Energy 158 (2018) 774–781

Contents lists available at ScienceDirect

Energy

journal homepage: www.elsevier.com/locate/energy

Short term electricity load forecasting using a hybrid model

Jinliang Zhang [a, *], Yi-Ming Wei [b], Dezhi Li [c], Zhongfu Tan [a], Jianhua Zhou [d]

[a] School of Economics and Management, North China Electric Power University, Beijing, 102206, China
[b] Center for Energy and Environmental Policy Research, Beijing Institute of Technology, Beijing, 100181, China
[c] China Electric Power Research Institute, Beijing, 100192, China
[d] Shanghai Electric Power Company, Shanghai, 200122, China

ARTICLE INFO

Article history:
Received 31 December 2017
Received in revised form
27 May 2018
Accepted 3 June 2018
Available online 13 June 2018

Keywords:
Electricity load forecasting
IEMD
ARIMA
WNN
FOA

ABSTRACT

Short term electricity load forecasting is one of the most important issue for all market participants. Short term electricity load is affected by natural and social factors, which makes load forecasting more difficult. To improve the forecasting accuracy, a new hybrid model based on improved empirical mode decomposition (IEMD), autoregressive integrated moving average (ARIMA) and wavelet neural network (WNN) optimized by fruit fly optimization algorithm (FOA) is proposed and compared with some other models. Simulation results illustrate that the proposed model performs well in electricity load forecasting than other comparison models.

1. Introduction

Accurate short term electricity load forecasting has a significant role in power system, which is useful for making optimal decisions to ensure the secure, reliable and economic operation of the power system [1]. Besides, power grid planning, investment and transaction are also based on accurate electricity load forecasting. However, due to many factors affecting electricity load, it turns out to be a challenging work. In recent years, researchers have proposed many models to forecast electricity load. In general, these models can be classified into three major categories: time series models, artificial intelligence models and hybrid models.

Time series models have been used for electricity load forecasting, such as linear regression [2], seasonal autoregressive [3], ARIMA [4], threshold autoregressive [5], kalman filtering [6], seasonal autoregressive integrate moving average (SARIMA) [7], etc. As mentioned above, electricity load is affected by many other factors, using only time series model does not produce a good result [8]. Thus, artificial intelligence models have been presented to forecast electricity load, such as expert systems [9], support vector machines [10], fuzzy logic [11], artificial neural networks [12], echo state networks [13], support vector regression [14]. Although artificial intelligence model can consider other influence factors, it still has its own shortcoming such as local minimum and over training. Thus, hybrid model integrating different models will be a good choice for improving the forecasting accuracy. The reason can be mainly attributed to two aspects: First, the model combined with suitable models can capture the different features of electricity load. Second, the model combined with other models can overcome its own defects. Therefore, the hybrid model may be superior to the single model. In general, the hybrid model for electricity load forecasting can be classified into two main categories.

For the first category model, electricity load was predicted separately by different models. Then, the weight of each model was calculated by a suitable method. The final forecasting value was the sum of each model's forecasting value multiplied by its weight [15–23]. In Refs. [24,25], the single model such as back propagation neural network (BPNN), genetic algorithm back propagation neural network (GABPNN), WNN, radical basis function neural network (RBFNN), general regression neural network (GRNN), ARIMA, support vector machine (SVM) was used separately to predict the electricity load. Then, the multi-objective flower pollination algorithm (MOFPA) is applied to optimize the weight of each model. The final prediction value is the product of each model's value and its weight. Lusis et al. [26] developed a load prediction model combined with regression trees, neural networks, and support vector

* Corresponding author.
E-mail address: zhangjinliang1213@163.com (J. Zhang).

https://doi.org/10.1016/j.energy.2018.06.012
0360-5442/

regression. Chahkoutahi et al. [27] have proposed a hybrid method combined with multi-layer perceptron (MLP), adaptive neuro-fuzzy inference system (ANFIS) and SARIMA for load forecasting. Tong et al. [28] proposed a deep learning based model combined with support vector regression model to forecast electricity load. Although the performance of the first category model is better than single model, there is a problem in calculating the weight of each model, which will affect the forecasting accuracy. Therefore, the second category model has been proposed by many researchers.

For the second category model, electricity load was decomposed into several components. Then each component was predicted by a suitable model. The final forecasting results were the sum of each component's forecasting results. In Ref. [29], the original electricity load was decomposed by wavelet transform into multiple terms. Then each term was predicted by a hybrid model combined with finite newton lagrangian support vector machine (NLSSVM) and ARIMA. Li et al. [30] used wavelet transform to decompose the original electricity load into several components. Then each component was predicted by extreme learning machine combined with partial least squares regression. In Ref. [31], electricity load was decomposed by wavelet transform into some detailed sub-series, then each sub-series was predicted by boundary network node (BNN) model. In Ref. [32], the wavelet transform was applied to filter the high frequency components, then the remaining component was predicted by grey model (GM) and particle swarm optimization (PSO).

Bessec et al. [33] established the hybrid model combined with wavelet transforms and neural network methods for short-run electricity load forecasting. Although wavelet transform can decompose the original electricity load into some more regular components, it lack the ability to extract the deep information as much as possible [34]. To increase the efficiency of decomposition, EMD was used by many researchers [35–37]. Fan et al. [38] used EMD to decompose the electricity load into some detail parts and an approximate part. Then the support vector regression (SVR) model was used for prediction. Qiu et al. [39] have proposed a hybrid model combined with EMD and deep belief network (DBN).

