Energy 166 (2019) 598–609

Contents lists available at ScienceDirect

Energy

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

# A new hybrid model to predict the electrical load in five states of Australia

Jinran Wu [a], Zhesen Cui [b], Yanyan Chen [c], Demeng Kong [d], You-Gan Wang [a, *]

[a] School of Mathematical Sciences, Queensland University of Technology, Brisbane, Queensland, 4001, Australia
[b] Department of Computer Science, Changzhi University, Changzhi, Shanxi, 046000, China
[c] Taobao (China) Software Co., Ltd, Hangzhou, Zhejiang, 311121, China
[d] School of Mathematics and Statistics, Lanzhou University, Lanzhou, Gansu, 730000, China

ARTICLE INFO

*Article history:*
Received 9 April 2018
Received in revised form
3 October 2018
Accepted 13 October 2018
Available online 15 October 2018

*Keywords:*
Electrical load
Forecast
Ensemble empirical mode decomposition
Extreme learning machine
Grasshopper optimization algorithm

ABSTRACT

Short-term electrical load forecasting is an important part in the management of electrical power because electrical load is an extreme, complex non-linear system. To obtain parameter values that provide better performances with high precision, this paper proposes a new hybrid electrical load forecasting model, which combines ensemble empirical mode decomposition, extreme learning machine, and grasshopper optimization algorithm for short-term load forecasting. The most important difference that distinguishes this electrical load forecasting model from other models is that grasshopper optimization can search suitable parameters (weight values and threshold values) of extreme learning machine, while traditional parameters are selected randomly. It is applied in Australia electrical load prediction to show its superiority and applicability. The simulation studies are carried out using a data set collected from five main states (New South Wales, Queensland, Tasmania, South Australia and Victoria) in Australia from February 1 to February 27, 2018. Compared with all considered basic models, the proposed hybrid model has the best performance in predicting electrical load.


## 1. Introduction

Accurate and efficient electrical load forecasting has become increasingly important for power systems and optimal decisions [1]. In general, according to the forecasting time scale, there are three main electrical load demand predictions. The first category is the annual electrical load demand prediction, which is the premise of formulating reasonable macro power planning [2]. As well as annual demand prediction, the daily electrical load demand prediction is in a comparative small interval, which can daily offer the administration some suggestions of allocating the power production [3]. The final category, half-hourly demand prediction, is the most efficient power management tool, which can help power administrative departments to reduce power generation costs and establish power construction plans [4]. However, the electrical load is regarded as an open, non-linear, dynamic, and complex system [5]. For this reason, electrical load forecasting has been a hot topic in power research, and many new approaches have been introduced to improve prediction.

### 1.1. Literature review

In recent years, data-driven methods have become popular for electrical load forecasting with the development of data mining techniques. More specifically, various data mining techniques are used to predict the electrical load time series. Boroojeni et al. introduced the autoregressive and moving-average components to the short-term electrical power demand forecasting [6]. As for grey models, Zhao and Guo proposed an annual power load predictive model based on grey model [7]. Mamlook et al. used a fuzzy inference model to improve the predictive accuracy of short term load [8]. Then, the linear regression model was introduced to promoting the predictive accuracy by Kamyab and Bahrami [9]. Besides, artificial intelligence techniques also are used to forecast the electrical load. For example, the support vector regression (SVR) was applied in short term electric load forecasting by Chen and his team [10]. Another case is Hu et al. proposed a model based on artificial neural network for electrical load prediction [11].

However, a single data mining technique can only capture a few characteristics of the electrical load time series. The hybrid model

* Corresponding author.
*E-mail address:* you-gan.wang@qut.edu.au (Y.-G. Wang).

https://doi.org/10.1016/j.energy.2018.10.076
0360-5442/

makes use of multiple data mining techniques, which can improve the performance of short-term forecasting. The two classical artificial intelligence approaches, SVR and ANN, are often used as the main algorithms to capture non-linear relationships between input variables and targets. On one hand, many researchers combined the SVR methods with other data mining techniques. Zhang, Wang and Zhang investigated hybrid approaches for forecasting based on SVR and cuckoo search algorithm [12]. Yang et al. proposed a hybrid approach based on SVR and a nested particle swarm optimization; the results showed the model can improve annual power load forecasting accuracy [13]. Zhang proposed an approach for short-term load forecasting using a cuckoo search algorithm based on SVR, and the empirical results demonstrated that the model has a better performance of prediction [14]. Hong proposed a hybrid model based on SVR—with nonlinear mapping capabilities of forecasting—and a chaotic particle swarm optimization algorithm to choose the suitable parameters [15]. Researchers have noticed that the ANN can play an important role in electrical load forecasting while the application of SVR is also widely accepted [16]. To improve the accuracy of electrical load forecasting, Xiao et al. employed the modify firefly algorithm to optimize the weights and bias of ANN to get a higher accuracy for power load forecasting [17]. Mosbah intended to determine the optimal parameters of ANN by stochastic fractal search to improve the performance of forecasting [18]. Xiao et al. proposed an ANN optimized by multi-objective optimization to enhance the efficiency and reliability of power system's operation [19].

However, the two classical data mining techniques usually require a substantial amount of time to train the forecasting model in engineer application. To deal with this problem, Huang developed an algorithm called extreme learning machine (ELM), which is based on the traditional single-hidden layer feedforward neural network [20]. Specifically, ELM can reduce the training time of neural networks and achieve a greater accuracy of forecasting. ELM is faster than traditional methods, including support vector machine and artificial neural networks [21]. ELM has thus become a popular forecasting technique for short-term load forecasting due to its optimal performance [22]. For high-quality short-term load forecasting of the New England Electricity Market, Li et al. developed the ELM, and algorithm's training efficiency and accuracy were superior [23]. Ertugrul proposed an approach to accurately predict power load based on ELM, and the results showed that ELM can be utilized in modeling dynamic systems effectively [24]. Zeng built a hybrid learning approach combining ELM with improved PSO for short-term load forecasting, which resulted in a higher accuracy than the traditional ANN [25].

Although ELM is obtained through the ordinary least squares instead of training the parameters (weights and bias), the ELM's performance may be affected by the parameters. Therefore, selecting suitable parameters is crucial for ELM to work. Fortunately, the development of optimization algorithms and meta-heuristic optimization algorithms have made it possible to select the optimal parameter values. These algorithms can bypass local optima and are, therefore, easy to implement [26]. The idea of meta-heuristic algorithms comes from the behavior of animals or physical phenomena [27]. In addition, the algorithms can be grouped into three main categories (Fig. 1). The process of determining optimal parameters can be divided into two phases: exploration and exploitation [28]. More specifically, in the exploration phase, search agents are targeted to investigate the search space globally. In the exploitation phase, search agents are employed to search the best results by following the exploration phase. In this paper, a new meta-heuristic optimization algorithm grasshopper optimization algorithm (GOA) is utilized to search for ELM parameters for short-term load forecasting because the latest optimization algorithm has the most outstanding performance among the former classical optimizers [29].

The electrical load series is complicated, and it is difficult to establish the system between inputs and output. To solve this problem, some decomposition methods are introduced to improve the performance of short-term load forecasting. There are three main methods to decompose the original sequence to low frequent sub-sequences. The first decomposition method is wavelet decomposition (WA), which is a classical method; it can decompose the original signals by wavelet functions [30]. The second technique, empirical mode decomposition (EMD), can make the instantaneous frequency meaningful and eliminate the need for spurious harmonics to represent nonlinear and non-stationary signals through the Hilbert transform [31]. The final method is ensemble empirical mode decomposition (EEMD), which improve the robustness of EMD by adding extra white noise in the original signals [32]. The main idea of the methods is that the sub-sequences are established by data mining techniques individually, and all the results are integrated together as the final prediction. This serves as the final system for power load forecasting. The decomposition and ensemble methods have been popularly applied in electrical problems [33]. Ghofrani et al. proposed a model combining WA and ANN to forecast the short-term load in New England [34]. Zhang et al. developed a hybrid model based on EMD, autoregressive integrated moving average and ANN for electric load forecasting, and the results showed it can lower errors in forecasting [35]. Li and his team combined EEMD with ANN to establish a hybrid model to predict the electric load; the model performed better than other models [36].

