IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 23, 2019, accepted July 29, 2019, date of publication August 5, 2019, date of current version August 19, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2932999

An Optimized Heterogeneous Structure LSTM Network for Electricity Price Forecasting

SIYU ZHOU[1], LIN ZHOU[1], MINGXUAN MAO[1], (Member, IEEE), HENG-MING TAI[2], (Senior Member, IEEE), AND YIHAO WAN[1]

[1]State Key Laboratory of Power Transmission Equipment and System Security, Chongqing University, Chongqing 400044, China
[2]Department of Electrical and Computer Engineering, The University of Tulsa, Tulsa, OK 74104, USA

Corresponding author: Siyu Zhou (zhousy@cqu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 51477021, and in part by the Graduate Research and Innovation Foundation of Chongqing, China, under Grant CYS18007.

ABSTRACT Electricity price is an important indicator of the market operation. Accurate prediction of electricity price will facilitate the maximization of economic benefits and reduction of risks to the power market. At the same time, because of the excellent performance of deep learning models, using long-short term memory neural network (LSTM) and other deep learning models to predict time series has gradually become a research hotspot. In this paper, an optimized heterogeneous structure LSTM model is proposed to solve the problems of the single network structure and hyperparameter selection existing in the current research on LSTM. The heterogeneous structure LSTM is constructed based on the decomposed and reconstructed electricity price data, and sequence model-based optimization (SMBO) is used to optimize its hyperparameters. In order to verify the proposed model, online hourly forecasting and day-ahead hourly forecasting are tested on the electricity markets of Pennsylvania–New Jersey–Maryland (PJM). The experimental results show that the performance of the proposed model is much better than that of the general LSTM model and traditional models in accuracy and stability, providing a new idea for the use of LSTM for time series prediction.

INDEX TERMS Long short-term memory neural network, neural network structure, hyperparameter optimization, time series analysis, electricity price forecasting.

LIST OF ABBREVIATIONS

LSTM	Long-short term memory neural network
SMBO	Sequence Model-Based Optimization
PJM	Pennsylvania–New Jersey–Maryland
EPF	Electricity price forecasting
AI	Artificial intelligence algorithms
ANN	Artificial neural network
SVM	Support vector machine
RNN	Recurrent neural network
GRU	Gated recurrent units
DNN	Deep neural network
CNN	Convolutional neural network
EMD	Empirical mode decomposition
EEMD	Ensemble empirical mode decomposition
IMF	Intrinsic mode function
SVR	Support vector regression
BPNN	Backpropagation neural networks
GTB	Gradient boosting regressor
DTR	Decision tree regressor
LSTM-HF	LSTM model for predicting high-frequency components of sequences
LSTM-LF	LSTM model for predicting low-frequency components of sequences
EEMD-LSTM	The heterogeneous network which consists of LSTM-HF and LSTM-LF is used to predict the high and low-frequency components obtained by EEMD
EEMD-LSTM-SMBO	SMBO is used to optimize the hyperparameter of EEMD-LSTM
APE	Absolute Percentage Error
MAPE	Mean Absolute Percentage Error
RMSE	Root Mean Square Error
IQR	Interquartile Range

The associate editor coordinating the review of this manuscript and approving it for publication was Bilal Alatas.

I. INTRODUCTION

Electricity price is an important factor in the electricity market. Accurate electricity price forecasting (EPF) is critical to all parties in the power market competition [1]. The decision-making in the power market highly relies on the electricity prices, making the EPF model a crucial component for the orderly and efficient operation of the electricity market. EPF has a certain degree of periodicity, it also possesses highly random and time-varying characteristics. Thus, it is a very challenging task to establish a model to predict electricity prices. The accuracy of EPF in the electricity market affects the efficiency and rationality of energy resource optimization. Moreover, accurate prediction of the electricity price enables the balance between power supply and demand, and facilitates the stable operation of the power market [2].

The existing EPF methods include statistical methods [3] and artificial intelligence algorithms (AI). The AI approach can better fit the nonlinear data and has shown good prediction results in electricity price prediction, solar radiation prediction [4], wind power prediction [5], stock market prediction [6] and other fields [7]. The major AI models include the artificial neural network (ANN) [8], [9], support vector machine (SVM) [10], ensemble model [11], and the meta-heuristic optimization algorithm [12].

The ANN and SVM models establish a nonlinear relationship between the output and the input through a large amount of historical data. They transform the dynamic time modeling problem into a static space modeling problem. The electricity price can be modeled as a typical time series. It is nonlinear and has dynamic characteristics, that is, the output of the system is not only related to the input at the current moment but also related to the past inputs. Hence, the traditional artificial intelligence algorithms exhibit limited accuracy for electricity prices prediction.

Recently, deep learning has been widely studied due to its excellent performance. Recurrent neural network (RNN) is a deep learning algorithm with a recurrent feedback network framework. Compared with traditional AI algorithms, it can consider the temporal correlation of time series, which can perform more comprehensive and complete modeling of time series. LSTM [13] is a special RNN model with special structure design, which can effectively avoid the problems of gradient disappearance and gradient explosion in the process of RNN training. For electricity price prediction, LSTM has been used to predict the day-ahead electricity price for the Australian market in the Victoria region and Singapore market. Mean Absolute Percentage Error (MAPE) is improved by 47.3% compared with other traditional machine learning models [14]. LSTM is also used for short-term residential load forecasting [15]. Two other neural network architectures, TD-CNN and C-LSTM, are used to model the load data based on different time dimensions to extract the information contained in the time series [16]. Single gated recurrent units (GRU) network structure is used for prediction. GRU neuron structure is simpler than LSTM, which makes its calculation speed faster [17]. ANN-LSTM hybrid model is utilized to predict the electricity price, that is, the predicted results are the average of the predicted values of ANN and LSTM [18]. Empirical mode decomposition (EMD) is utilized to decompose time series, and LSTM is used to predict each subsequence, which fully explores the connotation information of time series [19]. A large number of experiments were conducted, and four kinds of deep learning network structures deep neural network (DNN), LSTM-DNN, GRU-DNN and convolutional neural network (CNN) were used to predict the electricity price [20].

The aforementioned researches have two main drawbacks. Firstly, the selection of network hyperparameters is based on experience or spending a lot of time experimenting to select parameters [14]–[19], but the hyperparameters obtained by these methods may not be optimal, resulting in the limited performance of models [21]. Secondly, some researches use a single network structure to model time series [14]–[20]. However, a large number of studies have shown that because of the strong volatility and nonlinearity of electricity prices, if the similar time series are decomposed in multiple scales, and the prediction model is adjusted according to the different components with a different nonlinear degree, the prediction performance will be significantly improved compared with sole forecasting models [22]–[25]. Besides, in paper [19], Multiscale components are obtained through EMD, while the sub-sequences (IMFs) are not classified, and all sub-models need to be trained separately, resulting in a huge workload.

This paper proposes an optimized heterogeneous structure LSTM hybrid forecasting model for electricity price prediction. The Intrinsic Mode Functions (IMFs) obtained by ensemble empirical mode decomposition (EEMD) method are reconstructed into high and low-frequency components using the fine to coarse method. Heterogeneous LSTM network structures are designed for different components. The smaller the nonlinear degree of the component is, the simpler the network structure is. Then the SMBO is utilized to optimize the hyperparameters of the heterogeneous LSTM, to further improve the performance of the model. Contributions of this study include:

- Fully explore the information contained in the time series by decomposing and reconstructing the electricity price time series into high and low-frequency components.
- Improve the prediction performance of the forecasting model by designing appropriate LSTM structures for different components, adjusting the nonlinear degree of the network (Construction of heterogeneous structure).
- To further improve the performance and reduce the cost of training, SMBO is utilized to optimize the hyperparameters of deep neural networks.

The remainder of this paper is organized as follows. The characteristics analysis of electricity price is described in Section II. Section III introduces the internal structure of LSTM and the impact of network structure on model performance. Section IV describes the selection of hyperparameters and the SMBO algorithm. Data decomposition and

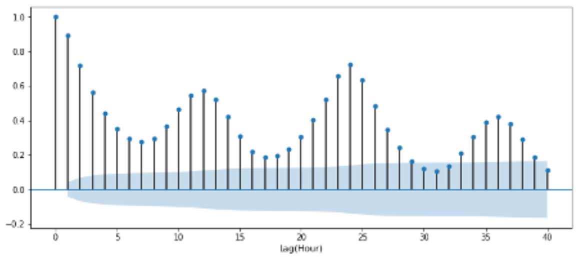

FIGURE 1. **ACF plots of the electricity price.**

reconstruction based on EEMD are presented in Section V. Procedure and design of the proposed model is introduced in Section VI. Section VII describes the settings of models, forecasting results and comparative analysis. Finally, the conclusions are drawn in Section VIII.