Although EMD can provide a better decomposition results than wavelet transform, it has an end effect, which becomes the main factor affecting the accuracy of EMD. To deal with this problem, IEMD is used to eliminate the end effect in this paper. Then the characteristics of electricity load can be extracted more accurately and effectively. By the IEMD, the original electricity load was decomposed into several components. Unlike other published literature, only two component are predicted by a suitable model in this paper. The reason is that the more component is, the greater accumulative error is. Besides, the cost of the hybrid model is increased accordingly. The first component is the residual term extracting from the original electricity load. As we know, the residual term is the low frequency component and follows the growth trend of electricity load. Thus, the residual term shows a linear trend, which can be seen in Fig. 2. Therefore, the ARIMA is used for the first component, which performs well in dealing with linear series. The second component is obtained by subtracting the first component from original electricity load, which avoids the loss of electricity load information. We know that the second component is a nonlinear series with calendar effects (week effect, month effect, holiday effect). Thus, WNN is used because it can capture the characteristics of calendar effects associated with nonlinear series [40]. To avoid adding hidden nodes and over-training problems, FOA is applied to optimize the WNN. Therefore, a hybrid model combined with IEMD, ARIMA and WNN optimized FOA is proposed in this paper, which is established based on capturing the different characteristics and overcoming the shortcoming of single model. Thus, the combination of different model could outperform the single one. The main contribution of this paper can be summarized as follows.

(1) The improved EMD is used for reducing the loss of information, which is rarely discussed in the literature related to electricity load decomposition. Thus the characteristics of electricity load can be extracted more accurately and effectively.
(2) Each component decomposing from the original electricity load is predicted by a suitable model according to its features. Thus, different characteristics associated with the original electricity load can be captured.
(3) Unlike other published literature, this paper selects the main influence factors for each component, respectively. Therefore, the characteristics of each component can be better captured by the suitable model.
(4) The hybrid model combined with IEMD, ARIMA, WNN and FOA is firstly proposed for short term electricity load forecasting and displays significant superiority based on the results of multiple robustness checks.

The rest of the paper is organized as follows: Section 2 introduces the proposed model. Section 3 describes the data and forecasting results. Section 4 concludes the paper.

2. Methodologies

2.1. IEMD

Compared with the traditional wavelet transform, EMD is a highly adaptable signal processing method, which is extremely applicable for non-stationary and nonlinear series. The procedures of the EMD can be summarized as following. More details of EMD can be found in Ref. [41].

(1) First, all the local extreme values of electricity load $L(t)$ must be identified. Then the values are used by a cubic spline line to construct the upper and lower envelopes. The average of the upper and lower envelopes is defined as m_1. The difference between $L(t)$ and m_1 is defined as h_1.

$$h_1 = L(t) - m_1 \tag{1}$$

If h_1 satisfies the definition of intrinsic mode functions (IMF), then it is selected as the first IMF. Otherwise, the above steps are repeated until $c_1(t)$ is an IMF.

(2) In the next step, the first IMF will be subtracted from the original electricity load.

$$L(t) - c_1(t) = r_1(t) \tag{2}$$

(3) Repeat the above process until the residual function $r_n(t)$ is a monotone function.
(4) By using the EMD, the original electricity load can be decomposed as follows:

$$L(t) = \sum_{i=1}^{n} c_i(t) + r_n(t) \tag{3}$$

Although EMD has many advantages, the end effect was produced in the decomposition process, which increased the deviation of the final result [42]. To obtain more accuracy component, mirror method is used, which can eliminate the end effect. The detailed progress of mirror method can be found in Ref. [43]. The procedures of mirror method used for EMD can be summarized as following:

(1) For the beginning of $L(t)$, add local minimum $Min(0)$ as well as local maximum $Max(1)$ by mirror symmetry.
(2) For the end of $L(t)$, add local maximum $Max(n+1)$ as well as local maximum $Min(n)$ by mirror symmetry.
(3) The newly obtained $Min(0)$ and $Max(n+1)$ are used for construction of the upper and lower envelops.
(4) Then the next step can follow the steps of the EMD.

2.2. ARIMA

ARIMA model is widely used in modeling of all kinds of linear series, which is the extension of ARMA (Box and Jenkins, 1970) and can be defined as follows:

$$\Delta^d y_t = \alpha_0 + \sum_{i=1}^{p} \beta_i \Delta^d y_{t-i} + \varepsilon_t + \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j} \tag{4}$$

where $\Delta^d y_t$ denotes the decomposition component of electricity load after d second difference. ε_t is the random error at time t, which is a white noise. $\alpha_0, \alpha_j, \beta_i$ are the parameters, p and q are the order.

Generally, the construction process of ARIMA includes three stages: model identification, parameter estimation and diagnosis checking.

(1) Model identification. As stationary is essential for ARIMA model, data should be often stationary. Differencing operation is used to remove the trend of data and stabilize the variance.
(2) Parameter estimation. Autocorrelation (ACF), partial autocorrelation (PACF), Akaike's Information Criterion (AICC) and Schwarz's Bayesian Information Criterion (BIC) are used for parameter estimation.
(3) Diagnostic checking. The model's accuracy and error stationary are checked. The best model is determined by some forecasting error criteria.

2.3. WNN optimized by FOA

2.3.1. WNN

Wavelet neural network is based on wavelet transform and neural network, which fully inherit the excellent time-frequency localization properties of wavelet transform and self-learning characteristic of neural networks to implement strong nonlinear approximation ability [44]. A detailed structure of WNN with n inputs and one output is shown in Fig. 1. There are four layers in the model.

Layer 1: This layer is input layer, which can deliver the decomposition component of electricity load to the second layer.

Layer 2: This layer is mother wavelet layer. The Gaussian function is select as the mother wavelet function for each node. Thus, the input at t moment can be determined as follows:

$$u_{ik} = x_i(t) + \beta_{ik}\varphi_{ik}(t-1), i \in [1, n] \pm \tag{5}$$

where u_{it} represents the output, β_{ik} is the self-feedback information, φ_{ik} is the Gaussian function, ik represents the kth wavelet of the ith input variable.

Layer 3: This layer is the product layer, which can be defined as follows.

$$\varphi_i = \prod_{k=1}^{n} \varphi_{ik}(z_{ik}) \tag{6}$$

where $z_{ik} = (u_{ik} - m_{ik})/d_{ik}$, m_{ik} represents the shift factor.