### 1.2. Contribution

The aim of this paper is to develop hybrid models based on ELM, which is known to be robust, highly accurate, and computationally efficient. The newly developed algorithm also makes use of the GOA and EEMD. The GOA is known to perform well in searching the global optimum parameters, and the EEMD is a robust decomposition method which has led to superior performance in our new hybrid algorithms. This paper proposes a novel hybrid model, EEMD-ELM-GOA, for 1-step ahead electrical load forecasting in five states of Australia (New South Wales, Queensland, Tasmania, South Australia and Victoria) to verify the accuracy of the new hybrid model.

### 1.3. Organization of the paper

This paper mainly consists of three parts: (a) introduction of EEMD, ELM, GOA; (b) description of the experiment's data set, the evaluation criteria for model, construction of the hybrid model, performance of all models evaluated by simulations; and (c) summary and conclusion.

## 2. The proposed approach

In this section, the basic theory of the proposed approach is described, and the three machine learning techniques (EEMD, ELM and GOA) are elaborated on.

### 2.1. Extreme learning machine

The extreme learning machine was developed by Huang to solve two important bottleneck problems: 1) the slow gradient-based learning algorithms used to train neural network; 2) all parameters of the networks being tuned iteratively by using such learning algorithms [20]. Unlike traditional neural networks, ELM is one of

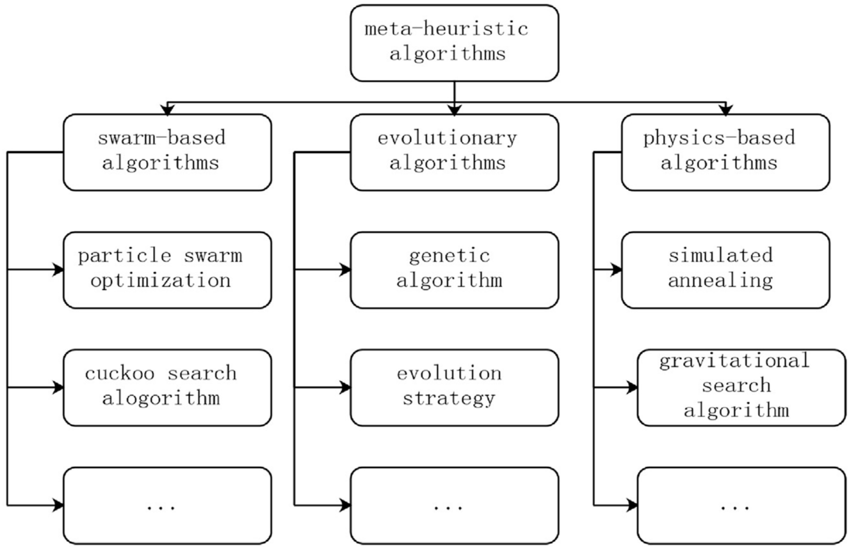


**Fig. 1.** Classification of meta-heuristic algorithms.

the new types of single-hidden layer feedforward neural networks (SLFNs), which choose the connection weights and bias of SLFNs randomly. The structure of ELM is shown in Fig. 2.

Given a data set $(\boldsymbol{x_q}, \boldsymbol{t_q})$, where $\boldsymbol{x_q}=[x_{q1}, \ldots, x_{qn}]\in\boldsymbol{R}^n$, and $\boldsymbol{t_q}=[t_{q1}, \ldots, t_{qm}]\in\boldsymbol{R}^{\boldsymbol{m}}$, extreme learning machine with $n$ input nodes, $L$ hidden nodes, $m$ output nodes, and activation function $g(\cdot)$ is modeled as,

$$T' = H \bullet \beta, \tag{1}$$

where $H = \begin{bmatrix} g(w_1 \cdot x_1 + b_1) & \cdots & g(w_l \cdot x_1 + b_L) \\ \vdots & \vdots & \vdots \\ g(w_1 \cdot x_Q + b_1) & \cdots & g(w_1 \cdot x_Q + b_L) \end{bmatrix}_{Q\times L}$, $\beta = \begin{bmatrix} \beta_{11} & \cdots & \beta_{1m} \\ \vdots & \vdots & \vdots \\ \beta_{L1} & \cdots & \beta_{Lm} \end{bmatrix}_{L\times m}$ and $T=[t_1, \ldots, t_Q]_{m\times Q}$ ($Q$ is the number of samples), and in which $w_i=[w_{i1}, w_{i2}, \ldots, w_{in}]$ is the weight vector connecting the i-th hidden node and the input nodes, and $\beta_{ik}$ is the weight vector connecting the i-th hidden node and the k-th output node, $b_j$ is the bias of the j-th hidden node. Here, a reference for Theorem 1 and 2 needs to be introduced to dealing with formula (2) [20].

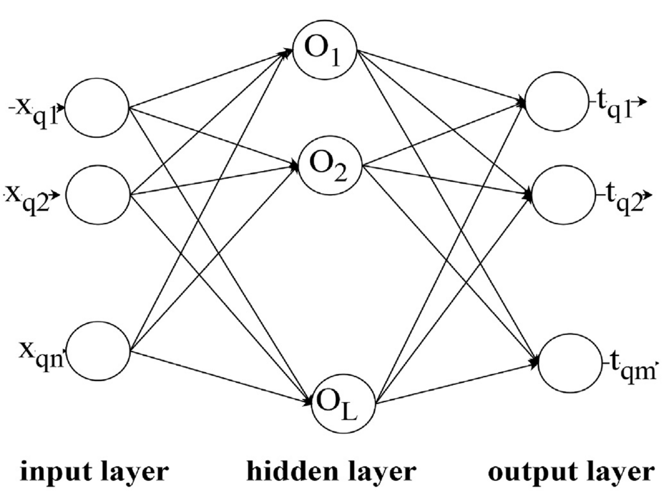


**Fig. 2.** The structure of ELM.

**Theorem 1.** Given a standard SLFN with L hidden nodes and activation function g: $R\rightarrow R$ which is infinitely differentiable in any interval, for Q arbitrary distinct samples $(x_q, t_q)$, where $x_q \in R^n$ and $t_q \in R^m$, for any $w_i$ and $b_{ik}$ randomly chosen from any intervals of $R^n$ and R, respectively, according to any continuous probability distribution with probability one, if Q=L, the hidden layer output matrix $H$of the SLFN is invertible and $\|H\cdot\beta - T\|=0$.

**Theorem 2.** Given any small positive value ε>0 and an activation function g: $R\rightarrow R$ which is infinitely differentiable in any interval, there exists Q < L such that for Q arbitrary distinct samples $(x_q, t_q)$, where $x_q \in R^n$ and $t_q \in R^m$, and for any $w_i$ and $b_{ik}$ randomly chosen from any intervals of $R^n$ and $R$, respectively, according to any continuous probability distribution with probability one, $\|H\cdot\beta - T\| < \varepsilon$.

Based on Theorems 1 and 2, there is an extremely simple and efficient method, which could train ELM by the least square method [37]. The ordinary least square (OLS) method to solve the linear system can be shown as follows,

$$\widehat{\beta} = H^{-1} \cdot T, \tag{2}$$

where $\boldsymbol{H^{-1}}$ is the Moore-Penrose generalized inverse of matrix $\boldsymbol{H}$ [38].

### 2.2. Grasshopper optimization algorithm

Grasshoppers are considered pests because they damage crops and hinder agricultural production. The life cycle of a grasshopper is shown in Fig. 3. The unique aspect of grasshopper swarms occurs both during the nymph and adulthood stages [39]. Specifically, the grasshopper nymphs jump and move like rolling cylinders, while adult grasshoppers form a swarm and migrate over large distances.

The grasshopper optimization algorithm's design originates from the insect's behavior in different phases. Slow movement and small steps are the main characteristics of grasshoppers in the larval phase, and long range and abrupt movement are the main features of the insect in adulthood. The meta-heuristic optimization algorithm divides the search process into two tendencies: exploration and exploitation [27]. The search agents move abruptly

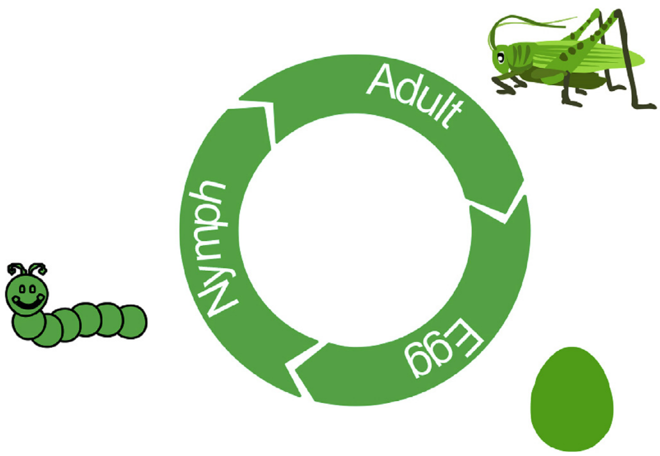


**Fig. 3.** Life cycle of grasshoppers.

in exploration and locally in exploitation. Therefore, the new meta-heuristic optimization algorithm is designed based on the behaviors of a grasshopper swarm [27].