II. DATA FEATURES

In the power market, the short-term electricity price change rule is different from the long-term electricity price, and it has the following features:

- Obvious periodicity, generally taking 24 hours as a cycle;
- Strong time regularity and self-correlated to a certain extent;
- Greatly affected by various environmental factors, showing a certain degree of non-stationary randomness.

The LSTM model can make full use of the memory characteristic when predicting the time series with autocorrelation. In each processing process, recursive architecture can keep or pass on the memory information when the weight is updated. Besides, the unit architecture of the LSTM model achieves long-term persistence based on short-term persistence, which is superior to RNN. Therefore, highly accurate prediction can be obtained by utilizing the autocorrelation existing in time series.

Considering the requirement of day-ahead hourly forecasting, it is necessary to continuously forecast the electricity price of the next 24h based on historical information. In this paper, batch forecasting is used, that is, a single forecasting batch is created within the whole forecasting time length, which is different from the forecasting executed in one or more iterative steps and directly predicts the electricity price of the next 24h. However, the batch prediction is only effective if the autocorrelation lasts for more than 24h. As mentioned above, it is necessary to conduct an autocorrelation analysis of electricity price.

The electricity price of the PJM power market from the periods 2018.2-2018.4 is selected [26]. Fig. 1 shows the autocorrelation function graph of electricity price. It can be seen that the autocorrelation coefficient still exceeds the confidence boundary at the 24th order, representing strong autocorrelation. Therefore, we can construct the LSTM model based on the high autocorrelation lag.

III. NEURAL NETWORK STRUCTRUE

A. LSTM MODEL

An LSTM cell includes forgetting gates, input gates, and output gates, and a flow of information representing long-term memory is added to form a black box of input x and state output s. These two features allow LSTM to be trained more effectively so that historical sequence information is fully utilized. The structure of an LSTM cell is shown in Fig. 2.

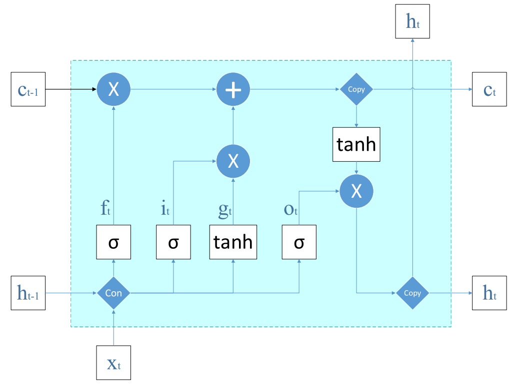

FIGURE 2. **The detailed structure within an LSTM cell.**

x_t is the current input; h_{t-1}/h_t are hidden layer previous state and current state; c_{t-1}/c_t are previous and current cell memory information; W_c, W_i, W_f, W_o are weight matrixes connecting the input signal x and the hidden layer output signal y respectively; b_i, b_c, b_f, b_o are offset vectors; σ is a sigmoid activation function.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{1}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{2}$$

$$\tilde{c}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{3}$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t \tag{4}$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{5}$$

$$h_t = o_t \cdot \tanh(C_t) \tag{6}$$

The weights in the LSTM network can be updated by the input training set to minimize the prediction error. Since LSTM is a deep neural network and the information transmitted is the information flow containing historical memory, key information can be retained by adjusting the weight. Therefore, LSTM has a significant advantage in processing the time series data.

B. NETWORK STRUCTURE CHARACTERISTICS

The nonlinear degree of the neural network is mainly affected by the network structure. Fig. 3 A and B show the single-layer neural network and the two-layer neural network structure respectively. Structure A has two layers, namely the input layer and output layer. The input cell in the input layer only transmits data and does not perform calculations. The output cell in the output layer needs to calculate the input of the previous layer, which is called computing layer, and the network with a computing layer is called "single-layer neural

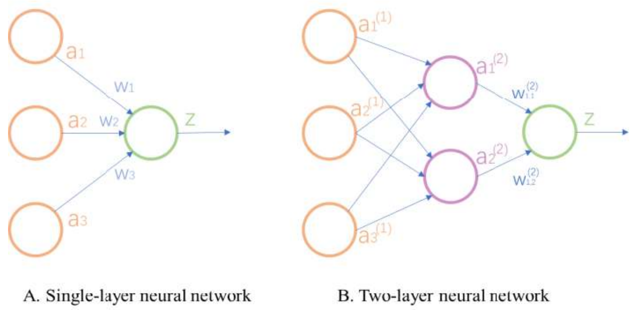

FIGURE 3. Neural network structure.

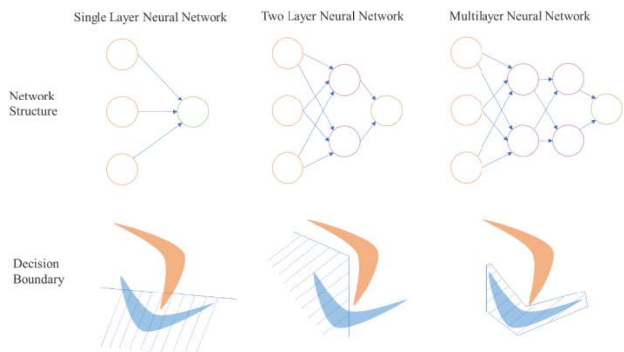

FIGURE 4. The fitting performance of different network structures.

network''. w is the weight value, g is the activation function, and z is the output function.

$$z = g(a_1 \cdot w_1 + a_2 \cdot w_2 + a_3 \cdot w_3) \tag{7}$$

For the two-layer neural network, an intermediate layer is added in addition to the input layer and the output layer. The upper corner mark represents the number of layers, $w_{m,n}^{k}$ represents the weight of the nth neuron in the kth layer connected to the mth neuron in the kth+1 layer, and z is the output function.

$$\begin{aligned} a_1^{(2)} &= g(a_1^{(1)} \cdot w_{1,1}^{(1)} + a_2^{(1)} \cdot w_{1,2}^{(1)} + a_3^{(1)} \cdot w_{1,3}^{(1)}) \\ a_2^{(2)} &= g(a_1^{(1)} \cdot w_{2,1}^{(1)} + a_2^{(1)} \cdot w_{2,2}^{(1)} + a_3^{(1)} \cdot w_{2,3}^{(1)}) \end{aligned} \tag{8}$$

$$z = g(a_1^{(2)} \cdot w_{1,1}^{(2)} + a_2^{(2)} \cdot w_{1,2}^{(2)}) \tag{9}$$

By the formula (7) and (9), it can be seen that the nonlinear degree of the function increases with the increase of the number of layers. As shown in Fig. 4, from the single-layer neural network to the two-layer neural network and then to the multi-layer neural network, the ability to fit the decision-making boundary is continuously enhanced. The essence of the neural network is a method to simulate the real relation function between features and targets. More parameters mean that its fitting function can be more complex and have more capacity to fit the real relation, but it also means that the probability of overfitting increases accordingly, especially when the model is highly complex and the sample data has a low degree of nonlinearity.

Thus, there are two points to note when designing the network structure.

- For high-frequency data with a high degree of nonlinearity, the shallow neural network has poor fitting performance.
- For low-frequency data with a low degree of nonlinearity, the deep neural network has high complexity and is prone to over-fitting, resulting in poor ductility.

For the electricity price data with high complexity and unstable volatility, it is a combination of two types of sequences [22]–[25]. Therefore, using a single structure neural network to fit the electricity price data will restrict the prediction performance to some extent.

In addition to the number of layers, the number of neurons is also an important parameter of the network structure. However, because it is a hyperparameter, the parameter space is large and the change is more flexible, its selection is based on the SMBO introduced in the next section.

IV. OPTIMIZATION OF HYPERPARAMETERS – SEQUENCE MODEL-BASED OPTIMIZATION

The hyperparameters of the deep learning model must be selected properly before the model can be applied to a new data set. Studies have shown that the influence of hyperparameters on different network architectures exhibits a complex relationship. Hyperparameters that provide significant performance improvement in simple networks do not have the same effect to complex architectures. Conventional methods for hyperparameter selection usually depend on prior experience and experimental errors. This leads to two drawbacks. One is that the final model prediction still exhibits low accuracy and unsatisfactory robustness. Another is the computational cost of the algorithm training is high.