Layer 4: This layer is the output layer, which includes the first and third inputs, and can be expressed as follows.

$$y = \sum_{i=1}^{n} W_i \varphi_i \tag{7}$$

where W_i is the weight of output φ_i, y is the forecasting results of decomposition component of electricity load.

There are two important things in the design process of WNN. The first is the selection of the parameters of wavelet function. The second thing is the selection of connection weights. Thus, optimization algorithm is required to select the most appropriate parameters.

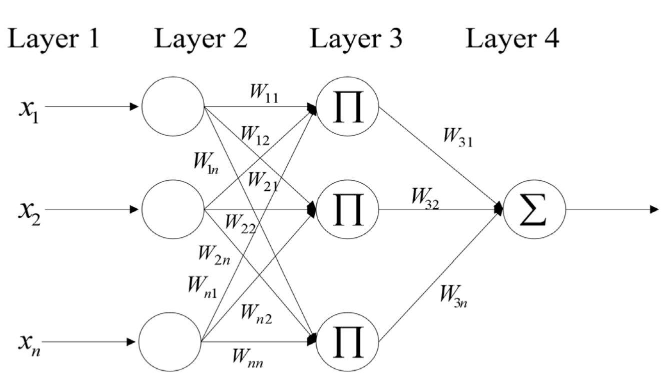

Fig. 1. The structure of WNN.

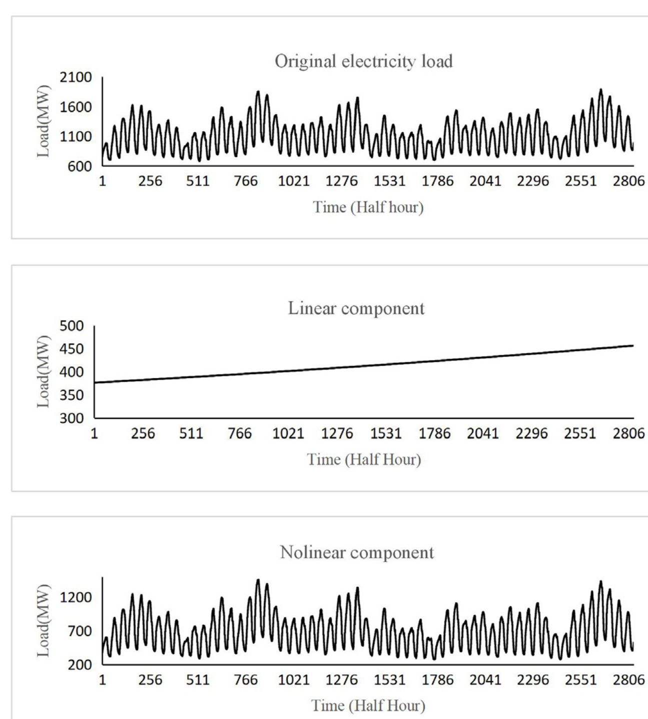

Fig. 2. Electricity load comprise of linear and nonlinear components in Australian market (January 1, 2010–February 28, 2010).

2.3.2. FOA

FOA is a new swarm intelligence algorithm proposed by Pan et al. [45], which performs better than ant colony optimization (ACO), PSO and other swarm intelligence algorithm [1]. The procedure of FOA are given bellows:

Step 1: Initialize the parameters. The parameters of the maximum number of generation, the size of population, and the center location are initialized.

Step 2: Initialize step size. Set the random direction and distance of each fruit fly to search food by the sense of smell.

Step 3: Calculate smell concentration. Estimate the distance between the location and the origin, then compute the smell concentration.

Step 4: Search best smell. Find out the fruit fly by the highest smell concentration, then the fruit fly swarm flies toward the location.

Step 5: Update swarm location. If the best food resource is better than the current swarm location, the latter is updated.

Step 6: Check termination criterion. When the termination condition is satisfied, computation is stopped. Otherwise, repeat step 2 to step 5.

2.3.3. WNN-FOA

As we known that the parameters of scale factor(a), position factor(b) and weights(W) affect the efficiency and performance of WNN. Thus the FOA is used to optimize these parameters. The procedure of WNN-FOA is presented as follows.

Step 1: Parameters Setting. Set the population size and the maximum number of iterations. The population size is the parameters to be optimized. Random initial population location, which can be defined as follows:

$$x_{o,m} = rand(LR), y_{o,m} = rand(LR) \tag{8}$$

where$m$represents the parameters to be optimized. $rand$is a random number generation function. $LR$is the location range.

Step 2: Evolution stating. Give the random direction and distance for an individual fruit fly using the sense of smell.

$$x_{i,m} = x_{o,m} + rand(FR), y_{i,m} = y_{o,m} + rand(FR) \tag{9}$$

where$i$denotes the$i$th fruit fly. $FR$is the random flight range.

Step 3: Preliminary calculation. Calculate the flight distance and the smell concentration judgement value$S_{i,m}$.

$$S_{i,m} = 1\Big/\sqrt{x_{i,m}^2 + y_{i,m}^2} \tag{10}$$

Step 4: Smell concentration calculation. Take smell concentration judgement value to smell concentration detection function, which is defined as follows. Then the fruit fly's individual location$smell_{i,m}$is obtained.

$$Function(S_{i,m}) = \frac{1}{N}\sum_{t=1}^{N}\left(y_t - \widehat{y_t}\right)^2 \tag{11}$$

where N is the number of prediction periods, $y_t$and $\widehat{y_t}$are the true value and predicted value at time t, respectively.

Step 5: Highest smell concentration selection. Find the fruit fly with the highest smell concentration among fruit fly swarm.

$$[bestsmell_m, bestindex_m] = \max(smell_{i,m}) \tag{12}$$

Step 6: Location updating. Keep the best smell concentration and coordinate. Then the fruit fly swarm uses vision to fly towards the position.

$$smellbest_m = bestsmell_m, x_{o,m} = x(bestindex_m), y_{o,m} = y(bestindex_m) \tag{13}$$

Step 7:Termination criterion checking. When the termination condition is satisfied, computation is stopped. The optimal parameter values of WNN are obtained. Otherwise, repeat step 2 to step 6.

