

# *Market-based Short-term Electricity Price Forecast Based on EEMD and ARIMA*

Jun Dong, Xihao Dou\*, Dongran Liu, Dongxue Wang

School of Economics and Management

North China Electric Power University

No.2, Beinong Road, Huilongguan, Changping District, Beijing, 102206

\*2268728128@qq.com

**Abstract**—With the development and implementation of pilots in the Chinese spot market, the importance of electricity price forecasting to various entities in the electricity market is also constantly emerging. At the same time, with the continuous deepening of China's electricity market construction, the electricity market structure of each province gradually tends to be market-oriented, and the original electricity price forecasting method has been unable to adapt to the future changes in the structure of the electricity market. Based on the characteristics of real-time electricity prices in the spot market, this study proposes a short-term electricity price mixed forecasting model based on ARIMA and EEMD to predict the 24-hour electricity price. This hybrid model uses the ARIMA model to predict the time series decomposed by EEMD, which improves the stationarity of the electricity price series and improves the accuracy. Using MATLAB and SPSS tools, reference the electricity price data of the America PJM day-ahead market to verify the proposed prediction model. The results show that the hybrid model can significantly improve the forecast accuracy of electricity prices compared with the traditional ARIMA method.

**Keywords**—*Electricity price forecast; EEMD; ARIMA; Mixed model; Electricity market*

## I. INTRODUCTION

With the development of the electricity market, the important role of electricity price forecasting has become increasingly prominent. Under the general environment of China's power system reform, the power market structure has undergone tremendous changes. In particular, with the launch of the pilot spot market, many factors are uncertain, and the original prediction methods have certain limitations. Therefore, the electricity price forecast method applicable to the new electricity market environment plays a crucial role for all market players [1]. According to the electricity price forecast results, power generation companies can adjust the power generation plan and formulate the optimal power generation strategy and quoting strategy. Power users can adjust their electricity consumption plans in a timely manner to improve their peak load and reduce the cost of electricity. The power system can reverse the user's electricity consumption behavior according to electricity price forecast, reduce the operating cost of the power system, and ensure safe and reliable operation of the power grid.

The traditional short-term electricity price forecast models include regression model method, trend extrapolation method and time series method. At present, the main methods are neural networks, time series and gray models [1-4]. The most commonly used methods are time series models and neural networks [5]. The literature [2-3] uses different forecasting models, but they are based on BP neural network for short-term electricity price forecasting. The literature [4] uses artificial neural network, BP neural network and other neural network methods to model the nonlinearity of the electricity price sequence. The prediction accuracy depends on the number of hidden layer nodes and the selected input data characteristics. The neural network for power price prediction requires a large amount of data for effective training to achieve high prediction accuracy. For the time series method, the most commonly used models are the following: Autoregressive (AR) model, Moving Average (MA) model, Autoregressive Moving Average (ARMA) model, etc. The literature [5] uses the ARIMA and ARCH double model based on time series and neural network to predict the short-term electricity price, but does not convert the forecast result into data. The time series of electricity price forecasting requires less data, which can effectively predict changes in electricity price trends. However, a single model can easily lead to the generation of forecast deviations, and can't effectively reflect the authenticity of the forecast results. Therefore, it needs to be combined with other models.

China's electricity market reform has a certain degree of volatility. The short-term electricity price curve has strong non-linearity and non-stationarity. Therefore, the effective decomposition of the electricity price curve can reduce the non-stationarity of the electricity price curve and improve the accuracy of the forecast [6-7]. This paper uses the Cluster Empirical Mode Decomposition (EEMD) to decompose historical electricity prices and reduce the nonstationarity of historical electricity prices. In combination with the ARIMA model, the seasonality and deviation characteristics of the data are taken into account to predict the decomposed curves without excessive historical data. In this paper, the validity of the model is verified by using the 24-hour price data published by the PJM power market. Avoid the disadvantages of single use of EEMD or ARIMA, retain the advantages of its forecast model, and achieve better forecast results without affecting the original trend and volatility of historical electricity prices.

II. ENSEMBLE EMPIRICAL MODE DECOMPOSITION (EEMD)

The EEMD model is based on the improvement of the EMD model. EMD is essentially an adaptive signal screening method, which can gradually filter out different trend features hidden in the original sequence. These different curves with their own trend characteristics are called Intrinsic Mode Functions (IMF), and the intrinsic modal components must satisfy two conditions:

1) There are at most one difference between the number of poles and the number of poles in the entire natural modal component sequence.;

2) For any point, the envelope defined by the local minimum points and the envelope defined by the local maximum points have an average value of 0.