The formula employed to simulate swarm behaviors of grasshoppers is shown as follows,

$$X_i = S_i + G_i + A_i, \tag{3}$$

where $X_i$ is the position of the $i$-th grasshopper and $S_i$ defines the social interaction. Meanwhile, $G_i$ stands for the gravity force on the $i$-th grasshopper and $A_i$ shows the wind advection.

First, the $S$ component in formula (3) is calculated as follows,

$$S_i = \sum_{j=1, j\neq i}^{N} s(d_{ij}) \cdot \widehat{d}_{ij}, \tag{4}$$

where $d_{ij}$ is the distance between the $i$-th and the $j$-th grasshopper, which is calculated as $d_{ij}=|x_j\text{-}x_i|$. $\widehat{d}_{ij} = \frac{x_j - x_i}{d_{ij}}$ is a unit vector from the $i$-th grasshopper to the $j$-th grasshopper. The function $s(r)$ defining the strength of social forces is calculated as follows,

$$s(r) = f \cdot e^{-\frac{r}{l}} - e^{-r}, \tag{5}$$

where $f$ represents the intensity of attraction, $l$ stands for the attractive length scale and $r = \left|\widehat{d}_{ij}\right|$ [29].

Secondly, the $G$ component in formula (3) is calculated as follows,

$$G_i = -g \cdot \widehat{e}_g, \tag{6}$$

where $g$ is the gravitational constant and $\widehat{e}_g$ is a unity vector to the center of earth.

Thirdly, the $A$ component in formula (3) is calculated as follows,

$$A_i = u \cdot \widehat{e}_w, \tag{7}$$

where $u$ and $\widehat{e}_w$ stand for a constant drift and unity vector in the wind's direction.

Then, combining $S$ and $G$ with $A$, the expression can be obtained as follows,

$$X_i = \sum_{j=1, j\neq i}^{N} s(|x_j - x_i|) \cdot \frac{x_j - x_i}{d_{ij}} - g \cdot \widehat{e}_g + u \cdot \widehat{e}_w, \tag{8}$$

where $N$ stands for the number of grasshoppers.

Nevertheless, the grasshoppers quickly reach the comfort zone and the swarm does not converge to a specified point. A modified version to solve optimization problems is proposed as follows,

$$X_i^d = c\left(\sum_{j=1/j\neq i}^{N} c \cdot \frac{ub_d - lb_d}{2} \cdot s\left(\left|x_j^d - x_i^d\right|\right) \cdot \frac{x_j - x_i}{d_{ij}}\right) + \widehat{T}_d \tag{9}$$

where $ub_d$ and $lb_d$ represent the upper bound and lower bound in the $D$-$th$ dimension, respectively. $\widehat{T}_d$is the position of best solution found so far. The coefficient $c$ of each iteration is calculated as follows,

$$c = c\max - l \cdot \frac{c\max - c\min}{L} \tag{10}$$

where $cmax$ and $cmin$ represent the maximum value and minimum value, respectively. The $l$ indicates the current iteration and $L$ is the maximum number of iteration.

Finally, the process of the GOA algorithm is shown in Fig. 4.

### 2.3. Ensemble empirical mode decomposition

The EMD family is well-suited to handling non-stationary and non-linear data, since it can identify the intrinsic oscillatory modes and decompose the data into some oscillatory modes [41]. In general, most decomposition methods perform well when they are applied to data with certain features. For example, the wavelet decomposition method is suitable for data with non-stationary and linear characteristics. Hence, the ensemble empirical mode decomposition (EEMD), which improves upon EMD, is selected to decompose the electrical load. Firstly, the EMD assumes that the data consist of different modes of oscillations simultaneously due to intrinsic complexities hidden in data [40], and such intrinsic mode function (IMF) would be extracted by EMD from the original data [41]. So, the original data series $X(s)$ can be decomposed as some IMFs and a residual,

$$X(s) = \sum\nolimits_{j=1}^{m-1} C_j(s) + R_m(s), \tag{11}$$

where $s$, $R_m(s)$ and $Cj(s)$ stand for the number of IMFs, final residual and IMF, respectively.

However, the mode mixing problem generates in the sifting process, and the reliability of EMD is poor [42]. To address this problem, white noise is added to the data, and the new method EEMD can resolve the mode mixing problem [32]. The white noise must conform to the following rule [42],

$$\varepsilon_{ne} = \frac{\varepsilon}{\sqrt{NE}}, \tag{12}$$

where $NE$ is the ensemble number, $\varepsilon$ is the added noises, and $\varepsilon_{ne}$ is the final standard deviation of errors. The process of EMD and EEMD decomposition is shown in Fig. 5.

### 2.4. The establishment of the hybrid model

In this section, the approach of establishing the hybrid model EEMD-ELM-GOA is introduced in detail. Additionally, the values of parameters in the hybrid model are provided for electrical load