2.4. The hybrid model

Hybrid models are commonly used for short term electricity load forecasting by more and more researchers, which can achieve more accurate and reliable forecasting results. Since electricity load contains both linear and nonlinear components, a hybrid model combined with linear and nonlinear model can better capture the characteristics of electricity load. Thus, the first thing is to separate the linear and nonlinear components from original electricity load. By using IEMD, the residual component representing the linear trend is extracted from electricity load, defined as$R(t)$. To avoid the loss of electricity load information, the nonlinear component is obtained by subtracting the residual component from the original electricity load, defined as$N(t)$. Therefore, the original electricity load$L(t)$can be expressed as the following equation with no information loss, which is shown in Fig. 2.

$$L(t) = N(t) + R(t) \tag{14}$$

There are many factors affecting electricity load forecasting. However, taking all the factors into account is impossible, and the influencing factors for each component is different. Unlike other published literature, this paper selects the main influencing factors for each component according to its different features. We think that electricity load forecasting is mainly affected by historical influence factors and future weather conditions. Because, other historical influencing factors are included in historical electricity load, which carries enough information for prediction. Therefore, electricity load forecasting is actual affected by its historical electricity load and future weather conditions.

As seen from Fig. 2, the residual component is a linear time series, which is mainly influenced by its historical electricity load. Thus, an ARIMA model is applied to predict the residual component. It also can be seen that, the nonlinear component has high volatility and calendar effects, which is mainly influenced by its historical electricity load and future weather conditions (such as temperature, humidity, wind speed, etc). In this paper, temperature is selected as the only factor of future weather conditions. The reason is that there is high correlation between electricity load and temperature. The other reason is that other factor's data cannot be obtained from public sources. Then, the WNN-FOA model with two exogenous variables is used to predict the nonlinear component, which does not need assumption of any functional relationship between electricity load and temperature. Thus, the temperature variable can be directly incorporated into the WNN-FOA model.

The next work is to select input features for each component. Feature selection is the process of removing irrelevant features and selecting the representative features that are necessary and sufficient for establishing a forecasting model. Appropriate feature selection improves the forecasting accuracy and speed by reducing over-fitting and solving the dimensionality reduction problem. The lagged values of load and temperatures are of great use for load forecasting. However, some lagged values may be redundant or even irrelevant for the forecasting model. Thus, Pearson's correlation coefficient (PC) is used to select the appropriate candidates in this paper, which is used to calculate the strength of relationship among two vectors. The correlation coefficient between two variables X_i and Y_j is defined as:

$$P(X_i, Y_j) = \frac{\mathrm{cov}(X_i, Y_j)}{\sqrt{\mathrm{var}(X_i)\mathrm{var}(Y_j)}} \tag{15}$$

where$P(X_i, Y_j)$is the correlation coefficient, $\mathrm{cov}(X_i, Y_j)$is the covariance. $\mathrm{var}(X_i), \mathrm{var}(Y_j)$is the standard deviation, respectively.

Hence, the hybrid model combined with IEMD, ARIMA, WNN and FOA is proposed for electricity load forecasting. The structure of the proposed model is shown in Fig. 3.

3. Numerical results

3.1. Accuracy assessment

There are many measurements to evaluate the forecasting accuracy. However, to make a fair comparison, MAPE, MAE, MPE and RMSE are used as the measurement index, which are defined as follows:

$$MAPE = \frac{1}{N}\sum_{t=1}^{N}\frac{\left|L_t - \widehat{L}_t\right|}{L_t} \times 100\% \tag{16}$$

$$MAE = \frac{1}{N}\sum_{t=1}^{N}\left|L_t - \widehat{L}_t\right| \tag{17}$$

$$MPE = \frac{1}{N}\sum_{t=1}^{N}\left(\left(\widehat{L}_t - L_t\right)\Big/L_t\right) \times 100 \tag{18}$$

$$RMSE = \frac{1}{N}\sqrt{\sum_{t=1}^{N}\left(L_t - \widehat{L}_t\right)^2} \tag{19}$$

where$L_t$and $\widehat{L}_t$are the actual and predicted electricity load, respectively. N is the prediction horizon.

3.2. Case studies

In this paper, we focus on short-term electricity load forecasting, which mainly refers to the prediction period for few hours, one day, one week and one month. The purpose of this paper is to propose an accurate and robust model that can be used for different types of hours, different types of days and different types of markets, which will be very attractive for all market participants. The electricity load data (www.aemo.com.au) and temperature data (www.bom.gov.au) from Australian, and the electricity load data (www.nyiso.com) and temperature data (www.weather.gov) from New York City are used to evaluate the performance of the proposed model.

3.2.1. Comparison results for Australian market

Electricity load (half-hour data) is collected from November 1, 2011 to December 31, 2012 and is divided into training set and testing set. For a fair comparison, each month of 2012 is used as testing set, while the two month before testing month is used as training set. The prediction results of proposed model are compared with results obtained in Ref. [22], which are given in Table 1.