EEMD is white noise that each contains a signal and an additional finite amplitude. The ensemble empirical mode decomposition steps are as follows:

1) The white noise with a mean of 0 and a standard deviation of the amplitude is added to the original time series  $x(t)$  multiple times:

$$x_i(t) = x(t) + n_i(t)$$

Where  $x_i(t)$  is the signal after adding the white noise at the i-th time;  $n_i(t)$  is the white noise added at the i-th time.

2) The newly formed sequence  $s_i(t)$  is decomposed separately, and then the respective IMF is  $c_{ij}(t)$ , with a remainder  $r_i(t)$ ,

where  $r_i(t)$  is the first time to add white noise after EMD decomposition to get the j-th IMF.

3) Integrate the corresponding IMF in the above decomposition into the average value, and get the final IMF after EEMD decomposition, namely:

$$C_j(t) = \frac{1}{N} \sum_{i=1}^n c_{ij}(t)$$

Where N is the number of times that white noise is added;  $C_j(t)$  represents the j-th IMF obtained after EEMD decomposition of the original signal.

III. AUTO REGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA)

In ARIMA (p, d, q), AR is "autoregressive", p is the number of autoregressive terms, MA is "smoothing average," q is the number of sliding averaged terms, and d is the number of difference times that make it a stationary sequence (order). The ARIMA (p, d, q) model is an extension of the ARMA (p, q) model.

Process of ARIMA model application:

1) The stationarity of time series is identified by scatter plot, autocorrelation function and partial autocorrelation function diagram.

2) The stationary time series data are processed smoothly until the processed autocorrelation function and the partial autocorrelation function are nonsignificant nonzero.

3) Establish a corresponding time series model based on the identified characteristics. After the smoothing process, if the partial autocorrelation function is truncated and the autocorrelation function is trailing, an AR model is established. If the partial autocorrelation function is trailing and the autocorrelation function is truncated, the MA model is established; if the partial autocorrelation function and the autocorrelation function are trailing, the sequence is suitable for the ARMA model.

4) Parameter estimation, and check for statistical significance.

5) Hypothesis test, judge (diagnose) whether the residual sequence is white noise sequence.

6) Using a validated model to predict.

In the process of modeling, ARIMA can effectively test the periodicity and seasonality of the electricity price. Therefore, it is necessary to establish a time series model with time characteristics to predict the electricity price in different time periods, and to realize the seasonal and periodic performance in the process of electricity price prediction.

IV. SIMULATION COMPARISON

A. Model establishment

ARIMA model can get a more accurate prediction based on SPSS. In order to improve the accuracy of prediction, ARIMA and EEMD are combined into a new prediction method EEMD-ARIMA. The main steps of the combined model are: "Decomposition-Prediction-Combination". The historical electricity price curve is decomposed by EEMD, which can decompose different internal fluctuation curves and trend characteristics of the data. It can also improve the smoothness of short-term historical electricity prices, simplify the time series model's ability to identify electricity price signals, and improve the accuracy of forecasting. "Combination" is to restore the authenticity of short-term historical electricity prices. In the current study, less literature used the EEMD and ARIMA models to improve the accuracy of traditional ARIMA model predictions in electricity prices. The main model flow is as follows:

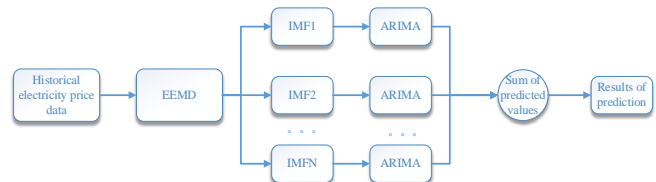


Fig. 1 Prediction framework of EEMD-ARIMA model electricity price

The specific steps of the EEMD-ARIMA prediction model are as follows:

(1) Inputting the time series X (t) of the historical electricity price of the original signal into EEMD, and the original signal

is decomposed into several IMF and residual  $r_n(t)$ . Residual is the trend item T;

(2) Since each IMF and T are independent, and the sum is equivalent to the input original signal, ARIMA modeling is performed for each IMF, and the corresponding predicted values of each MF and T are obtained respectively;

(3) Sum the predicted values of each IMF and T, and the sum is the final electricity price forecast data.