The Bayesian optimization, known as active optimization, uses a surrogate model to fit the real objective function and proactively selects the most ''potential'' evaluation points based on the fitting results to avoid unnecessary sampling. The Bayesian optimization framework effectively employs the complete historical information to improve the search efficiency and can be expressed as

$$p(f|D_{1:t}) = \frac{p(D_{1:t}|f)p(f)}{p(D_{1:t})} \tag{10}$$

In (11), f denotes an unknown objective function; $D_{1:t} = \{(x_1, y_1), (x_2, y_2), \ldots, (x_t, y_t)\}$ denotes the observed set which is composed of different hyperparameter sets and corresponding prediction results, x_t are the decision vectors, $y_t = f(x_t) + \varepsilon_t$ denotes the observed value, ε_t denotes the observation error, $p(D_{1:t}|f)$ denotes the likelihood distribution of y; $p(f)$ denotes the prior probability distribution of f; $p(D_{1:t})$ represents the marginal likelihood distribution of f, and it is mainly used to optimize hyperparameter in Bayesian optimization; $p(f|D_{1:t})$ represents the posterior probability distribution of f, and it describes the confidence of the unknown objective function after correction by the observed data set.

The SMBO is used in this study which is an improved algorithm based on Bayesian optimization. The SMBO method builds a model based on historical data to evaluate the performance of the hyperparameters, then selects new hyperparameters based on this model for testing. Replacing the previous distribution with a non-parametric density means to obtain a posterior probability distribution $p\left(f|D_{1:t}\right)$ containing more data information according to formula (8).

The acquisition function $\alpha : \chi \times \mathbb{R} \times \Theta \rightarrow \mathbb{R}$ is an active strategy for selecting the next evaluation point. This function maps from input space $\mathcal{X}$, the observation space $\mathbb{R}$, and the hyperparameter space Θ to the real space. It is constructed from the posterior distribution of the observed data set $D_{1:t}$ and is maximized as a reference for selecting the next evaluation point

$$x_{t+1} = \max_{x \in \chi} \alpha(x; D_{1:t}) \tag{11}$$

The expected improvement (EI) strategy [27] is adopted to quantify the probability so that the observed value of x could improve the current optimal objective function value and reflect different improvement amount as well. The acquisition function of the EI strategy can be expressed as

$$EI_{y^*}(x) = \int_{-\infty}^{y^*} (y^* - y)p(y|x)dy \tag{12}$$

y^* is the threshold of the objective function, x is the suggested hyperparameter set, and y is the actual value of the objective function using the hyperparameter set x. The goal is to maximize the expected improvement of x, that is to find the best hyperparameter under the surrogate function $p(y|x)$.

The optimization framework is an iterative process (as shown in Algorithm 1.), consisting of three steps. Step 1 selects the next most potential point x_t according to the maximum acquisition function. Step 2 evaluates the objective function value $y_t = f(x_t) + \varepsilon_t$ according to the selected evaluation point x_t. Step 3 adds the new input-observation value $\{x_t, y_t\}$ to the historical observation set $D_{1:t-1}$ and updates the probabilistic surrogate model to prepare for the next iteration.

Algorithm 1 Hyperparameter Optimization

1: for t = 1, 2 . . . do
2: Maximize the acquisition function to get the next evaluation point: $x_t = \arg\max_{x \in \chi} \alpha\left(x|D_{1:t-1}\right)$
3: Evaluate the objective function value $y_t = f(x_t) + \varepsilon_t$;
4: Data Integration:$D_t = D_{t-1} \cup \{x_t, y_t\}$, Update probabilistic surrogate model
5: end for

V. DECOMPOSITION AND RECONSTRUCTION OF SEQUENCES

This section describes the method of preprocessing electricity price data. The essence includes the decomposition and reconstruction of the sequence using the EEMD algorithm and the fine-to-coarse scheme to avoid the excessive decomposition.

A. EEMD

EMD is based on the cycling alternation algorithm [28]. However, due to the mode overlapping in EMD, an IMF has multiple sequences with different frequencies, or sequences with the same frequency may appear in different IMFs. The EEMD algorithm was proposed to overcome these defects, it adds the white noise to the sequence, obtaining the local mean of the upper and lower envelopes accurately, thereby promoting anti-aliasing decomposition and avoiding the mode mixing defects [29]. Let $X(t)$ be the original time series, then the EEMD process can be described by the following steps.

1) Add a set of white noise sequence $\omega(t)$ to the original time series $X(t)$ to obtain a new set of sequence $x(t)$.

$$x(t) = X(t) + \omega(t) \tag{13}$$

2) Identify the local minima and local maxima in the original time series $x(t)$.
3) Determine the upper envelope $e_{max}(t)$ of $x(t)$ according to all the maximum points in the sequence $x(t)$ using the cubic spline interpolation function. Similarly, the lower envelope $e_{max}(t)$ of the sequence can be determined according to the corresponding minimum points.
4) Remove the low-frequency sequence by subtracting the sequence $m(t) = (e_{max}(t) + e_{min}(t))/2$.
5) Replace the original time series $x(t)$ with the newly created $d(t) = x(t) - m(t)$ and repeat Step 1 to Step 3 until the $d(t)$ is an IMF. Denote $d(t)$ as $c_1(t)$ and let the residual $r(t) = x(t) - c_1(t)$ be the new $x(t)$;
6) Repeat steps 1) to 4) to filter out the multiple IMFs until the new IMF can no longer be filtered out from $X(t)$. Under this circumstance, the original time series $X(t)$ decomposed into multiple IMFs and a trend term component $r(t)$ and can be expressed as

$$x(t) = \sum c_i(t) + r(t) \tag{14}$$

B. FINE TO COARSE

The obtained IMFs are reconstructed into high and low-frequency sequence by the fine-to-coarse method to avoid the accumulation of errors caused by too many subsequences and the excessive decomposition [30]. It has a positive effect on preventing error propagation caused by multiple steps. Suppose that the original time series $x(t)$ is decomposed into N IMFs. Procedure of the reconstruction is:

1) Let $s_i = \sum_{k=1}^{i} IMFs, i = 1, 2, \ldots, N$
2) Using the T-test method to determine the corresponding value of i when the mean value of s_i is significantly different from 0;
3) Since each IMF is independent and orthogonal to each other, IMF_1 to IMF_i can be added up to be the

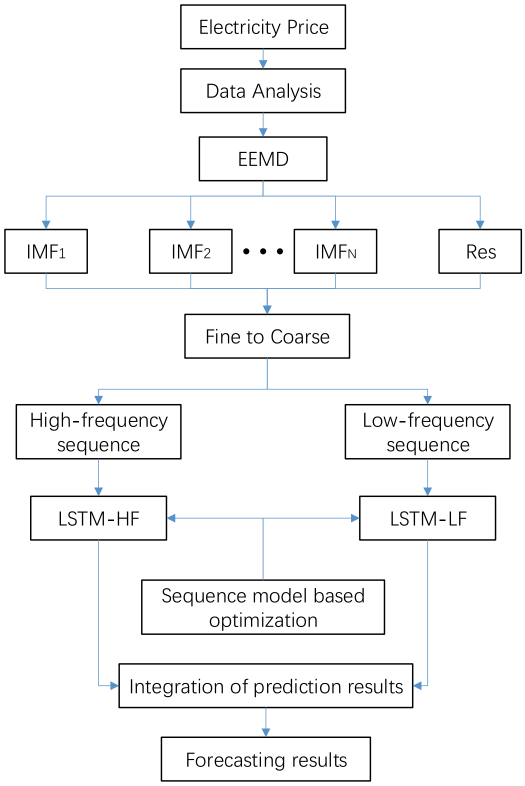

FIGURE 5. The proposed electricity price forecasting model.

high-frequency sequence and the sum of IMF_i to IMF_N can be used as the low-frequency sequence.

VI. THE PROPOSED ELECTRICITY PRICE FORECASTING MODEL

The proposed electricity price forecasting method is shown in the flowchart of Fig. 5. Detailed procedure is given below.

1) The autocorrelation analysis is conducted to determine whether the price data could be predicted in batches based on the LSTM model, and to provide a reference for determining the looking back steps of LSTM model.
2) According to the characteristics of non-linear, non-stationary and multi-frequency superposition of electricity price, EEMD is used to decompose the electricity price sequence.
3) Concerning the IMFs and Res. obtained by decomposition, fine to coarse method is used to classify them into high and low-frequency components.
4) LSTM network structure is adjusted according to the non-linearity of high and low-frequency sequences. LSTM-HF and LSTM-LF are obtained respectively.
5) Based on the network structure designed in step 3, the hyperparameters of LSTM-HF and LSTM-LF are optimized by SMBO to further improve the performance of the model
6) The predicted values of high and low-frequency sequences are added together to obtain the final predicted value.