As seen in Table 1, all evaluation values of the proposed model are lower than comparison models in testing months. Due to the high volatility, the forecasting accuracy of electricity load in Summer and Fall are lower than in Spring and Autumn. The MAPE values of the proposed model are fluctuated from 0.76% to 0.92%. However, when the fluctuation of the electricity load is more stable,

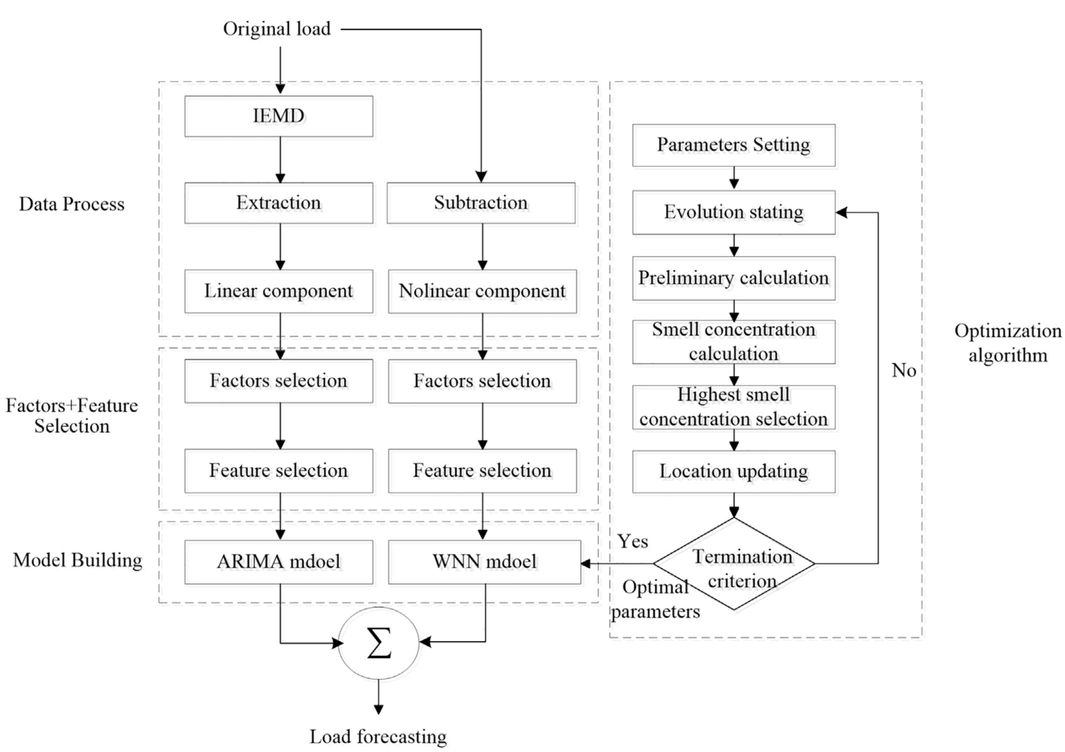

Fig. 3. The structure of the proposed model.

then the prediction accuracy will be higher. In order to further illustrate the performance of the proposed model, other comparison results are presented in Table 2.

It can be seen from Table 2, the average values of MAPE, MAE and RMSE obtained by the proposed model are also lower than comparative models. The average value of MAPE for ENN (2.34%) is improved by 64.95% with the proposed model (0.82%). The improvement of MAPE for SVM and ELM is about 26.78% and 9.89%, respectively. We also see that the standard deviation values of MAPE, MPE and RMSE obtained by the proposed model are lower than other models, which indicates that the stability of the hybrid model is the best one. Both Tables 1 and 2 demonstrate that the proposed model can provide a more accurate and stable results in comparing with other models.

To verify the effectiveness of the proposed model, the last week of each month in 2013 are used as the testing set. 30 days before the testing week are used as the training set. The proposed model is compared with some other models, such as IEMD combined with ARIMA (IEMDA), IEMD combined WNN-FOA (IEMDW), EMD combined with ARIMA and WNN-FOA (EMDAW). Table 3 presents the MAPE for 12 weeks of the Australian market in 2013. As an example, the forecasting results of the last week of January 2013 are also given in Fig. 4.

From Table 3, it is obviously that the prediction accuracy of the proposed model is higher than comparison models. The prediction error of IEMDA is in the range of 1.53%–1.75%. After combining with WNN-PSO, the average MAPE of IEMDA reducing from 1.65% to 0.68%. This results indicate that WNN-PSO can capture the characteristics of nonlinear in electricity load. The prediction error of IEMDW is in the range of 1.43%–1.69%. After combining with ARIMA, the average MAPE of IEMDW reducing from 1.58% to 0.68%, which demonstrates that ARIMA can fit the linear trend associated with electricity load. Besides, the proposed model was also compared with EMDAW, with the average MAPE reducing from 0.79% to 0.68%. The results confirm that electricity load can be extracted more accurately and effectively by IEMD. Thus, the forecasting results of the proposed model are more reasonable and accurate. Fig. 4 illustrates that the proposed model can follow the trend of the actual load, even when the electricity load shows a high volatility. All these examinations indicate that the proposed model can capture the complex characteristics of electricity load.

3.2.2. Comparison results for New York market

Since the accuracy of electricity load forecasting is also affected by the types of market, electricity load acquired from New York is used to evaluate the performance of the proposed model. The results of the proposed model are also compared with the results obtained form [16,31]. For a fair comparison, the first day of July for the year 2004 are considered as testing set, while 30 days previous to the testing day are used to build the prediction model. The comparison results for the testing day are shown in Table 4 and Fig. 5.

In Table 4, the MAPE of WTNNEA, WGMIPSO is 2.06% and 0.68%, respectively. However, the MAPE of the proposed model is only 0.60%, which is obviously less than WTNNEA and WGMIPSO. Meanwhile, the MAE and MPE obtained by the proposed model are also reduced than the comparison models. This finding proves the superiority of the proposed model. As can be seen from Fig. 5, the curve between the actual value and the forecast value is very close, which also demonstrates the superiority of the proposed model.

To illustrate the performance of the proposed model for different types of day (such as weekday, weekend, special day), each day of July is used as the testing set. For a fair comparison, 30 days previous to the testing day are used to build the prediction

Table 2
The average value and standard deviation of different evaluation criterion in 2012.