(4) The mean absolute percent error is a prediction error index expressed by the relative number form percentage, and it is one of the important indexes to measure the precision of prediction, which is expressed as:

$$MAPE = \frac{1}{m} \sum_{t=1}^m \left[ \frac{|\hat{y}_t - y_t|}{y_t} \right]$$

Where MAPE represents the average absolute percentage error, m represents the number of points of the prediction data,  $\hat{y}_t$  is the original data, and  $y_t$  is the data predicted by the model. The smaller the value, the smaller the error of the model and the higher the fitting degree.

**B. Sample selection**

To verify the feasibility of the proposed method, the historical electricity prices from May 31 to June 11 in the PJM electricity spot market and their corresponding historical load data were selected for simulation. China currently does not have a complete electricity spot market. The PJM electricity market is mature and various systems are sound. Many provinces use PJM as a template for the construction of the spot market. The use of PJM data has a certain forward-looking, can better reflect the power market supply and demand characteristics and changes in market structure, the original price sequence diagram shown in Fig. 2.

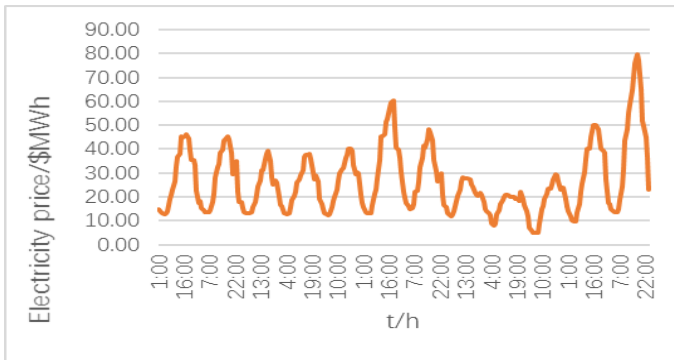


Fig. 2 Historical electricity prices for one week in the PJM spot market

**C. Result analysis**

In order to compare the accuracy of the model, a contrast group was established. First, the electricity price is decomposed using EEMD, then simulation is performed using ARIMA as model 1; and ARIMA is used to predict the undecomposed original sequence as model 2. The error uses MAPE as an index to evaluate the performance of the model.

The original electricity price column of EEMD decomposition of model 1 is shown in Fig. 3. EEMD decomposes the original electricity price series into six groups of IMF and T. It can be seen from Fig. 4 that the signal after decomposition has a different point change trend and is gradually leveled down, and the decomposed data is fitted and predicted by ARIMA. Model 2 directly uses ARIMA to perform fitting prediction on pjm's raw data. The main difference between the two lies in whether or not to preprocess the original sequence. Model 1 is to decompose the original data. The decomposition curve fluctuates more slowly than the original curve, but it does not change the characteristics of the original sequence.