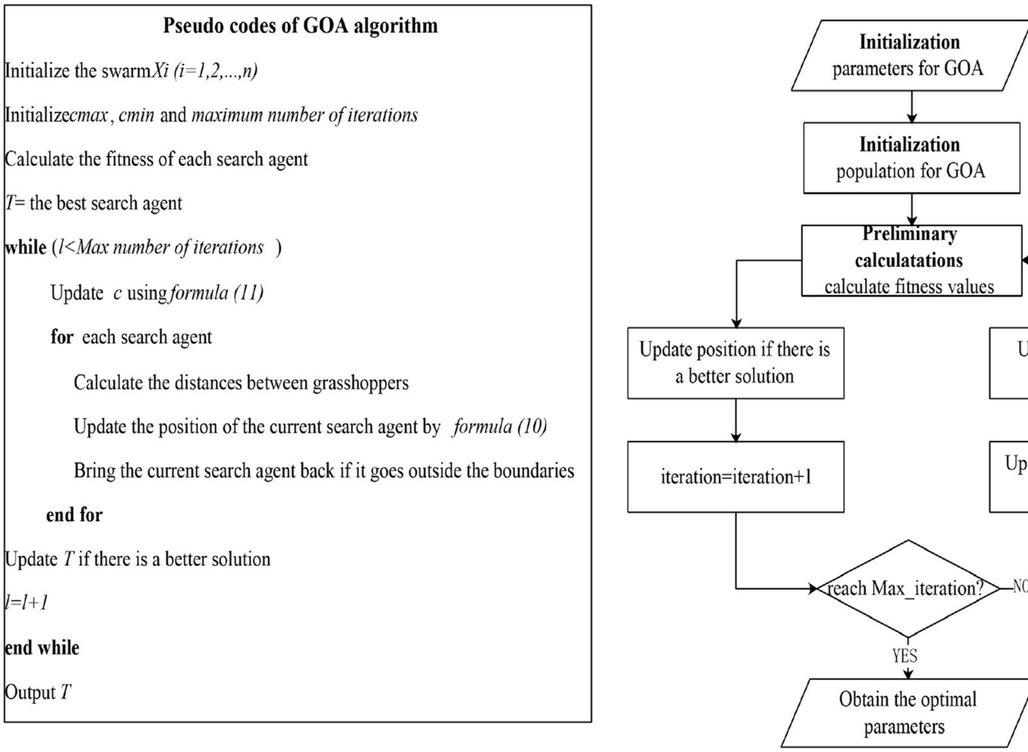


**Fig. 4.** Process of GOA algorithm.

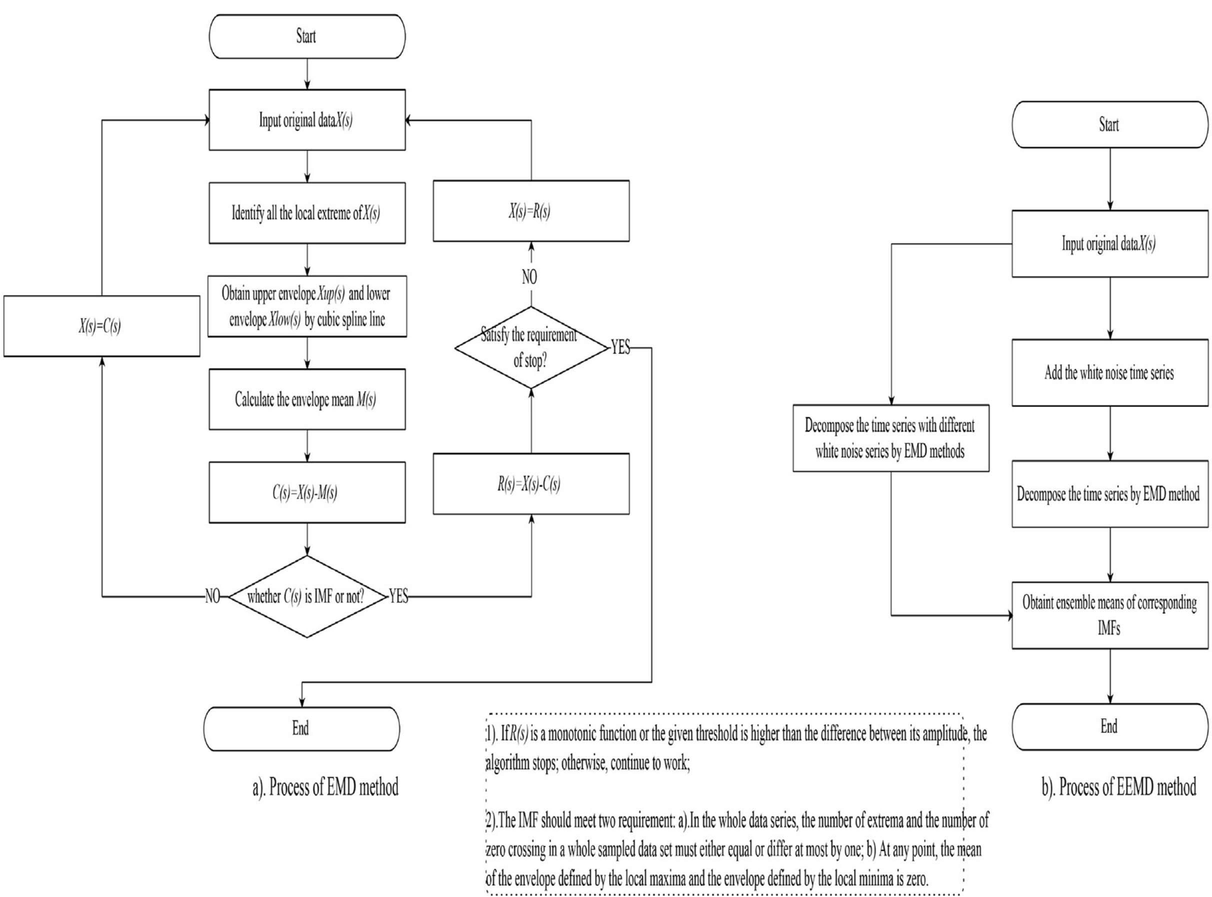


**Fig. 5.** The process of EMD and EEMD.

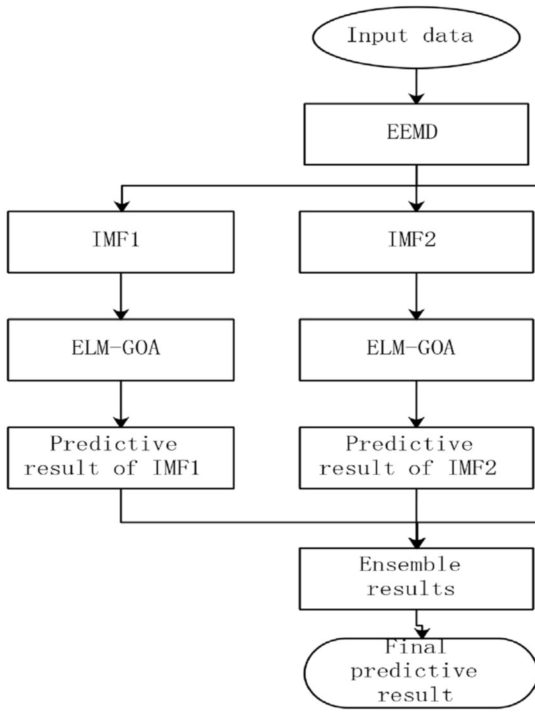


**Fig. 6.** The flow of EEMD-ELM-GOA for electrical load forecasting.

forecasting. The new hybrid model EEMD-ELM-GOA is illustrated in Fig. 6.

First, the EEMD is used to decompose the original electrical load. While training the data set, GOA is then selected to choose suitable parameters for ELM. Finally, ELM is employed to establish the non-linear model between the input and target. In the test data set, the hybrid model EEMD-ELM-GOA obtained by the training data set is tested.

As for electrical load forecasting, the last twelve half-hourly load data ($x_{n-12}$, $x_{n-11}$, $x_{n-10}$, …, $x_{n-2}$, $x_{n-1}$) are selected as the input of EEMD-ELM-GOA, and the target is $x_n$. Therefore, the mapping relationships can be shown as follows,

$$(x_{n-12}, x_{n-11}, x_{10}, ..., x_{n-2}, x_{n-1}) \rightarrow (x_n). \tag{13}$$

The selection of parameters of GOA is performed in Table 1. It must be mentioned that the radial basis function is selected as the active function in the hidden node of ELM in the experiment.

## 3. Simulation results

In this section, simulation results are presented to demonstrate the performance of the proposed model. The proposed hybrid model EEMD-ELM-GOA has the best performance among all nine basic models (SVR, ANN, ELM, WA-ELM, EMD-ELM, EEMD-ELM, EEMD-ELM-DA, EEMD-ELM-GWO, and EEMD-ELM-PSO).

**Table 1**
Selection of parameters of GOA.

| Parameter | value |
|---|---|
| *f* | 0.5 |
| *l* | 1.5 |
| *cmax* | 1 |
| *cmin* | 0.00001 |

### *3.1. Data sets*

To verify the effectiveness of the proposed model, the data sets of electrical load from February 1, 2018 to February 27, 2018 in five regions (NSW, QLD, SA, TAS and VIC) are used as the experimental data (Fig. 7). The five regions are major states in Australia with large populations and active economies. Since electricity is a key demand for both industrial and residential interests, studying the electrical demand in these areas has great significance.

The data sets of electrical load *(MW)* are retrieved from the website of energy security for all Australians *(*http://www.aemo.com.au/*)*. The sample data used in this study are half-hourly electrical loads, and the total number of observations in each data set of these five regions is *1248* (Table 2). Each data set is divided into two subsets: the training data set including *1152* data points (from *2018/2/1 0:30* to *2018/2/25 0:00*), and the remainder used as the testing data set (from *2018/2/25 0:30* to *2018/2/27 0:00*).

### *3.2. Evaluation criteria*

The performance of the referred models can be evaluated in terms of several statistical indicators which measure error between observed and forecasted electrical load values. In this study, the models' performances are evaluated by using the mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE). The performance of the model is more reliable when their value is closer to zero.