Model	Criterion	Average value	Standard deviation
ENN [22]	MAPE(%)	2.34	0.04
	MAE(MW)	204.19	5.36
	RMSE	7.05	0.23
SVM [22]	MAPE(%)	1.12	0.06
	MAE(MW)	99.26	6.94
	RMSE	3.59	0.25
ELM [22]	MAPE(%)	0.91	0.04
	MAE(MW)	80.23	5.78
	RMSE	2.88	0.21
Proposed model	MAPE(%)	0.82	0.04
	MAE(MW)	69.03	3.70
	RMSE	2.15	0.15

Table 3
MAPE (%) for 12 weeks of Australian market in 2013.

	IEMDA	IEMDW	EMDAW	Proposed model
January	1.67	1.56	0.86	0.71
February	1.72	1.62	0.85	0.70
March	1.53	1.43	0.74	0.66
April	1.56	1.46	0.76	0.68
May	1.60	1.48	0.73	0.65
June	1.73	1.71	0.91	0.75
July	1.75	1.68	0.87	0.71
August	1.72	1.63	0.82	0.68
September	1.59	1.55	0.71	0.65
October	1.62	1.59	0.75	0.67
November	1.66	1.62	0.75	0.67
December	1.73	1.69	0.82	0.71
Average	1.65	1.58	0.79	0.68

Table 1
Comparison results of different evaluation criterion for each month in 2012.

Model	Criterion	Jan.	Feb.	Mar.	Apr.	May.	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
ENN [22]	MAPE(%)	1.79	1.87	2.14	2.37	2.50	3.03	2.83	2.85	2.33	2.11	1.98	2.32
	MAE(MW)	160.1	164.0	172.2	191.5	226.2	297.2	274.1	267.0	197.5	166.5	157.1	176.9
	RMSE	5.40	5.48	5.70	6.38	8.09	10.58	9.75	9.63	6.90	5.52	5.26	5.87
SVM [22]	MAPE(%)	0.94	0.98	1.01	1.20	1.26	1.28	1.26	1.37	1.19	0.99	1.00	0.97
	MAE(MW)	84.19	87.02	84.98	100.1	114.4	125.6	122.8	128.4	102.4	80.16	82.00	79.04
	RMSE	2.93	3.07	2.99	3.77	4.24	4.63	4.50	4.73	3.80	2.75	2.82	2.83
ELM [22]	MAPE(%)	0.84	0.82	0.88	1.04	0.94	0.98	0.96	0.95	0.95	0.84	0.82	0.92
	MAE(MW)	75.59	73.43	73.77	87.05	84.79	95.42	92.03	87.95	81.60	68.50	68.00	74.68
	RMSE	2.63	2.62	2.64	3.41	3.13	3.34	3.21	3.21	3.03	2.63	2.36	2.66
Proposed model	MAPE(%)	0.78	0.76	0.80	0.92	0.86	0.87	0.85	0.86	0.85	0.79	0.76	0.80
	MAE(MW)	67.51	65.76	65.82	72.59	71.21	75.43	72.35	73.08	70.05	64.34	64.11	66.17
	RMSE	2.03	2.01	2.05	2.20	2.17	2.19	2.15	2.61	2.15	2.04	2.13	2.08

Note: ENN is the abbreviation of elman neural network, ELM is the abbreviation of extreme learning machine.

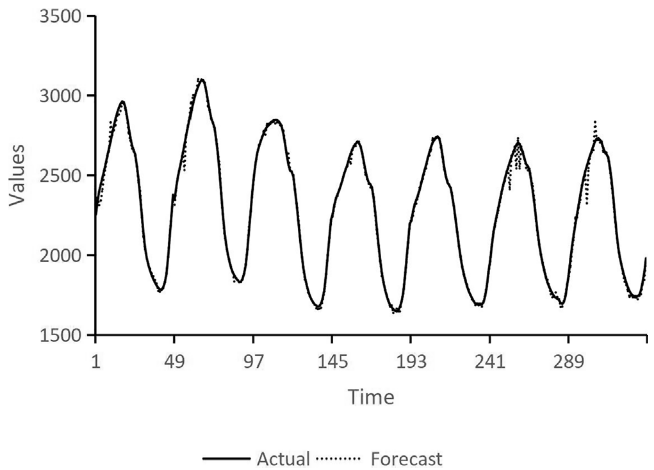

Fig. 4. Forecasting results of the proposed model for the last week of January in 2013.

Table 4
Comparison results for the first day of July 2004.

	MAPE(%)	MAE(MW)	MPE(%)
WTNNEA [16]	2.06	139.95	−0.62
WGMIPSO [30]	0.68	46.62	−0.06
Proposed model	0.60	38.61	−0.04

Note: WTNNEA is the hybrid model combined with wavelet transform, neural network and evolutionary algorithm, WGMIPSO is wavelet transform combined with grey model improved by particle swarm optimization.

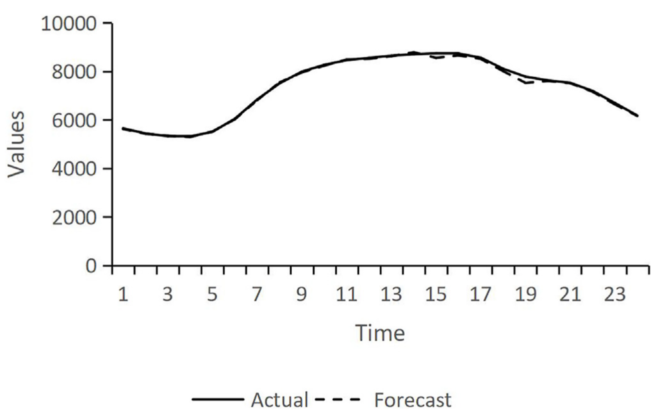

Fig. 5. Forecasting results of the proposed model for the first day of July in 2004.

model. Table 5 presents the MAPE for all days of July 2004.

