lect. Puliang Du, Ph.D.
Shanghai University of Electric Power
School of Economics and Management
Shanghai, China
E-mail: dupuliang@163.com
ORCID: https://orcid.org/0000-0001-7881-3616

Miaoheng Yang
Shanghai University of Electric Power
School of Economics and Management
Shanghai, China
E-mail: yangmh7247@126.com
ORCID: https://orcid.org/0009-0001-6898-5972

Assoc. Prof. Fei Xie, Ph.D. (*corresponding author*)
Shanghai University of Finance and Economics
School of Finance
Shanghai, China
E-mail: xiefei@mail.shufe.edu.cn

Prof. Wei Hu, Ph.D.
Shanghai University of Electric Power
School of Economics and Management
Shanghai, China
E-mail: hwshiep@shiep.edu.cn
ORCID: https://orcid.org/0000-0003-4978-2376

Jiyu Li
Economic and Technical Research of State Grid Fujian Electric Power Co., Ltd
Fujian, China
E-mail：1583344845@qq.com

This work was supported by the Shanghai Science and Technology Plan Program (Sailing Program) (Grant Number 23YF1414300), and the National Natural Science Foundation of China (Grant Number 72401182).