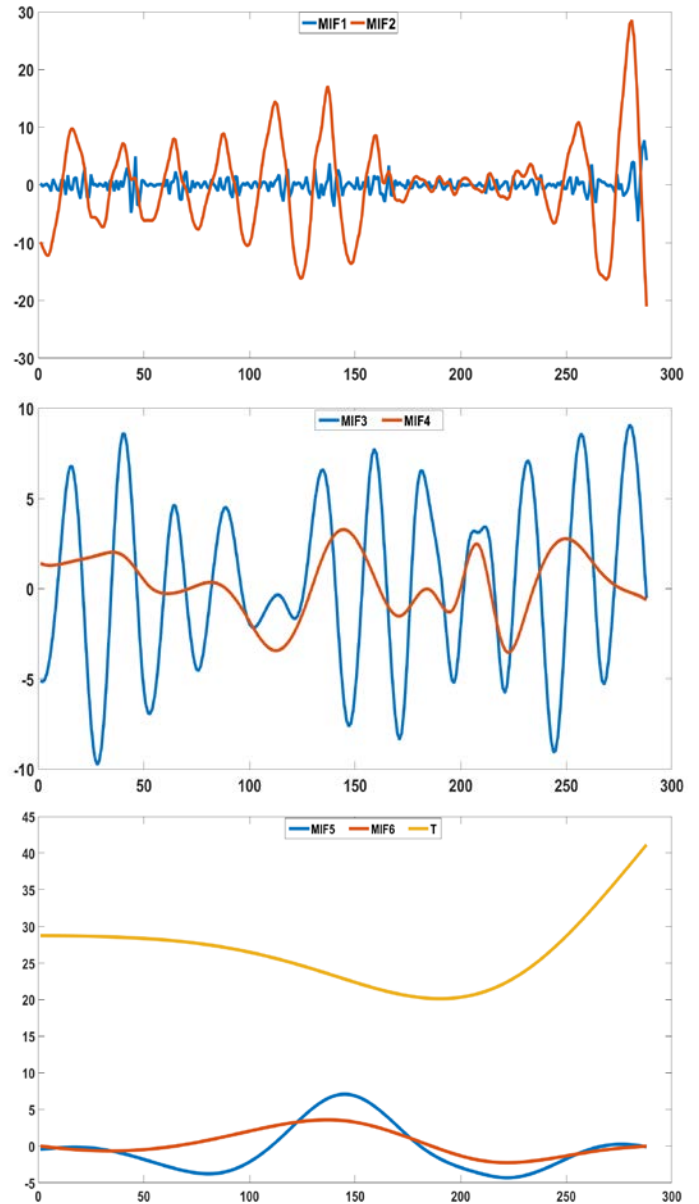


Fig. 3 The result of EEMD electricity price decomposition

TABLE I HISTORICAL DATA AND FITTING VALUES FOR A CERTAIN PERIOD

Period	Original value (\$/MWh)	EEMD-ARIMA prediction (\$/MWh)	ARIMA prediction (\$/MWh)
1:00	12.01	11.26684	13.90
2:00	7.20	8.143687	7.74
3:00	5.77	5.37696	5.18
4:00	5.00	5.951234	4.64
5:00	5.00	4.722827	3.89
6:00	5.00	4.592499	4.54
7:00	5.00	5.724703	5.09
8:00	9.06	9.048185	5.43

TABLE II COMPARISON OF MAPE ERROR BETWEEN TWO MODELS AT THE SAME TIME PERIOD

period	model 1_MAPE	model 2_MAPE
1:00	0.06606	0.1360
2:00	0.3637	0.4104
3:00	0.7648	0.9595
4:00	1.6591	1.5921
5:00	2.3673	3.4219
6:00	3.3732	4.7201
7:00	4.8215	5.6352
8:00	5.5207	11.7959
9:00	6.4396	14.7065
10:00	7.9408	18.0045
11:00	9.1367	21.1873
12:00	10.0938	24.6456
13:00	11.3597	27.8598
14:00	12.7934	30.5383
15:00	13.8196	33.7020
16:00	14.7786	36.7720
17:00	16.0210	40.9595
18:00	17.0757	43.4509
19:00	18.1795	47.6747
20:00	19.8909	52.7640
21:00	21.4698	56.2153
22:00	22.9393	60.5851
23:00	24.6949	67.3694
0:00	25.7861	72.1345

From the electricity price data predicted by the two models in Table 1, we can see that the model 2 is prone to deviations in the prediction of the historical price, especially in the point where the fluctuation of the electricity price is larger. While the history electricity price after the EEMD decomposition is simulated by ARIMA, the time series prediction of the electricity price is closer to the historical data than that of the simultaneous segment, which does not decompose. It can also reflect the volatility of the original electricity price, and basically do not deviate from the track of the historical price. From the comparison of the MAPE errors of the two models in Table 2, it can be found that compared with the results of direct ARIMA model prediction without decomposition, the power price prediction error after decomposing with EEMD is smaller, indicating that the prediction accuracy of model 1 is higher. From the comparison of the two kinds of prediction effect maps in Fig. 5, it can be seen that the results obtained by using ARIMA prediction model 2 directly show that: At each turning point and the extremum of the fluctuation, basically an outlier will be generated, which will deviate from the historical track of historical price. Therefore, in the case of large fluctuations in the short-term price, the forecasting results cannot reflect the authenticity of the original sequence. For Model 1, after EEMD decomposition, the ARIMA model is forecasted. The results show that both the historical electricity price turning point and the fluctuation point can well reflect the volatility of the original sequence, and are closer to the true value of the historical electricity price. Compared with the time series forecasting of historical electricity prices directly, this model reduces the occurrence of abnormal jumping points.