As seen from Table 5, the MAPE, MAE and MPE of the proposed model are approximately 29.70%, 8.10% and 98.76% less than WTNNEA, respectively. The different evaluation values of the proposed model are also less than WGMIPSO. This results demonstrate that the proposed model outperforms the comparison models. Fig. 6 shows the MAPE of the proposed model for all testing days in July. The MAPE are fluctuated from 0.60% to 1.73%. July 1 has the lowest value of MAPE, and July 23 has the highest value. Generally speaking, the MAPE of weekday is less than weekend and holidays. July 5 is a legal holiday in New York. Although the MAPE of some days are bigger than other days, the average of MAPE is still lower (1.13%).

Table 5
Comparison results for all day of July 2004.

	MAPE(%)	MAE(MW)	MPE(%)
WTNNEA [16]	2.02	123.62	−0.81
WGMIPSO [30]	1.82	122.47	−0.01
Proposed model	1.42	113.61	−0.01

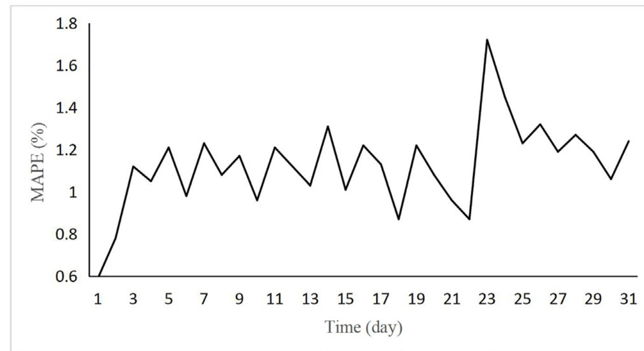

Fig. 6. The MAPE for all days in July.

Table 6
Forecasting results for different hours of July 9, 2004.

	Actual load	Forecast load	MAPE(%)
Hour (00:00:00)	9473	9415	0.61
Hour (01:00:00)	8859	8812	0.53
Hour (02:00:00)	8372	8331	0.49
Hour (08:00:00)	10320	10245	0.73
Hour (09:00:00)	11033	10937	0.87
Hour (10:00:00)	11468	11386	0.71

To further illustrate the accuracy and robustness of the proposed model, electricity load acquired from peak hour and off-peak hour of July 9 in New York City are used for testing. The results is given in Table 6. As shown in Table 6, the MAPE of the peak hour is bigger than off-peak hour. The reason can be attributed to the high volatility caused by weather forecast, consumption customers, etc [31]. The average MAPE for the peak hours is 0.66%, which is a reasonable result.

4. Conclusions

This paper proposes a hybrid model for short term electricity load forecasting. First, the trend component is extracted from the original electricity load. Second, the nonlinear component is obtained by subtracting the trend component from original electricity load. Third, the two components are predicted separately by the suitable model. Last, each component's forecasting results are added up to obtain the final forecasting results. Electricity load data from Australian and New York electricity markets are used to demonstrate the performance of the proposed model. All case studies indicate that the proposed model can improve the forecasting accuracy of electricity load than the comparison models.

Three facts emerge clearly from the results: (1) The linear component of electricity load can be extracted more accurately and effectively by the IEMD. (2) The ARIMA model can well fit the linear component of the original electricity load. (3)The WNN optimized by FOA has a strong ability to fit the nonlinear component of the original electricity load. By using each model's advantage, the hybrid model can capture the different characteristics associated with electricity load. Therefore, the proposed model can provide a robust, stable and accurate prediction results.

Advanced models can be used to select suitable input variables for electricity load forecasting in the future. Besides, some other future influencing factors such as psychological expect can be added in the hybrid model as future research.

Acknowledgement

This work was supported by the China National Science Foundation (No.71774054), the Fundamental Research Funds for the Central Universities (2017MS081) and Science and Technology Project of State Grid Corporation of China (Grant no. YDB17201600102). We are grateful to the China Scholarship Council (the International Clean Energy Talent Programme, 2017).