These three criteria are defined by the following equations:

$$MAE = \sum_{i=1}^{n} |y_i - \widehat{y}_i| / n, \tag{14}$$

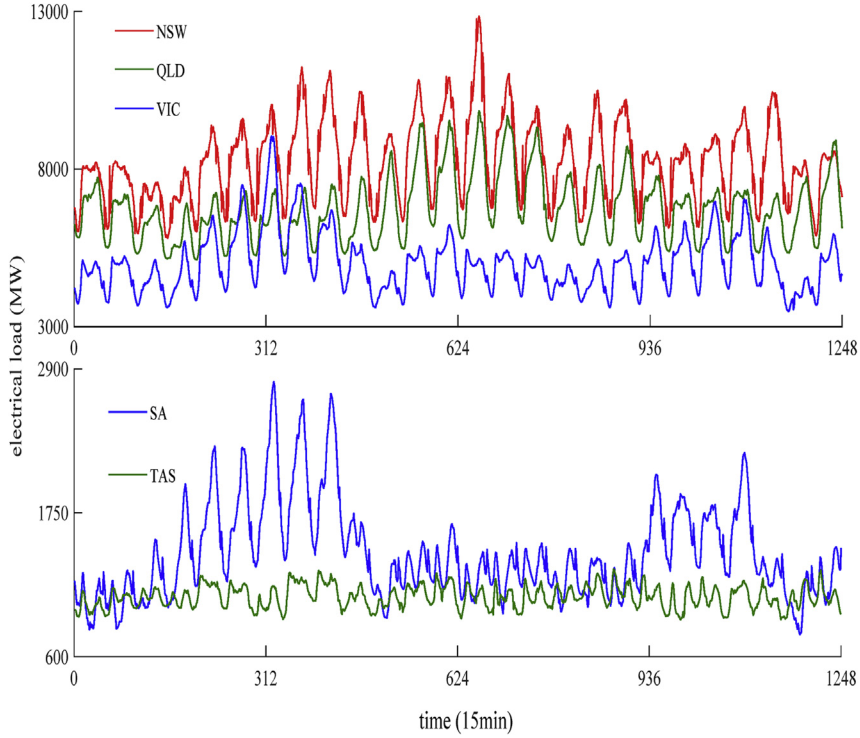


**Fig. 7.** The data sets of five regions, Australia.

**Table 2**
Statistical properties of each data set.

| Region | Data set | N | Min | Max | Mean | Std | S | K |
|---|---|---|---|---|---|---|---|---|
| NSW | train_data | 1152 | 5809 | 12846 | 8288 | 1307 | 0.35 | −0.21 |
| | test_data | 96 | 5884 | 8563 | 7618 | 755 | −0.69 | −0.64 |
| | all_data | 1248 | 5809 | 12846 | 8236 | 1285 | 0.39 | −0.11 |
| QLD | train_data | 1152 | 5127 | 9840 | 6738 | 1021 | 0.73 | 0.20 |
| | test_data | 96 | 5337 | 8911 | 6822 | 1060 | 0.23 | −1.04 |
| | all_data | 1248 | 5127 | 9840 | 6744 | 1023 | 0.69 | 0.08 |
| SA | train_data | 1152 | 816 | 2798 | 1448 | 378 | 1.13 | 1.15 |
| | test_data | 96 | 778 | 1480 | 1134 | 179 | 0.03 | −0.57 |
| | all_data | 1248 | 778 | 2798 | 1423 | 376 | 1.17 | 1.31 |
| TAS | train_data | 1152 | 896 | 1302 | 1082 | 88 | 0.04 | −0.84 |
| | test_data | 96 | 903 | 1294 | 1072 | 90 | 0.29 | −0.26 |
| | all_data | 1248 | 896 | 1302 | 1081 | 89 | 0.05 | −0.80 |
| VIC | train_data | 1152 | 3602 | 9045 | 5049 | 927 | 1.15 | 1.95 |
| | test_data | 96 | 3482 | 5945 | 4466 | 692 | 0.51 | −0.92 |
| | all_data | 1248 | 3482 | 9045 | 5004 | 924 | 1.13 | 1.96 |

Min:minimum, Max: maximum, Std: standard deviation, S: skewness, K: kurtosis.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}, \tag{15}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right| * 100\%. \tag{16}$$

In these expressions, $y_i$ is the observed value, $\hat{y}_i$is the predicted value of $y_i$, $\bar{y}$ is the average of the observed value, and $n$ is the number of observations.

### 3.3. The selection of main algorithm

To choose the best data mining algorithm for obtaining the relationship between the system's input and output for short-term load forecasting, three main algorithms (SVR, ANN and ELM) have been selected to acquire the system. In addition, the former situations $(x_t, x_{t-1}, \ldots, x_{t-11})$ are utilized as the input of the system, and the following situation $x_{t+1}$ is used for the target of the system. Then, the electrical load from five major states in Australia is applied to verify a suitable main algorithm for short-term load forecasting (Fig. 8 and Table 3).

Table 3 summarizes the performances of SVR, ANN, and ELM for 1-step ahead electrical load forecasting of NSW, QLD, SA, TAS, and VIC. From the results, it can be observed that the models' performances based on neural network are better than the SVR model. Especially, almost three criteria for SVR are significantly larger than those of ANN and ELM. For example, the RMSE, MAPE, and MAE of SVR are *2351*, *2.51,* and *187,* respectively, while these figures for ANN are *146*, *1.48*, and *112*. However, it is apparent that the performance of ANN is almost as the same as that of ELM. Specifically, the three criteria of ANN are close to those of ELM. On the other hand, it should not be ignored that ELM costs less time than ANN because it needs many experimental samples to train the network until the error of the model is limited to a preconcerted range. Meanwhile the weights of ELM can be determined by OLS method, which need not any training.

In short, based on the performances of the three main algorithms, ELM is selected as one part of a hybrid model for short-term electrical load forecasting.

### 3.4. The selection of decomposition method

To choose the best decomposition method of electrical load, three approaches (WA, EMD and EEMD) are applied to decompose the original series. The most suitable technique can be confirmed through the simulation of short-term load forecasting based on the data set from Australia. Besides, the *db3* is regarded as the wavelet function for wavelet decomposition (WA) [2], and the ensemble member and standard deviation of white noise are set to *50* and *0.4*,

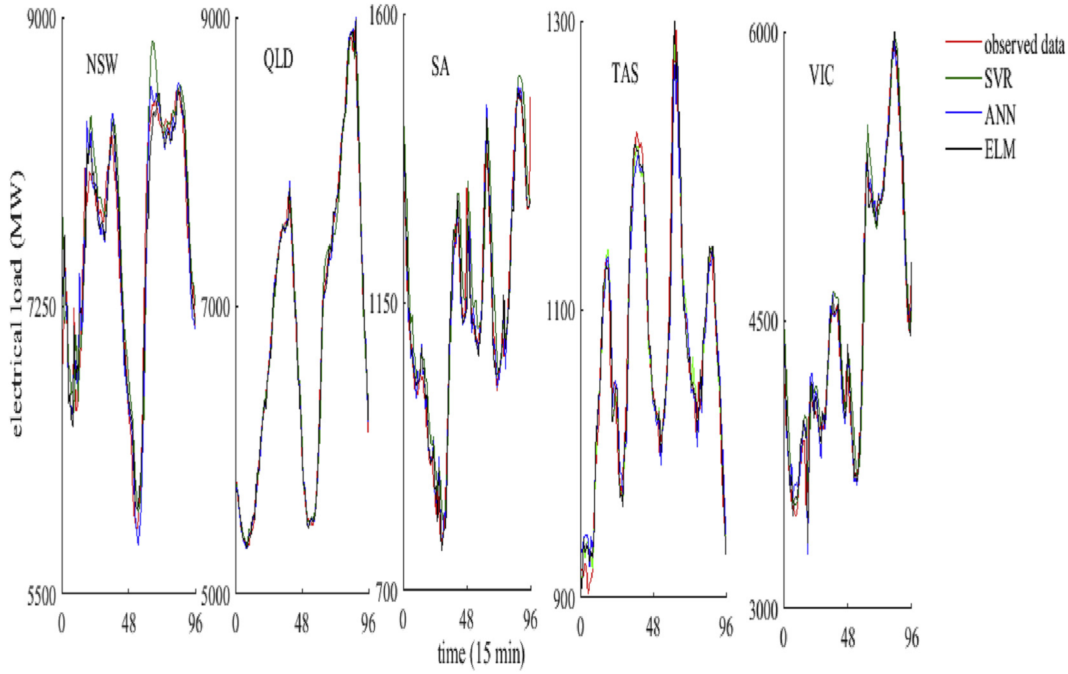


**Fig. 8.** Results of three main algorithms for load forecasting.

**Table 3**
Indicators of three main algorithms for electrical load forecasting.

| Criteria | SVR | | | ANN | | | ELM | | |
|---|---|---|---|---|---|---|---|---|---|
| | RMSE | MAPE | MAE | RMSE | MAPE | MAE | RMSE | MAPE | MAE |
| NSW | 2351 | 2.51 | 187 | 1468 | 1.48 | 112 | 1651 | 1.80 | 135 |
| QLD | 907 | 1.01 | 71 | 717 | 0.75 | 52 | 705 | 0.79 | 54 |
| SA | 552 | 4.07 | 46 | 380 | 2.41 | 27 | 381 | 2.24 | 25 |
| TAS | 168 | 1.26 | 13 | 192 | 1.40 | 15 | 162 | 1.19 | 13 |
| VIC | 1323 | 2.42 | 105 | 975 | 1.73 | 73 | 961 | 1.70 | 73 |

respectively. To lower the complexity of calculation in the hybrid model, the original series are decomposed to three *IMFs*. For example, the results of decomposition based on the EEMD technique are shown for each region in Fig. 9.