VII. CASE STUDIES

The data of the electricity price in the last weeks of March, June, September, and December are utilized as the test set. The electricity price of the test set for the past half-year before the corresponding months is utilized as the training set and the data for the whole year of 2017 is utilized as the validation set for parameter optimization. In addition to the proposed method, the models used for performance comparison include support vector regression (SVR), backpropagation neural networks (BPNN), gradient boosting regressor (GTB), and decision tree regressor (DTR), LSTM series models include shallow LSTM, stacked LSTM, EEMD-LSTM and EEMD-LSTM-SMBO.

The experiment includes the on-line hourly forecasting which is the prediction of the electricity price in the next hour with the prediction time step of 1 hour and the day-ahead hourly forecasting which is a batch prediction of the electricity price in the next 24 hours, with the prediction time step of 24 hours. Performance of the proposed forecasting method is evaluated by the following three metrics: Absolute Percentage Error (APE), Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE). These metrics can be expressed as

$$APE = \frac{|y_i - \widehat{y}_i|}{y_i} \times 100\% \tag{15}$$

$$MAPE = \frac{1}{N}\sum\nolimits_{i=1}^{N} \frac{|y_i - \widehat{y}_i|}{y_i} \times 100\% \tag{16}$$

$$RMSE = \sqrt{\frac{1}{N}\sum\nolimits_{i=1}^{N} (y_i - \widehat{y}_i)^2} \tag{17}$$

N is the number of predicted points; y_i is the actual price value of the i^{th} predicted point; $\widehat{y}_i$ is the predicted electricity price value of thei^{th} predicted point.

A. DATA PREPROCESSING

The activation function of the neurons is sensitive to whether the data are in [-1,1]. The efficiency and the effectiveness of the neural network training will be greatly accelerated if values of the input data are in the interval of [-1,1]. Thus, it is preferred to normalize the input data. The normalized input variable x' can be expressed as

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{18}$$

where max(x) and min(x) are the maximum and minimum values of the input variable, respectively.

In the next stage, data reconstruction, data from February to April are selected to show the process. The normalized electricity price data are decomposed into nine independent IMFs and one trend component Res, as shown in Fig. 6. Plots of IMF_1 to IMF_6 denote the high-frequency sequences and that of IMF_7 to IMF_9 constitute the low-frequency sequences. It can be seen from Fig. 6 that fluctuation of the IMFs becomes smaller as the index of IMF are larger. Then, the fine-to-coarse is applied to reconstruct the sub-sequences.

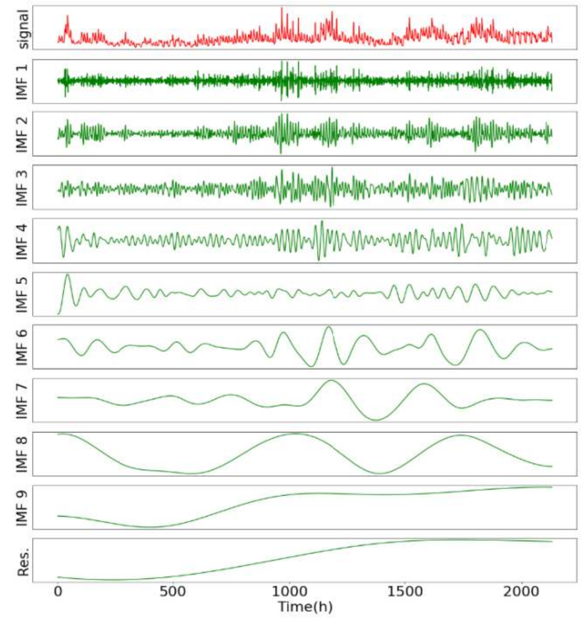

FIGURE 6. The EEMD results of electricity price.

The reconstructed high and low-frequency sequence results are shown in Fig. 7.

B. ALGORITHM SETTINGS

Traditional machine learning models include SVR, BPNN, GTB, and DTR. Grid search and k-fold cross-validation are used for the selection of traditional machine learning models' parameters. It is worth noticing that although the type of research data is time series, in order to use the machine learning model, time series prediction problem should be re-framed as supervised learning problems. Time series data are converted to supervised data during the data processing stage. K-fold cross-validation is therefore used and divides all the samples into K groups called folds. K-1 folds are used for training, the K fold is used for testing and this process would be performed K times. That is, search in the hyper-parametric domain specified in grid search and take k-fold cross-validation results as corresponding parameters evaluation. Finally, the optimal parameters are obtained through this process.

Based on the above methods, SVR selects the Gaussian kernel function and the loss function is quadratic function. The BPNN utilizes a single hidden layer structure. The number of neurons in the hidden layer is 30, the activation function of the hidden layer is the tanh function and the activation function of the output layer is a linear function. The learning rate is 0.001 and the learning target is 0.01. The loss function is defined as the MSE and the Adam algorithm is used. For GTB, the number of boosting stages is 300, the maximum depth of estimators is 4, the minimum number of samples required to be at a node is 1, the learning rate is 0.01 and the loss function is the least square method. For DTR, the maximum depth is 3, the minimum number of samples required to be at a node is 1, the learning rate and loss function are the same as GTB. Their inputs are all the electricity price of the two days before the prediction date.

LSTM series models include shallow LSTM (Since the commonly used LSTM models are shallow structures, LSTM is used to refer to the shallow LSTM in the following results presentation and description), stacked LSTM, EEMD-LSTM, and EEMD-LSTM-SMBO. The purpose of using both shallow and stacked LSTM is to test the content of Section III.B, that is, neural networks with different depths perform variously on data sets with different characteristics. EEMD-LSTM combines the first two single network structures for different components, then for EEMD-LSTM-SMBO, SMBO is used to further improve the performance.

The preliminary parameter design of each LSTM model is established. The number of neurons is generally a power of 2, and the number of neurons in the first layer cannot be too small, otherwise it is impossible to learn the rule of data set. Based on the above considerations, Shallow LSTM adopts two-layer network structure, the first layer contains 128 neurons, the second layer 64 neurons. The tanh function is selected as the activation function the Dropout set to 0.2, the number of iterations is 150, and the optimization algorithm is Adam, the batch size is 64, the number of epochs is 150. As for the determination of lag order, there are no clear rules for reference at present, which are mainly determined according to specific application scenarios. Considering that the experiment includes continuous prediction of data of the next 24 orders, the lag order of LSTM should be greater than 24. Meanwhile, according to the application experience of LSTM in various fields, the order should not exceed the required order too much. Combined with the autocorrelation analysis in section II, the 36th order with high correlation coefficient is finally selected as the lag order. For Stacked LSTM, it adopts a six-layer network structure. From shallow to deep, the number of neurons in each layer is 128, 128, 64, 64, 32, 32, and the dropout of each layer is set at 0.3. Other parameters are the same as those of the shallow LSTM.

After the price sequence is processed by EEMD, the electricity price is reconstructed into high and low-frequency components using the Fine to coarse method. EEMD-LSTM is utilized, for low-frequency components, to reduce the nonlinearity of the model, the structure of LSTM needs to be simplified, so shallow LSTM (LSTM-LF) is used to predict this part. For high-frequency component, stacked LSTM (LSTM-HF) is used. The batch size is 64 and the number of epochs is 150. Besides, the learning rate in the LSTM series model is set to make the adaptive adjustment according to the value loss at each training. The number of acceptances of model performance not improved is set as 5, each time reduced by 0.2. The learning rate is at least 0.001.

For the optimized LSTM model, based on the above network architecture, SMBO is used to search for optimal parameters in the preset hyperparameter space to optimize the LSTM model, and the preset hyperparameter space is shown in Table 1.

TABLE 1. Hyperparameter search space.

The number of neurons in hidden layer	16, 32, 64, 128, 256
Dropout	Between 0 and 0.5
Batch size	16, 32, 64, 128, 256
Epochs	50, 100, 150, 200

TABLE 2. Hyperparameters of EEMD-LSTM-SMBO.