References

[1] Hu R, Wen SP, Zeng ZG, Huang TW. A short-term power load forecasting model based on the generalized regression neural network with decreasing step fruit fly optimization algorithm. Neurocomputing 2017;221:24–31.
[2] Papalexopoulos AD, Hesterberg TC. A regression-based approach to short term system load forecasting. IEEE Trans Power Syst 1990;4:1535–47.
[3] Mbamalu G, El-hawary M. Load forecasting via suboptimal seasonal autoregressive models and iteratively reweighted least-squares estimation. IEEE Trans Power Syst 1993;1:343–8.
[4] Chen JF, Wang WM, Huang CM. Analysis of an adaptive time-series autoregressive moving-average (ARMA) model for short-term load forecasting. Elec Power Syst Res 1995;3:187–96.
[5] Huang S. Short-term load forecasting using threshold autoregressive models. IEEE Proc-Gener Transm Distrib 1997;5:477–81.
[6] Zhang M, Bao H, Yan L, Cao JP, Jian-Guang DU. Research on processing of short-term historical data of daily load based on kalman filter. Power Syst Technol 2003;10:1–9.
[7] Tarsitano A, Amerise IL. Short-term load forecasting using a two-stage sarimax model. Energy 2017;133:108–14.
[8] Che J, Wang J, Wang G. An adaptive fuzzy combination model based on self-organizing map and support vector regression for load forecasting. Energy 2012;37:657–64.
[9] Rahman S, Bhatnagar R. An expert system based algorithm for short term load forecast. IEEE Trans Power Syst 1988;3:392–9.
[10] Pai PF, Hong WC. Forecasting regional electricity load based on recurrent support vector machines with genetic algorithms. Elec Power Syst Res 2005;74:417–25.
[11] Chenthur Pandian S, Duraiswamy K, Christober Asir Rajan C, Kanagaraj N. Fuzzy approach for short term load forecasting. Elec Power Syst Res 2006;76:541–8.
[12] Kouhi S, Keynia F. A new cascade NN based method to short-term load forecast in deregulated electricity market. Energy Convers Manag 2013;71:76–83.
[13] Deihimi A, Showkati H. Application of echo state networks in short-term electric load forecasting. Energy 2012;1:327–40.
[14] Yang YL, Che JX, Li YY, Zhao YJ, Zhu SL. An incremental electric load forecasting model based on support vector regression. Energy 2017;113:796–808.
[15] Kyung SB, Young SB, Dug HH, Gilsoo J. Short term load forecasting for the holidays using fuzzy linear regression method. IEEE Trans Power Syst 2005;1:96–101.
[16] Amjady N, Keynia F. Short-term load forecasting of power systems by combination of wavelet transform and nero-evolutionary algorithm. Energy 2009;34(1):46–57.
[17] Wang J, Zhu S, Zhang W, Lu H. Combined modeling for electric load forecasting with adaptive particle swarm optimization. Energy 2010;35:1671–8.
[18] Che JX, Wang JZ. Short-term load forecasting using a kernel-based support vector regression combination model. Appl Energy 2014;132:602–9.
[19] Xiao LY, Wang JZ, Hou R, Wu J. A combined model based on data pre-analysis and weight coefficients optimization for electrical load forecasting. Energy 2015;82:524–49.
[20] Coelho VN, Coelho LM, Coelho BN, et al. A self-adaptive evolutionary fuzzy model for load forecasting problems on smart grid environment. Appl Energy 2016;169:567–84.
[21] Xiao LY, Shao W, Liang TL, Wang C. A combined model based on multiple seasonal patterns and modified firefly algorithm for electrical load forecasting. Appl Energy 2016;167:135–53.
[22] Jiang P, Liu F, Song YL. A hybrid forecasting model based on date-framework strategy and improved feature selection technology for short-term load forecasting. Energy 2017;119:694–709.
[23] Chahkoutahi F, Khashei M. A seasonal direct optimal hybrid model of computational intelligence and soft computing techniques for electricity load forecasting. Energy 2017;140:988–1004.
[24] Xiao LY, Shao W, Yu MX, Ma J, Jin CJ. Research and application of a combined model based on multi-objective optimization for electrical load forecasting. Energy 2017;119:1057–74.
[25] Xiao LY, Shao W, Wang C, Zhang KQ, Lu HY. Research and application of a hybrid model based on multi-objective optimization for electrical load forecasting. Appl Energy 2017;180:213–33.
[26] Lusis P, Khalilpour KR, Andrew L. Liebman A.Short-term residential load forecasting: impact of calendar effects and forecast granularity. Appl Energy 2017;205:654–69.
[27] Chahkoutahi F, Khashei M. A seasonal direct optimal hybrid model of computational intelligence and soft computing techniques for electricity load forecasting. Energy 2017;140:988–1004.
[28] Tong C, Li J, Lang C, et al. An efficient deep model for day-ahead electricity load forecasting with stacked denoising auto-encoders. J Parallel Distr Comput 2018;6:267–73.
[29] Ghasemi A, Shayeghi H, Moradzadeh M, Nooshyar M. A novel hybrid algorithm for electricity price and load forecasting in smart grids with demand-side management. Appl Energy 2016;177:40–59.
[30] Li S, Goel L, Wang P. An ensemble approach for short-term load forecasting by extreme learning machine. Appl Energy 2016;170:22–9.
[31] Ghofrani M, Ghayekhloo M, Arabali A, Ghayekhoo A. A hybrid short term load forecasting with a new input selection framework. Energy 2017;81:777–86.
[32] Bahrami S, Hooshmand RA, Parastegari M. Short term electric load forecasting by wavelet transform and grey model improved by PSO (particle swarm optimization) algorithm. Energy 2014;72:434–42.
[33] Besseca M, Fouquaub JL. Short-run electricity load forecasting with combinations of stationary wavelet transforms. Eur J Oper Res 2018;1:149–64.
[34] Alickovic E, Kevric J, Subasi A. Performance evaluation of empirical mode decomposition, discrete wavelet transform, and waveletpacked decomposition for automated epileptic seizure detection and prediction. Biomed Sigmal Proces 2018;39:94–102.
[35] Fan GF, Qing S, Wang H, Hong WC, Li HJ. Support vector regression model based on empirical mode decomposition and auto regression for electric load forecasting. Energies 2016;6:1887–901.
[36] Xiong T, Bao Y, Hu Z. Interval forecasting of electricity demand: a novel bivariate EMD-based support vector regression modeling framework. Int J Electr Power Energy Syst 2014;63:353–62.
[37] Liu N, Tang QF, Zhang JH, Fan W, Liu J. A hybrid forecasting model with parameter optimization for short-term load forecasting of micro-grids. Appl Energy 2014;129:336–45.
[38] Fan GF, Peng LL, Hong WC, Sun F. Electric load forecasting by the SVR model with differential empirical mode decomposition and auto regression. Neurocomputing 2016;173:958–70.
[39] Qiu XH, Ren Y, Suganthan PN, Amaratunga GAJ. Empirical Mode Decomposition based ensemble deep learning for load demand time seriesforecasting. Appl Soft Comput 2017;54:246–55.
[40] Xu ZG, Zhang SY. High-frequency financial data calendar effects' wavelet neural network analysis. Math. Practice Theory 2007;15:1–6.
[41] Huang NE, Shen Z, Long SR, et al. The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings: mathematical. Phys Eng Sci 1998;454:903–95.
[42] Ning J, Chu CQ, Gao PX. Study on extending method of EMD. Comput Eng and Appl 2011;3:125–8.
[43] Rilling G, Flandrin P, Gonçalvés P. On empirical mode decomposition and its algorithms. IEEE-EURASIP Workshop on Nonlinear Signal and Image 2003;12:1–5.
[44] Abiyev H. Fuzzy wavelet neural network for control of dynamic plants. Int J Comput Intell 2004;1:441–4.
[45] Pan WC. A new fruit fly optimization algorithm: taking the financial distress model as an example. Knowl.-Based Syst 2012;26:69–74.