After the decomposition of electrical load, the ELM, which is selected in Section 4.3, is applied to obtain the system for each subsequence to predict the 1-step ahead electrical load for five regions. Then, the results of each series are added together to compute the final predictive value. The three criteria for predictive results can be acquired to show the errors of different decomposition methods, and the performances are shown in Table 4. Apart from that, the results of ELM with different decomposition methods are performed in Fig. 10.

The fitted curves of all five regions (NSW, QLD, SA, TAS, and VIC) for 1-step ahead electrical load prediction by the EEMD-ELM model are shown in Fig. 10. The results indicate that the electrical load values estimated by the EEMD-ELM model are closely matched with observed values. There are three primary methods to decompose the original sequence into sub-sequences. From the results in Table 4, it is clear that the performances of model EEMD-ELM for 1-step electrical load forecasting in all regions have superior accuracy to that of new models WA-ELM and EMD-ELM based on the three error indicators (MAE, RMSE, MAPE). Through the simulation of electrical load forecasting, EEMD is regarded as the ideal technique to establish the hybrid model in the following section.

### 3.5. The performance of proposed hybrid mode EEMD-ELM-GOA

When the experiments are carried out, the maximum number of iterations of all hybrid models is set at *30* and the number of search agents is *15*. Meanwhile, the number of input nodes $n$ and hidden nodes $L$ is *12* and *20*, respectively. According to the results, the performance of hybrid model EEMD-ELM-GOA is shown as follows.

Table 5 summarizes the performance of proposed hybrid model EEMD-ELM-GOA for 1-step ahead electrical load forecasting of NSW, QLD, SA, TAS, and VIC, Australia. The results demonstrate that the performance of EEMD-ELM-GOA model for QLD is better than that for the other four regions. In all five regions, the accuracy of forecasted results is different, but the values of the three criteria are quite small. For example, the RMSE, MAPE, and MAE of EEMD-ELM-GOA in NSW are *603*, *0.54,* and *40*, respectively, and *564*, *0.46,* and *32*, for Queensland. The results indicate that the forecast is appropriate. In terms of MAE, MAPE, and RMSE, a GOA-ELM model with relatively low error can provide acceptable forecast to 1-step ahead electrical load time series for the five states. The results of proposed the hybrid model EEMD-ELM-GOA are shown in more detail in Fig. 11.

### 3.6. Comparison of different optimization algorithms

In this section, different experiments are carried out to confirm that the combination of GOA and EEMD-ELM offers the best performance. This simulation is carried out using three optimizers to obtain the ELM's parameters. The first main optimization algorithm is PSO, which is proposed by Eberhart and Kennedy in 1995, to show the classical optimization performance for the ELM's parameter selection [16]. To compare with the recent meta-heuristic optimizer, the grey wolf optimizer (GWO), which presented in 2014, has been popular in parameter selection for machine learning algorithms [43]. The third optimization, dragonfly algorithm (DA), as the one of the latest optimizers, is used to perform the new GOA's outstanding performance in searching parameters for ELM [44]. To assess the ability of the novel

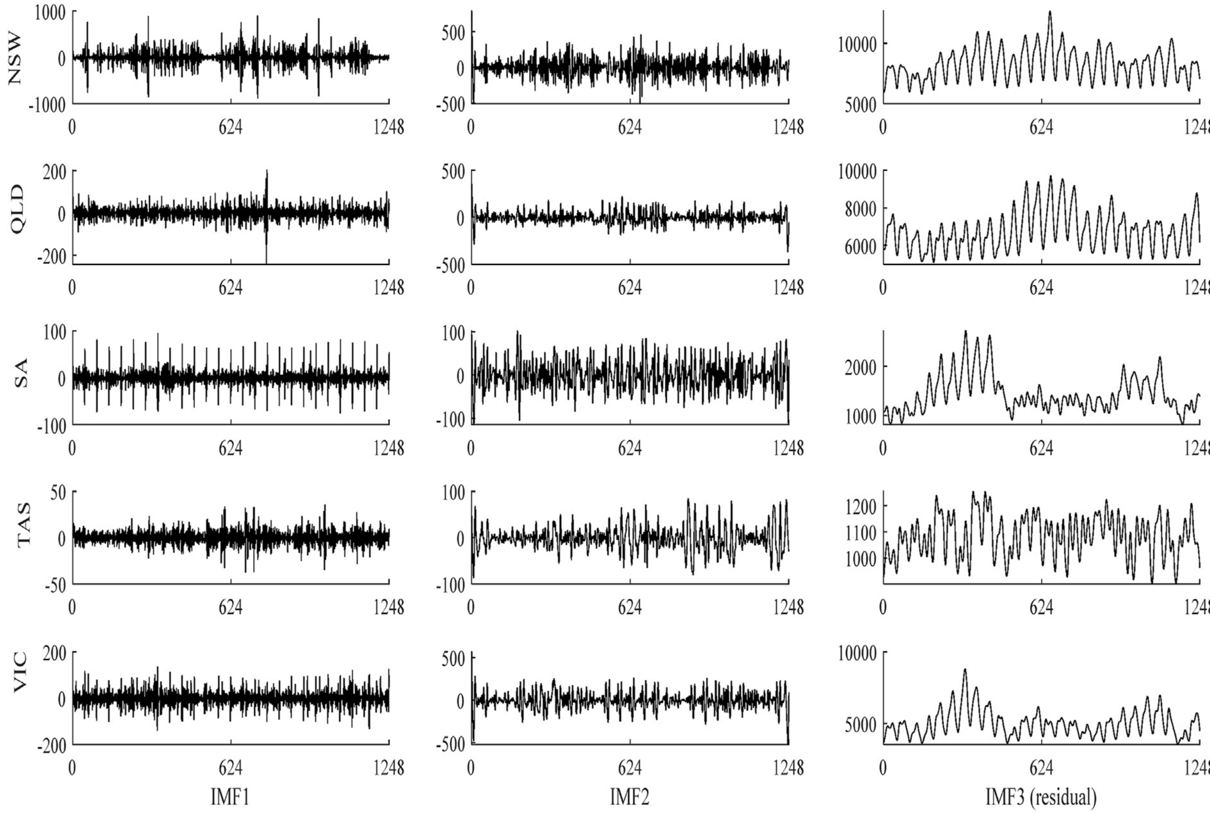


**Fig. 9.** The results of decomposition by EEMD for five states.

**Table 4**
Indicators of different decomposition methods for electrical load forecasting.

| Criteria | WA-ELM | | | EMD-ELM | | | EEMD-ELM | | |
|---|---|---|---|---|---|---|---|---|---|
| | RMSE | MAPE | MAE | RMSE | MAPE | MAE | RMSE | MAPE | MAE |
| NSW | 1150 | 1.27 | 93 | 1275 | 1.32 | 96 | 850 | 0.89 | 64 |
| QLD | 1582 | 1.64 | 112 | 1242 | 1.13 | 80 | 469 | 0.50 | 34 |
| SA | 467 | 3.59 | 40 | 494 | 2.89 | 33 | 253 | 1.82 | 20 |
| TAS | 213 | 1.64 | 18 | 194 | 1.50 | 16 | 102 | 0.77 | 8 |
| VIC | 1128 | 2.15 | 93 | 829 | 1.39 | 58 | 932 | 1.34 | 60 |

optimization algorithm GOA in relation to the DA, GWO, and PSO for 1-step ahead electrical load forecasting, the EEMD-ELM-DA, EEMD-ELM-GWO and EEMD-ELM-PSO models are developed. There are numerous optimization algorithms, which have been proposed in recent years, so the basic algorithms are selected from new optimization algorithms. To compare with classical algorithms, PSO is applied to build the basic model [45]. Moreover, since the optimization algorithms are applied while searching for the suitable parameters of ELM, the dimension $DI$ of the optimal results is calculated by the following formula (17),

$$DI = n \times L + L, \tag{17}$$

Thus, there are 260 parameters of ELM, which are searched by the optimizers.