EEMD-LSTM-SMBO			
LSTM-HF		LSTM-LF	
lstm-1: 128	dropout-1: 0.342	lstm-1: 64	dropout-1: 0.297
lstm-2: 64	dropout-2: 0.110	lstm-2: 32	dropout-2: 0.152
lstm-3: 64	dropout-3: 0.207		
lstm-4: 128	dropout-4: 0.124		
lstm-5: 64	dropout-5: 0.089		
lstm-6: 32	dropout-6: 0.116		
batch size: 64		batch size: 32	
Epochs: 200		epochs 150	

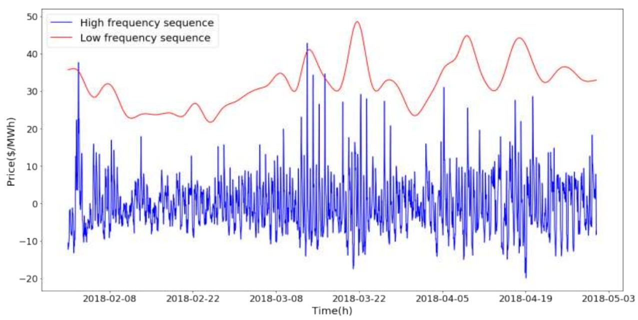

FIGURE 7. High and low-frequency sequences obtained after IMFs reconstruction.

The network structures modified by SMBO are shown in Table 2 (for lstm-i, dropout-i, i represents the i^{th} layer):

C. ON-LINE HOURLY FORECASTING

Fig. 8 shows the daily mean values of MAPE for different models in different months. It can be seen that BPNN and DTR models are unstable in continuous prediction. Their prediction accuracy at 3.30, 6.26 and 9.29, etc. is close to that of LSTM series models, but there is a big deviation on other days. The LSTM series models show good stability, but further observation reveals that single LSTM models with different structures perform dissimilarly. For example, according to the result on March 27-31, the Stacked LSTM outperforms the Shallow LSTM. While, in terms of the results at 6.25, 6.28 and 9.27, the Shallow LSTM's performance is better, and the Stacked LSTM becomes inaccurate, even inferior to the traditional machine learning models.

Table 3 and 4 show the comparison of the various models by seven-day averages of MAPE and RMSE in different months. It can be seen that the proposed method is superior to other LSTM series models and traditional machine learning models throughout the whole four stages. Combining Fig. 8 with Table 3 and 4, we can draw the following conclusions.

The single LSTM models could not maintain high accuracy on electricity price forecasting. The MAPE of Stacked LSTM in march is 6.28% lower than that of shallow LSTM, and RMSE decreases by 0.22%. While, in terms of the metric evaluation in June, September, and December, the shallow LSTM performs better, and the MAPE decreases by 30.60%, 7.24%, 44.74%, and the RMSE decreases by 16.11%, 4.67% and 33.60% respectively which means that the single LSTM model can only show good performance on data within some specific stages, but for other data with a large difference in nonlinear degree, it's accuracy is inferior to other LSTM models with different network structure. Besides, on the whole, the shallow LSTM has more advantages over the stacked LSTM in terms of electricity price prediction.

The performance of EEMD-LSTM is much better than single LSTM models, compared with shallow LSTM and stacked LSTM, the MAPE decreases by 1%∼10% and 10%∼50% respectively, and RMSE decreases by 1%∼5% and 1%∼40% respectively. It indicates that EEMD-LSTM with heterogeneous network structure can effectively overcome the poor stability of single network architecture, and can achieve good adaptability for data sets with complex characteristics.

After the optimization by SMBO, the performance is further improved. According to the result in September, the accuracy of LSTM series models in this stage is poor, which may be caused by the improper selection of hyperparameters. However, after optimization by SMBO, the model performance is greatly improved. It can be seen from metric evaluation in different months, the MAPE decreases by 0.36%, 2.30%, 15.08%, 4.78%, and RMSE decreases by 4.95%, 16.92%, 8.23%, and 2.20%, respectively. Also, the performance of LSTM series models is far better than traditional machine learning models generally, the MAPE of LSTM models is generally reduced by more than 20%.

For further comparison, Fig. 9 shows the plots of results for various methods on different days. It can be seen that when the price fluctuation is relatively small, the predicted values of the four models are close to the actual values, as shown in Fig. 9(b) 6:00-16:00. When the electricity price fluctuates violently, there is a large deviation on the predicted values of BPNN and SVR models, but LSTM series models can still maintain a certain high accuracy, as shown in Fig. 9(a) 12:00-20:00 and Fig. 9(d) 16:00-21:00. When the electricity price continues to fluctuate violently, the prediction bias of single LSTM models increases, but the proposed algorithm can still show excellent prediction accuracy under such severe conditions, as shown in Fig. 9(c) 06:00-20:00.

To further validate the capability of EEMD-LSTM-SMBO in forecasting electricity prices, the prediction error distribution of 24 hours of the LSTM and the proposed model in different months, are examined and shown

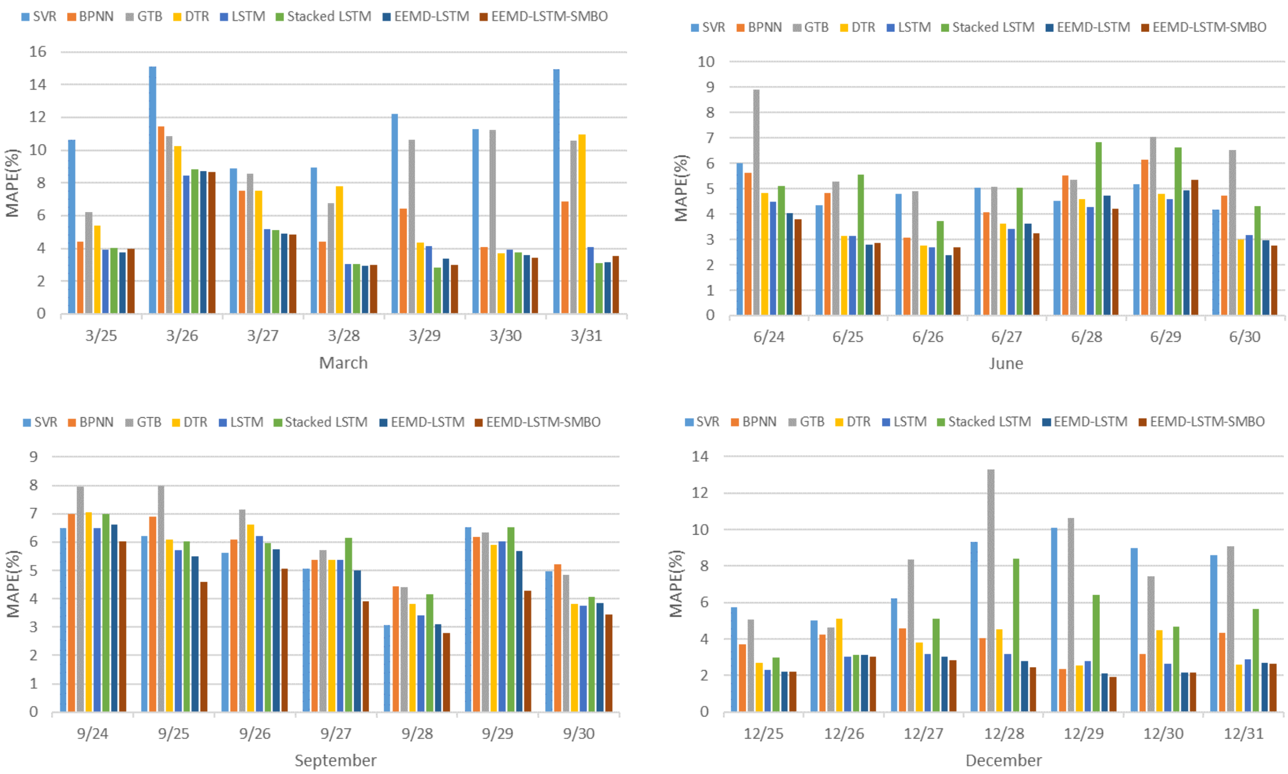

FIGURE 8. MAPE for different models in different month.

TABLE 3. Comparison of the on-line hourly forecasting various models by seven-day averages of MAPE in different months.

Month	SVR	BPNN	GTB	DTR	LSTM	Stacked-LSTM	EEMD-LSTM	EEMD-LSTM-SMBO
March	11.71	6.46	9.25	7.14	4.69	4.40	4.36	**4.34**
June	4.87	4.86	6.16	3.82	3.69	5.32	3.64	**3.56**
September	5.42	5.88	6.34	5.52	5.28	5.70	5.07	**4.31**
December	7.72	3.78	8.37	3.70	2.87	5.20	2.60	**2.47**

TABLE 4. Comparison of the on-line hourly forecasting various models by seven-day averages of RMSE in different months.