V. CONCLUSION

This paper proposes a new EEMD-ARIMA hybrid forecasting model aiming at the characteristics of market-oriented short-term electricity price fluctuations. Based on this model, the historical electricity prices from May 31 to June 11 in the PJM power market were used as prediction samples and compared with the traditional ARIMA single forecasting model. Based on the proposed hybrid prediction model, the EEMD is first used to decompose to reduce the volatility of the original sequence, and then ARIMA is used to perform sliding prediction on the decomposed curve data. The comparison result shows that the ARIMA time series model is easily affected by many factors, so only a single model such as the time series model can't effectively reflect the historical electricity price characteristics. However, the hybrid prediction model presented in this paper is more accurate than ARIMA time series forecasting, and it is more maneuverable. It also reflects the volatility of the original sequence more realistically and reduces the occurrence of jumping points. Thus, it can be shown that in the context of power system reforms, the EEMD-ARIMA model can obtain higher accuracy power price prediction data with less historical data. The application of the hybrid forecasting model provides a certain forecast reference for the market players trading in the new volatility new energy market.

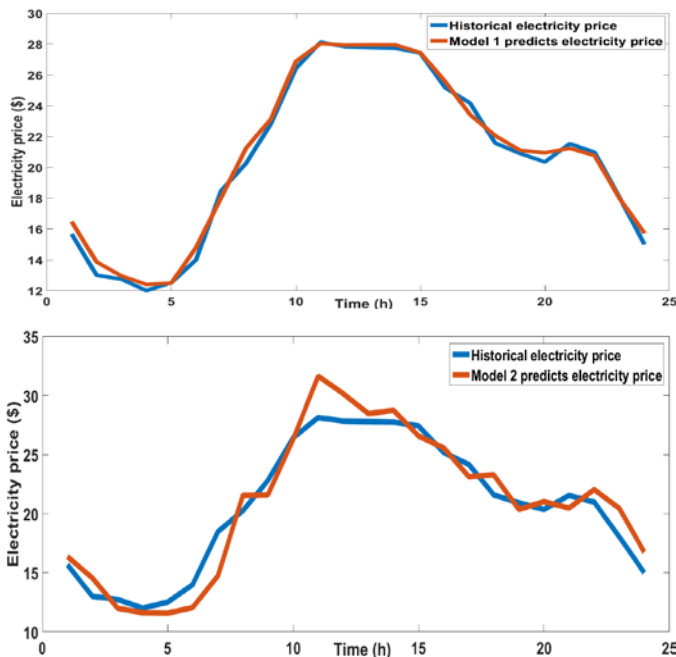


Fig. 4 Comparison of two models in a certain period of time

## REFERENCES

- [1] Song Y B, Yang P, Xu Z R, Huang Y J, Zhang Y J, Yu L, Huang J C, "Unit bidding strategy based on daily electricity price forecasting," Southern Power System Technology, vol. A11, pp. 57-62+70, 2017. In Chinese.
- [2] Niu L X, Wang Z F, Chuan C Z, Shang W L, Zhang S S, "A short-term price mixed forecasting model based on wavelet transform and ARIMA," Journal of Computer Applications, vol. A31, pp. 688-691. 2014. In Chinese.
- [3] Yang J, "Multi-factor gray forecasting model for short-term spot electricity price," Henan Science and Technology, vol.12, pp. 125-126, 2014. In Chinese.
- [4] Deng J J, Huang Y S, Song G F, "Application of non-parametric GARCH based time series model in daily electricity price forecasting," Power System Technology, vol. A36, pp. 190-196. 2012 In Chinese.
- [5] Huang G Q, Yang J J, "Short-term electricity price forecast based on BP neural network and Markov chain," Journal of Shanghai University of Electric Power, vol. A33, pp. 1-3+10.2017. In Chinese.
- [6] Li X B, "Food Combining Prediction Model Based on BP Neural Network and Weighted Fuzzy Markov Chain," Journal of Gansu Lianhe University (Natural Sciences Edition), vol. A27, pp. 8-13, 2013. In Chinese.
- [7] Fata E, Kadota I, Schneider I, "Comparison of Classical and Nonlinear Models for Short-Term Electricity Price Prediction". 2018.