The optimal results for the parameter of ELM by four different

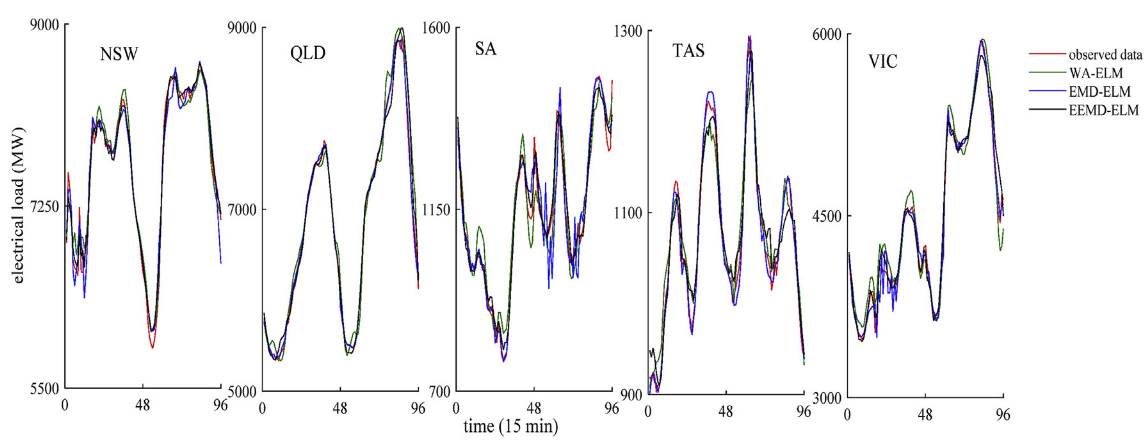


**Fig. 10.** Results of ELM with different decomposition methods.

**Table 5**
Indicators of different optimizers for electrical load forecasting.

| Model | Criteria | NSW | QLD | SA | TAS | VIC |
|---|---|---|---|---|---|---|
| EEMD-ELM-GOA | RMSE | 603 | 564 | 154 | 61 | 419 |
| | MAPE | 0.54 | 0.46 | 0.95 | 0.46 | 0.64 |
| | MAE | 40 | 32 | 11 | 5 | 28 |
| EEMD-ELM-DA | RMSE | 684 | 336 | 155 | 81 | 499 |
| | MAPE | 0.69 | 0.35 | 1.01 | 0.60 | 0.92 |
| | MAE | 50 | 25 | 11 | 6 | 39 |
| EEMD-ELM-PSO | RMSE | 839 | 941 | 261 | 257 | 877 |
| | MAPE | 0.77 | 1.09 | 1.71 | 1.90 | 1.36 |
| | MAE | 56 | 70 | 20 | 20 | 59 |
| EEMD-ELM-GWO | RMSE | 766 | 558 | 207 | 139 | 975 |
| | MAPE | 0.72 | 0.48 | 1.33 | 1.01 | 1.15 |
| | MAE | 52 | 34 | 15 | 11 | 52 |

optimization algorithms are presented in Supplementary Material A, B, C, and D, where there are significant gaps between the optimal results by different optimization algorithms. The results of these four considered hybrid models are performed in Fig. 11. Further, as shown in Table 5, the accuracy of EEMD-ELM-GOA is greater than those of EEMD-ELM-DA, EEMD-ELM-GWO and EEMD-ELM-PSO for the same electrical data. More importantly, it is evident that the ability to search for the parameters of GOA for ELM is more efficient than that PSO, DA, and GWO for ELM.

After all, the combination of GOA and EEMD-ELM can not only obtain a high accuracy for 1-step ahead electrical load prediction, but it can also deliver the highest efficiency among all basic models. Thus, the study chooses the hybrid EEMD-ELM-GOA model to forecast.

## 4. Conclusion

Accurate electrical load prediction is essential for the planning and maintenance of electrical power systems. In this paper, the hybrid model EEMD-ELM-GOA for forecasting 1-step ahead electrical load is constructed to enhance prediction reliability and accuracy.

There are four main findings which can be summarized as follows,

(1). Compared with SVR, the ELM has improved by 0.71% in predictive performance, while its accuracy also is 0.01% higher than traditional ANN. This is because the artificial neural network has a great structure to obtain the non-linear system. The ELM does not require the training process, so it is more computationally more efficient the ANN with the same accuracy;

(2). As for the decomposition techniques, the EEMD method has a great performance, whose accuracy is 0.99% higher than WA's. This means the Hilbert transform is superior to wavelet transform for electrical load time series. In addition, compared with EMD, the EEMD method has improved 0.55% accuracy due to the additional white noise;

(3). The state-of-the-art optimizer GOA is found to perform very well in searching the suitable parameters for ELM for electrical load forecasting with lower errors. Specifically, GOA's error is 0.10%, 0.76%, which is 0.33% smaller than DA's, PSO's and GWO's, respectively. The reason is that this optimizer has an extremely strong ability of exploring and exploiting the optimum solution for high-demission problems;

(4). Overall, the proposed hybrid model EEMD-ELM-GOA can provide satisfactory accuracy and the highest efficiency among all the basic models, which is promising for application in short-term electrical load forecasting.

Although the results indicate that the EEMD-ELM-GOA can be successfully applied to forecast short-term electrical load, there are a few limitations to consider. In this current work, EEMD-ELM-GOA is unable to predict abnormal events. For example, change in government policies may bring some influences on some large-scale cooperation, like mining companies. Extreme weather might also cause a heavier-than-normal electrical load. Further research can incorporate the effects of abnormal events by modeling these rare events as outliers.

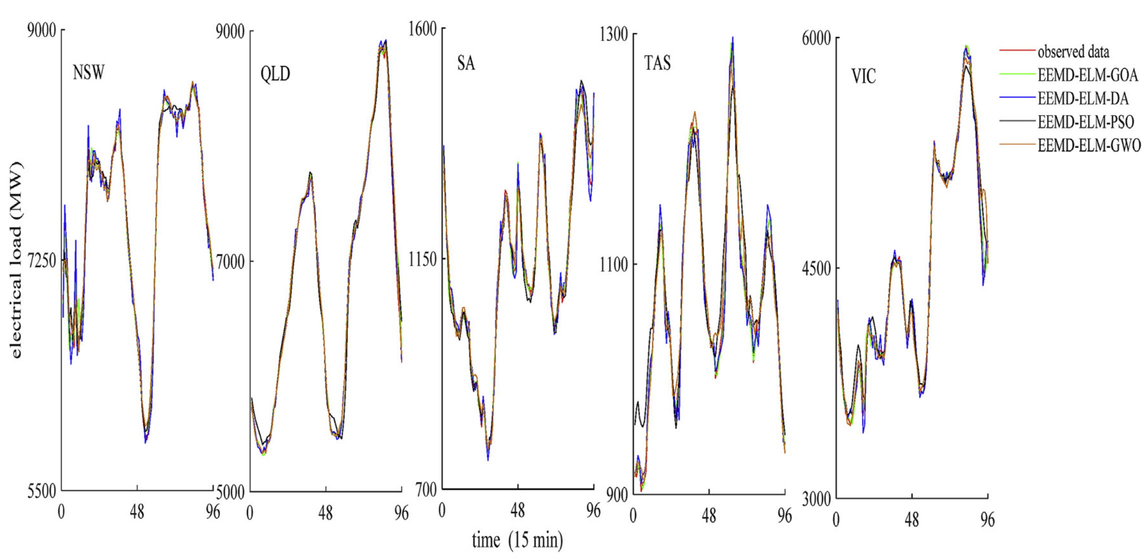


**Fig. 11.** The results of the hybrid model with different optimizers.

## Author contributions

Jinran Wu developed the first new model and implemented the algorithms; Zhesen Cui helped do some simulation experiments; Yanyan Chen drew the figures for this paper; Demeng Kong drafted the first manuscript version; You-Gan Wang has supervised this work, in particular, he contributed towards all the revision aspects.

## Conflicts of interest

The authors declare no conflict of interest.

## Acknowledgments

This research was partially funded by the Australian Research Council project (DP160104292).

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.energy.2018.10.076.