Month	SVR	BPNN	GTB	DTR	LSTM	Stacked-LSTM	EEMD-LSTM	EEMD-LSTM-SMBO
March	3.14	2.19	2.97	3.16	1.89	1.90	1.82	**1.73**
June	1.80	2.00	2.35	1.69	1.70	2.02	2.01	**1.67**
September	2.22	2.48	2.95	2.43	2.34	2.45	2.31	**2.12**
December	1.90	1.17	2.18	1.37	0.95	1.42	0.91	**0.89**

in Fig. 10. Combined with Fig. 9, it can be found that when severe fluctuation occurs, as illustrated in Fig. 10(c) 06:00-20:00, or evident trend of the data occurs, as illustrated in Fig. 10(d) 00:00-7:00 and Fig. 10(a) 00:00-6:00, the median and IQR (interquartile range) of proposed model's APEs are lower than those of the LSTM model. Small IQR reflects concentrated error distribution, which indicates good stability.

When the data changes gently, as shown in Fig. 10(b) 00:00-18:00, the proposed model can achieve almost the same excellent performance as LSTM (As mentioned above, the shallow LSTM has high prediction accuracy on data sets with low nonlinear degree), the median of APEs is less than 0.1% throughout the period.

Besides, Fig. 10(c) 6:00-7:00 shows that the trend of data changes suddenly, the proposed model can quickly respond to

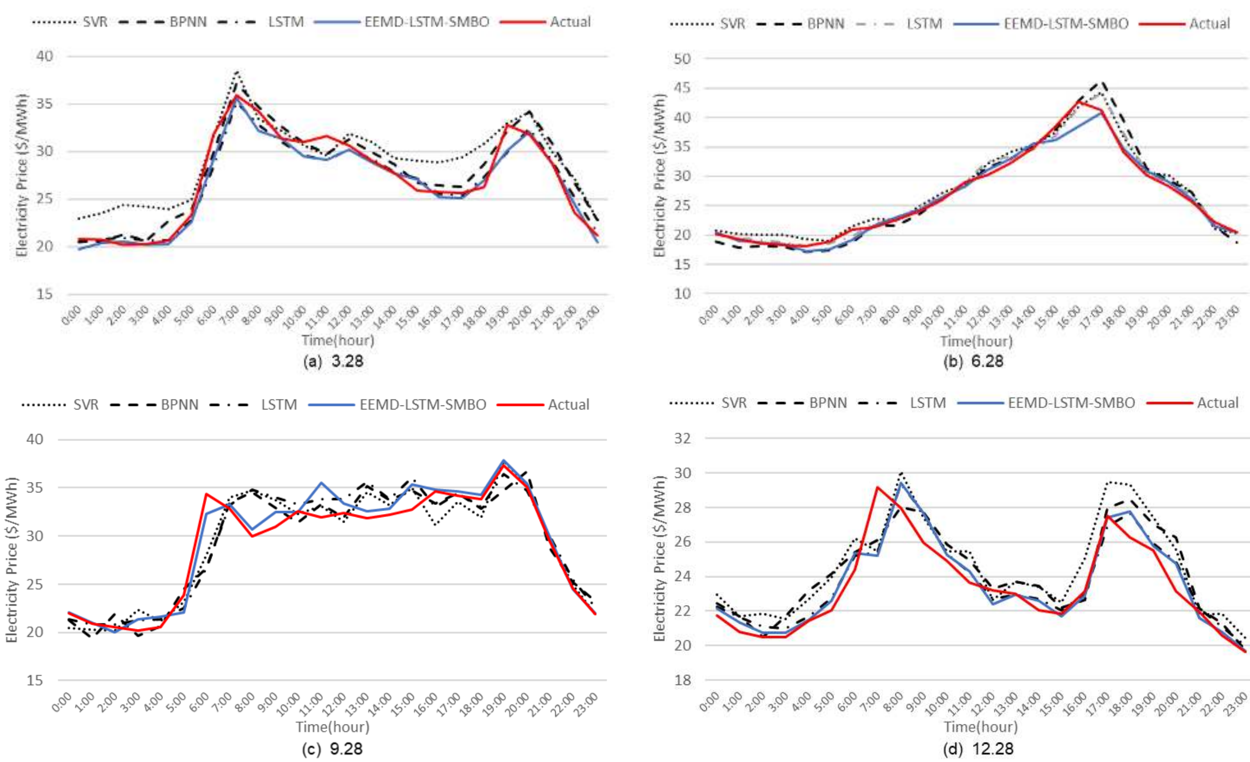

FIGURE 9. Forecasting results of different models on different days.

the changes of sequence which greatly improve the accuracy. However, we found that after a period of significant fluctuation, as shown in Fig. 10(c) 19:00-20:00, the median of proposed model's APEs remains at a low level, but the IQR suddenly increases, and the stability of the model is poor in this period. While it is worth noting that the proposed model is affected by this special case for a very short time and can quickly restore high stability.

D. DAY-AHEAD HOURLY FORECASTING

The data set remains unchanged, and the training and test sets are selected as well as the on-line hourly forecasting, However, models are set to forecast electricity prices of 24 hours based on the input data.

Table 5 and 6 show the MAPE and RMSE of different models in different months, and they are similar to on-line hourly forecasting. The performance of traditional machine learning models is worse than LSTM series models, the MAPE of SVR, BPNN, GTB and DTR increases by more than 30%, the RMSE generally increases by more than 20%. Furthermore, the heterogeneous LSTM (EEMD-LSTM) greatly improves the performance of single LSTM models, and after the optimization of hyperparameters by SMBO, the EEMD-LSTM-SMBO offers lowest MAPE and RMSE in general, maintaining high accuracy and stability.

Fig. 11 shows the error distribution of different models of different months, the median of Stacked LSTM's APEs is around 0.045, which is about 30% higher than that of shallow LSTM. The median of EEMD-LSTM's APEs is 25% and 45% lower than those of shallow LSTM and Stacked LSTM, respectively. Although the prediction accuracy of EEMD-LSTM is greatly improved compared with that of single LSTM, the IQR still fluctuates in a wide range from 0.028 to 0.061. The IQR of June and September is 0.045 and 0.061 respectively. However, after the hyperparameter optimization by SMBO, the IQR of these two months decreases significantly by 15.6% and 19.7% respectively. The median of APEs of the proposed model is less than 0.03 in the whole period, and the IQR is between 0.025 and 0.05, which means that it can maintain high prediction accuracy and stability over complex price data sets.

Fig. 12 compares the MAPE averages and the RMSE averages for four weeks in day-ahead hourly forecasting and on-line hourly forecasting. It can be seen that the accuracy of models performed in the day-ahead hourly forecasting has a different degree of decline compared with on-line hourly forecasting. The following states the observation from these figures.

- The conventional machine learning model transforms the regression problem of dynamic time series into the static space modeling problem, while the LSTM network directly carries out dynamic modeling of time series. The latter makes full use of historical data and is more suitable for the electricity price forecasting.

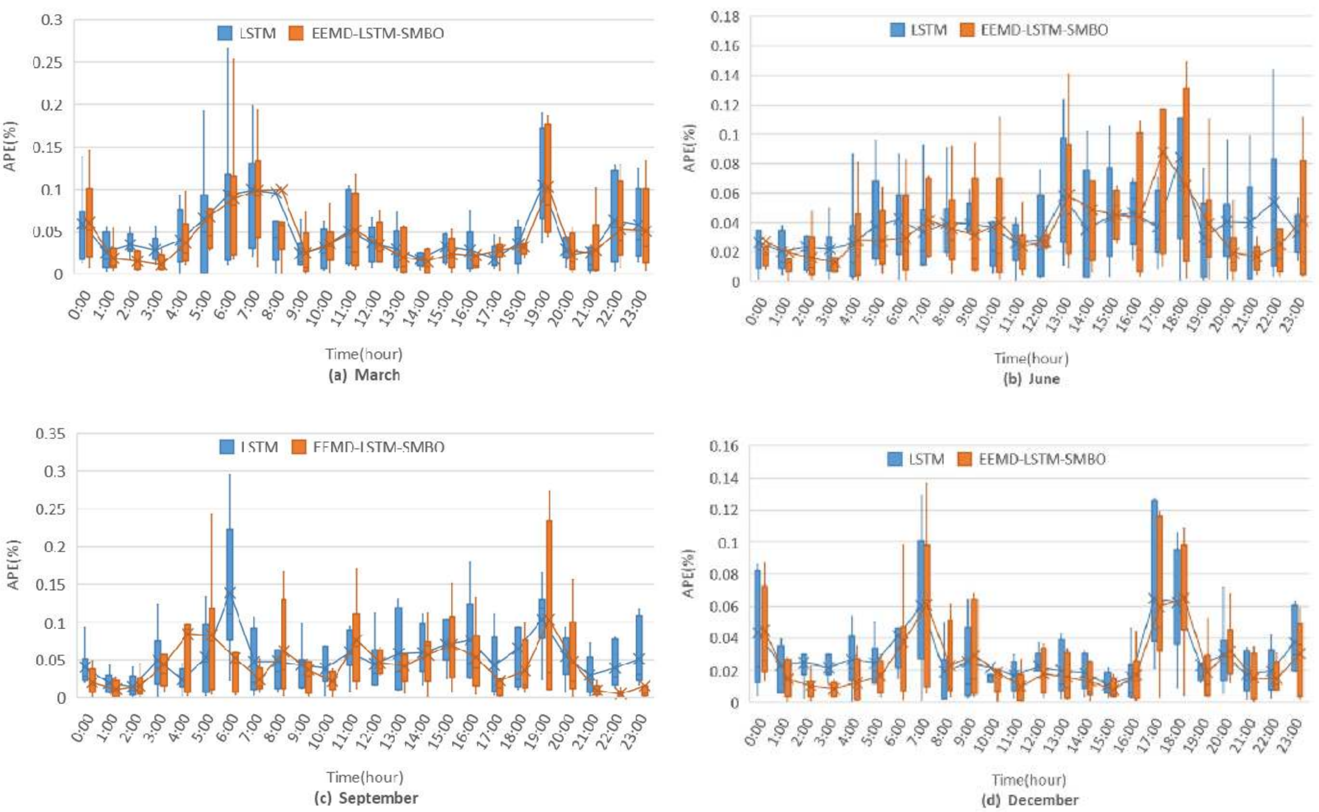

FIGURE 10. The error distribution of 24-hour of LSTM and EEMD-LSTM-SMBO in different months.

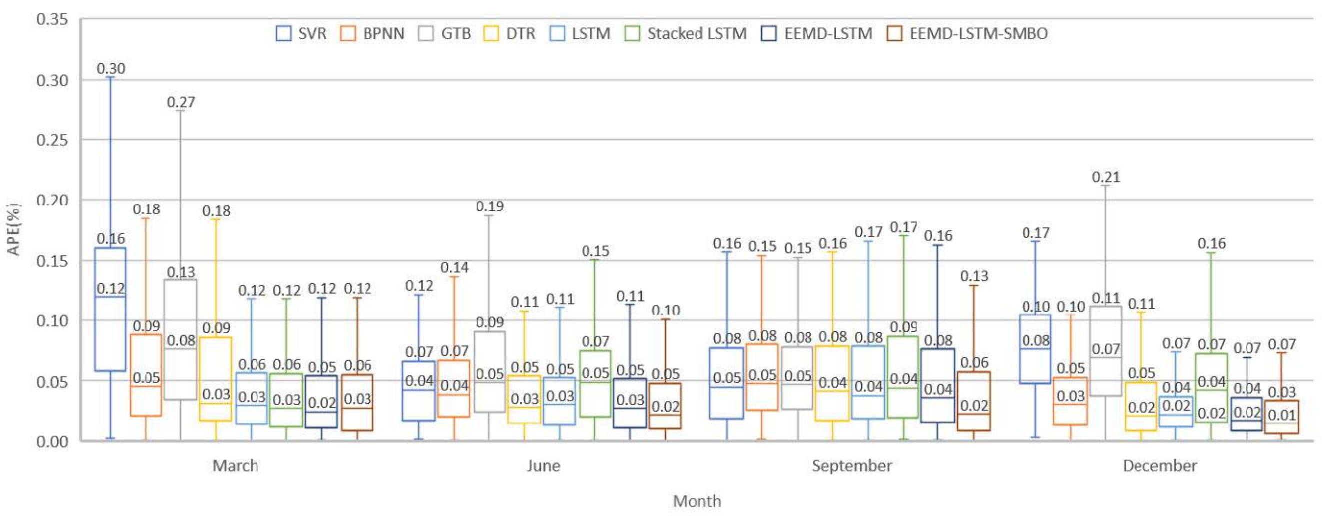

FIGURE 11. Error distribution of day-ahead hourly forecasting for different models.

- No matter the forecasting is the day-ahead hourly or the on-line hourly, the heterogeneous structure is effective. Both the MAPE and RMSE of the EEMD-LSTM are much lower than single LSTM. In addition, the SMBO can effectively optimize the hyperparameters of LSTM. Experimental results show that the use of SMBO lowers the MAPE of EEMD-LSTM by about 7%. The most important is that EEMD-LSTM-SMBO can maintain high accuracy in both cases, solving the problem of poor stability of single LSTM model, and achieve good adaptability of electricity price sequence data with complex nonlinear characteristics.

TABLE 5. Comparison of the various day-ahead hourly forecasting models by seven-day averages of MAPE in different months.

MAPE	SVR	BPNN	GTB	DTR	LSTM	Stacked LSTM	EEMD-LSTM	EEMD-LSTM-SMBO
March	12.03	6.71	9.52	7.65	4.78	4.66	4.57	**4.53**
June	5.15	5.06	6.40	4.03	3.78	5.42	3.78	**3.67**
September	5.81	6.12	6.67	5.81	5.59	5.95	5.23	**4.77**
December	8.05	3.80	8.92	3.86	2.97	5.57	2.66	**2.51**

TABLE 6. Comparison of the various day-ahead hourly forecasting models by seven-day averages of RMSE in different months.

RMSE	SVR	BPNN	GTB	DTR	LSTM	Stacked LSTM	EEMD-LSTM	EEMD-LSTM-SMBO
March	3.19	2.25	3.06	3.44	2.03	2.07	1.98	**1.96**
June	1.87	2.12	2.44	1.78	1.78	2.05	1.91	**1.77**
September	2.40	2.63	3.16	2.57	2.49	2.58	2.47	**2.21**
December	1.96	1.26	2.28	1.43	1.04	1.48	0.97	**0.93**

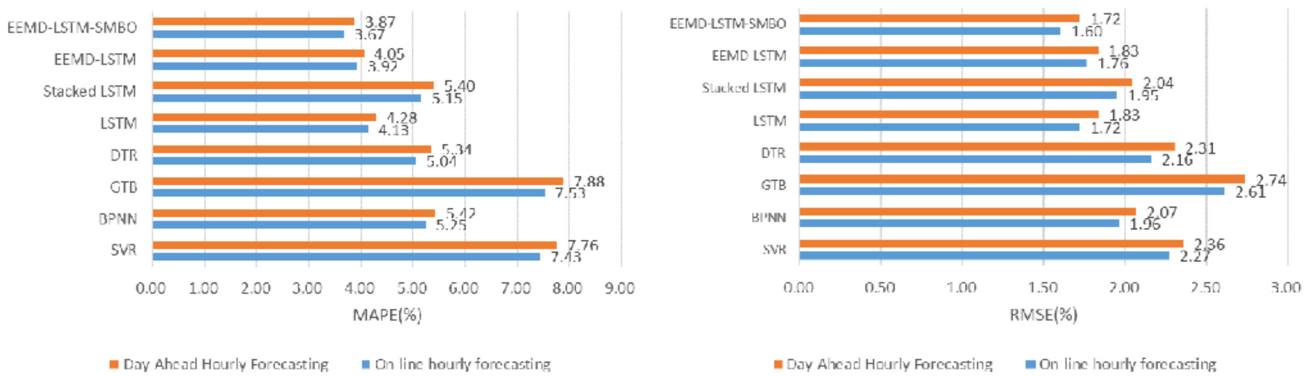

FIGURE 12. Comparison of MAPE, RMSE in day-ahead hourly and on-line hourly forecasting of different models.

VIII. CONCLUSION

This paper focuses on the multi-frequency analysis of the target time series, the non-linearity of deep neural network structures and the selection of hyperparameters, a heterogeneous structure LSTM with SMBO has been presented to predict the electricity price, which provides a new idea for the use of LSTM to predict time series.

The online hourly forecasting and the day-ahead hourly forecasting were performed. Simulation results demonstrate that LSTM outperforms traditional machine learning models greatly, the heterogeneous LSTM fully exploits the hidden information in the electricity price and overcomes the poor stability of single LSTM. The distribution of prediction errors of EEMD-LSTM-SMBO is more concentrated, which shows that the proposed model is more stable, and is more suitable for practical applications.

This study analyzes the price sequence without considering the influence of exogenous variables such as load and holidays. Moreover, the chosen LSTM network architecture is the basic sequential structure. In the future research, to further improve the performance of the model, a variety of variables can be considered to enrich the input of the model and some new network architectures can be tried in the selection of deep network structure.