## Abbreviations

| | |
|---|---|
| ANN | artificial neural network |
| DA | dragonfly algorithm |
| EEMD | ensemble empirical mode decomposition |
| ELM | extreme learning machine |
| EMD | empirical mode decomposition |
| GA | genetic algorithm |
| GOA | grasshopper optimization algorithm |
| GWO | grey wolf optimizer |
| MAE | mean absolute error |
| MAPE | mean absolute percentage error |
| NN | neural network |
| NSW | New South Wales |
| PSO | particle swarm optimization |
| QLD | Queensland |
| RMSE | root mean square error |
| SA | South Australia |
| SLFN | single-hidden layer feedforward neural network |
| SVR | support vector regression |
| TAS | Tasmania |
| VIC | Victoria |
| WA | wavelet decomposition |

## References


[1] Wang J, Chi D, Wu J, et al. Chaotic time series method combined with particle swarm optimization and trend adjustment for electricity demand forecasting. Expert Syst Appl 2011;38(7):8419–29.

[2] Li L, Wang H. A VVWBO-BVO-based GM (1,1) and its parameter optimization by GRA-IGSA integration algorithm for annual power load forecasting. PloS One 2018;13(5):e0196816. https://doi.org/10.1371/journal.pone.0196816.

[3] Wu D, Amin Amini, Ali Razban, et al. ARC algorithm: a novel approach to forecast and manage daily electrical maximum demand. Energy 2018;154(C):383–9.

[4] Xiao L, Shao W, Yu M, et al. Research and application of a combined model based on multi-objective optimization for electrical load forecasting. Energy 2017;119:1057–74.

[5] Li W, Kong D, Wu J. A novel hybrid model based on extreme learning machine, k-nearest neighbor regression and wavelet denoising applied to short-term electric load forecasting. Energies 2017;10(5).

[6] Boroojeni KG, Amini MH, Bahrami S, et al. A novel multi-time-scale modeling for electric power demand forecasting: from short-term to medium-term horizon. Elec Power Syst Res 2017;142:58–73.

[7] Zhao H, Guo S. An optimized grey model for annual power load forecasting. Energy 2016;107:272–86.

[8] Mamllo R, Badran O, Abdulhadi E. A fuzzy inference model for short-term load forecasting. Energy Pol 2009;37(4):1239–48.

[9] Kamyab F, Bahrami S. Efficient operation of energy hubs in time-of-use and dynamic pricing electricity markets. Energy 2016;106(1):343–55.

[10] Chen Y, Xu P, Chu Y, et al. Short-term electrical load forecasting using the Support Vector Regression (SVR) model to calculate the demand response baseline for office buildings. Appl Energy 2017;195:659–70.

[11] Hu R, Wen S, Zeng Z, Huang T. 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.

[12] Zhang X, Wang J, Zhang K. Short-term electric load forecasting based on singular spectrum analysis and support vector machine optimized by Cuckoo search algorithm. Elec Power Syst Res 2017;146:270–85.

[13] Yang Y, Che J, Li Y, Zhao Y, Zhu S. An incremental electric load forecasting model based on support vector regression. Energy 2016;113:796–808.

[14] Zhang X, Wang J, Zhang K. Short-term electric load forecasting based on singular spectrum analysis and support vector machine optimized by Cuckoo search algorithm. Elec Power Syst Res 2017;146:270–85.

[15] Hong WC. Chaotic particle swarm optimization algorithm in a support vector regression electric load forecasting model. Energy Convers Manag 2009;50(1):105–17.

[16] Khwaja AS, Zhang X, Anpalagan A, Venkatesh B. Boosted neural networks for improved short-term electric load forecasting. Elec Power Syst Res 2017;143:431–7.

[17] Xiao L, Shao W, Liang T, Wang C. A combined model based on multiple seasonal patterns and modified firefly algorithm for electrical load forecasting. Appl Energy 2016;167:135–53.

[18] Mosbah H, El-Hawary ME. Optimization of neural network parameters by Stochastic Fractal Search for dynamic state estimation under communication failure. Elec Power Syst Res 2017;147:288–301.

[19] Xiao L, Shao W, Wang C, Zhang K, Lu H. Research and application of a hybrid model based on muti-objective optimization for electrical load forecasting. Appl Energy 2016;180:213–33.

[20] Huang GB, Zhu QY, Siew CK. Extreme learning machine: theory and applications. Neurocomputing 2006;70(1):489–501.

[21] Ma Z, Luo G, Huang D. Short term traffic flow prediction based on on-line sequential extreme learning machine. In: Eighth international conference on advanced computational intelligence. IEEE; 2016. p. 143–9.

[22] Zeng N, Zhang H, Liu W, Liang J, Alsaadi FE. A switching delayed PSO optimized extreme learning machine for short-term load forecasting. Neurocomputing 2017;240:175–82.

[23] Li S, Goel L, Wang P. An ensemble approach for short-term load forecasting by extreme learning machine. Appl Energy 2016;170:22–9.

[24] Ömer Faruk Ertugrul. Forecasting electricity load by a novel recurrent extreme learning machines approach. Int J Electr Power Energy Syst 2016;78:429–35.

[25] Zeng N, Zhang H, Liu W, et al. A switching delayed PSO optimized extreme learning machine for short-term load forecasting. Neurocomputing 2017;240:175–82.

[26] Rashedi E, Nezamabadi-Pour H, Saryazdi S. GSA: a gravitational search algorithm. Elsevier Science Inc; 2009.

[27] Mirjalili S, Lewis A. The whale optimization algorithm. Adv Eng Software 2016;95:51–67.

[28] Olorunda O, Engelbrecht AP. Measuring exploration/exploitation in particle swarms using swarm diversity. In: Evolutionary computation. IEEE; 2008. p. 1128–34.

[29] Saremi S, Mirjalili S, Lewis A. Grasshopper optimisation algorithm: theory and application. Adv Eng Software 2017;105:30–47.

[30] Starck JL, Fadili J, Murtagh F. The undecimated wavelet decomposition and its reconstruction. IEEE Press; 2007.

[31] 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 Physical & Engineering Sciences 1998;454(1971):903–95.

[32] Wu Z, Huang NE. Ensemble empirical mode decomposition: a noise-assisted data analysis method. Adv Adapt Data Anal 2009;1(01):1–41.

[33] Wang D, Luo H, Grunder O, Lin Y, Guo H. Multi-step ahead electricity price forecasting using a hybrid model based on two-layer decomposition technique and BP neural network optimized by firefly algorithm. Appl Energy 2017;190:390–407.

[34] Ghofrani M, Ghayekhloo M, Arabali A, Ghayekhloo a. A hybrid short-term load forecasting with a new input selection framework. Energy 2015;81:777–86.

[35] Zhang J, Wei Y, Li D, Tan Z, Zhou J. Short term electricity load forecassting using a hybrid model. Energy 2018;158:774–81.

[36] Li W, Chang L. A combination model with variable weight optimization for short-term electrical load forecasting. Energy 2018;164:575–93.

[37] Banerjee KS. Generalized inverse of matrices and its applications. Technometrics 1971;15(1). 197-197.

[38] Serre D. Matrices : theory and applications. Mathematics 2002;(32):221.

[39] Aĭzenberg LA, Dautov ŠA. Mechanosensory-induced behavioral gregarization in the desert locust Schistocerca gregaria. J Exp Biol 2003;206(Pt 22):3991–4002.

[40] Huang NE, Wu MLC, Long SR, et al. A confidence limit for the empirical mode decomposition and Hilbert spectral analysis. Proceedings Mathematical Physical & Engineering Sciences 2003;459(2037):2317–45.

[41] Tang Ling, Dai Wei, Yu Lean, et al. A novel CEEMD-based EELM ensemble learning paradigm for crude oil price forecasting. Int J Inf Technol Decis Making 2015;14(01):141–69.

[42] Yeh JR, Shieh JS, Huang NE. Complementary ensemble empirical mode decomposition: a novel noise enhanced data analysis method. Adv Adapt Data Anal 2010;2(02):135–56.
[43] Mirjalili S, Mirjalili SM, Lewis A. Grey wolf optimizer. Adv Eng Software 2014;69(3):46–61.
[44] Mirjalili S. Dragonfly algorithm: a new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems. Neural Comput Appl 2016;27(4):1053–73.
[45] Li W, Kong D, Wu J. A new hybrid model FPA-SVM considering cointegration for particular matter concentration forecasting: a case study of kunming and yuxi, China. Comput Intell Neurosci 2017;2017(366, part 1):1–11.