REFERENCES

[1] D. S. Kirschen, "Demand-side view of electricity markets," *IEEE Trans. Power Syst.*, vol. 18, no. 2, pp. 520–527, May 2003.

[2] R. Weron, "Electricity price forecasting: A review of the state-of-the-art with a look into the future," *Int. J. Forecasting*, vol. 30, no. 4, pp. 1030–1081, 2014.

[3] A. J. Conejo, M. A. Plazas, R. Espinola, and A. B. Molina, "Day-ahead electricity price forecasting using the wavelet transform and ARIMA models," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 1035–1042, May 2005.

[4] G. Capizzi, C. Napoli, and F. Bonanno, "Innovative second-generation wavelets construction with recurrent neural networks for solar radiation forecasting," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 11, pp. 1805–1815, Nov. 2012.

[5] R. Jiao, X. Huang, X. Ma, L. Han, and W. Tian, "A model combining stacked auto encoder and back propagation algorithm for short-term wind power forecasting," *IEEE Access*, vol. 6, pp. 17851–17858, 2018.

[6] Y. Guo, S. Han, C. Shen, Y. Li, X. Yin, and Y. Bai, "An adaptive SVR for high-frequency stock price forecasting," *IEEE Access*, vol. 6, pp. 11397–11404, 2018.

[7] B. Li, J. Zhang, Y. He, and Y. Wang, "Short-term load-forecasting method based on wavelet decomposition with second-order gray neural network model combined with ADF test," *IEEE Access*, vol. 5, pp. 16324–16331, 2017.

[8] P. H. Kuo and C. J. Huang, "A high precision artificial neural networks model for short-term energy load forecasting," *Energies*, vol. 11, no. 1, p. 213, Jan. 2018.
[9] M. Rafiei, T. Niknam, and M. H. Khooban, "Probabilistic forecasting of hourly electricity price by generalization of ELM for usage in improved wavelet neural network," *IEEE Trans. Ind. Informat.*, vol. 13, no. 1, pp. 71–79, Feb. 2017.
[10] J. Shi, W.-J. Lee, Y. Liu, Y. Yang, and P. Wang, "Forecasting power output of photovoltaic systems based on weather classification and support vector machines," *IEEE Trans. Ind. Appl.*, vol. 48, no. 3, pp. 1064–1069, May/Jun. 2012.
[11] B. Neupane, W. L. Woon, and Z. Aung, "Ensemble prediction model with expert selection for electricity price forecasting," *Energies*, vol. 10, no. 1, p. 77, Jan. 2017.
[12] M. Rafiei, T. Niknam, and M. H. Khooban, "A novel intelligent strategy for probabilistic electricity price forecasting: Wavelet neural network based modified dolphin optimization algorithm," *J. Intell. Fuzzy Syst.*, vol. 31, no. 1, pp. 301–312, Jan. 2016.
[13] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
[14] S. Anbazhagan and N. Kumarappan, "Day-ahead deregulated electricity market price forecasting using recurrent neural network," *IEEE Syst. J.*, vol. 7, no. 4, pp. 866–872, Dec. 2013.
[15] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, "Short-term residential load forecasting based on LSTM recurrent neural network," *IEEE Trans. Smart Grid*, vol. 10, no. 1, pp. 841–851, Jan. 2019.
[16] L. Han, Y. Peng, Y. Li, B. Yong, Q. Zhou, and L. Shu, "Enhanced deep networks for short-term and medium-term load forecasting," *IEEE Access*, vol. 7, pp. 4045–4055, 2019.
[17] U. Ugurlu, I. Oksuz, and O. Tas, "Electricity price forecasting using recurrent neural networks," *Energies*, vol. 11, no. 5, p. 1255, May 2018.
[18] U. Ugurlu, O. Tas, A. Kaya, and I. Oksuz, "The financial effect of the electricity price forecasts' inaccuracy on a hydro-based generation company," *Energies*, vol. 11, no. 8, p. 2093, Aug. 2018.
[19] H. Zheng, J. Yuan, and L. Chen, "Short-term load forecasting using EMD-LSTM neural networks with a xgboost algorithm for feature importance evaluation," *Energies*, vol. 10, no. 8, p. 1168, Aug. 2017.
[20] J. Lago, F. De Ridder, and B. De Schutter, "Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms," *Appl. Energy*, vol. 221, pp. 386–405, Jul. 2018.
[21] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *J. Mach. Learn. Res.*, vol. 13, pp. 281–305, Feb. 2012.
[22] P. M. R. Bento, J. A. N. Pombo, M. R. A. Calado, and S. J. P. S. Mariano, "A bat optimized neural network and wavelet transform approach for short-term price forecasting," *Appl. Energy.*, vol. 210, pp. 88–97, Jan. 2018.
[23] P. Mandal, A. Ul Haque, J. Meng, A. K. Srivastava, and R. Martinez, "A novel hybrid approach using wavelet, firefly algorithm, and fuzzy ARTMAP for day-ahead electricity price forecasting," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1041–1051, May 2013.
[24] M. Rafiei, T. Niknam, and M. H. Khooban, "Probabilistic electricity price forecasting by improved clonal selection algorithm and wavelet preprocessing," *Neural Comput. Appl.*, vol. 28, no. 12, pp. 3889–3901, Dec. 2017.
[25] X. Qiu, P. N. Suganthan, and G. A. J. Amaratunga, "Short-term electricity price forecasting with empirical mode decomposition based ensemble kernel machines," *Procedia Comput. Sci.*, vol. 108, pp. 1308–1317, Jan. 2017.
[26] *Pennsylvania-New Jersey-Maryland (PJM) Electricity Markets, PJM*. Accessed: Apr. 1998. [Online]. Available: http://www.pjm.com
[27] R. Martinez-Cantin, N. de Freitas, E. Brochu, J. Castellanos, and A. Doucet, "A Bayesian exploration-exploitation approach for optimal online sensing and planning with a visually guided mobile robot," *Auton. Robots*, vol. 27, no. 2, pp. 93–103, Aug. 2009.
[28] Norden E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, and C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. Roy. Soc. London A, Math., Phys. Eng. Sci.*, vol. 454, pp. 903–995, Mar. 1998.
[29] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: A noise-assisted data analysis method," *Adv Adapt Data Anal.*, vol. 1, no. 1, pp. 1–41, Jan. 2009.
[30] X. Zhang, K. K. Lai, and S.-Y. Wang, "A new approach for crude oil price analysis based on empirical mode decomposition," *Energy Econ.*, vol. 30, no. 3, pp. 905–918, May 2008.

SIYU ZHOU received the B.S. degree in electrical engineering from the Hefei University of Technology, Hefei, China, in 2017. He is currently pursuing the M.S. degree in electrical engineering with Chongqing University, Chongqing, China. His current research interests include data mining and intelligent algorithm.

LIN ZHOU was born in Sichuan, China, in 1961. He received the Ph.D. degree in electrical engineering from the Department of Electrical Engineering, Chongqing University, Chongqing, China, in 2004, where he is currently a Professor with the State Key Laboratory of Transmission and Distribution Equipment and Power System Safety and New Technology. He is engaged in teaching and scientific research of power systems. His current research interests include the key technology of distributed generation and power quality control.

MINGXUAN MAO was born in Luoyang, China, in 1988. He received the B.S. degree from the Chongqing University of Arts and Sciences, in 2012, and the Ph.D. degree from Chongqing University, Chongqing, China, in 2017. From June 2016 to June 2017, he was a Visiting Scholar with the University of Leeds, Leeds, U.K., where he was engaged in research on renewable energy systems and optimal control of power electronic converters. In 2018, he joined the State Key Laboratory of Transmission and Distribution Equipment and Power System Safety and New Technology, Chongqing University, where he is currently a Postdoctoral Research Fellow and a Lecturer. His current research interests include advanced control for renewable energy systems, power electronics, and artificial intelligence algorithms.

HENG-MING TAI (S'82–M'82–SM'94) received the Ph.D. degree in electrical engineering from Texas Tech University, Lubbock, TX, USA, in 1987. He is currently a Professor of electrical and computer engineering with The University of Tulsa, Tulsa, OK, USA. His research interests include signal processing, industrial electronics, and applications to power systems.

YIHAO WAN received the B.S. degree in electrical engineering from the Wuhan University of Technology, Wuhan, China, in 2017. He is currently pursuing the M.S. degree in electrical engineering with Chongqing University. His current research interests include power system optimization and heuristic algorithm.