

# A Combined Model Based on Neural Networks, LSSVM and Weight Coefficients Optimization for Short-Term Electric Load Forecasting

Caihong Li<sup>(✉)</sup>, Zhaoshuang He, and Yachen Wang

School of Information Science and Engineering, Lanzhou University,  
Lanzhou 730000, People's Republic of China  
licaihong@lzu.edu.cn

**Abstract.** As an essential energy in the daily life, electricity which is difficult to store has become a hot issue in power system. Short-term electric load forecasting (STLF) which is regarded as a vital tool helps electric power companies make good decisions. It can not only guarantee adequate energy supply but also avoid unnecessary wastes. Although there exists quantity of forecasting methods, most of them are not able to make accurate predictions. Therefore, a forecasting method with high accuracy is particularly important. In this paper, a combined model based on neural networks and least squares support vector machine (LSSVM) is proposed to improve the forecasting accuracy. At first, three forecasting methods named generalized regression neural network (GRNN), Elman, LSSVM are utilized to forecast respectively. Among them, simulate anneal (SA) arithmetic is used to optimize GRNN. Then, SA is employed to determine the weight coefficients of each individual method. At last, multiplying all the three forecasting results with the corresponding weights, the final result of the combined model can be attained. Using the electric load data of Queensland of Australia as experimental simulation, case studies show that the proposed combined model works well for STLF and the results prove more accurate.

**Keywords:** Generalized regression neural network · Elman · Least squares support vector machine simulated annealing algorithm · Short-term electric load forecasting

## 1 Introduction

Electricity is of vital importance to every region as an essential energy resource in people's daily life. Especially in the underdeveloped areas where electricity is deficient, thus makes it more precious. Short-term electric load forecasting (STLF) is a fundamental tool for transmission dispatching, unit commitment and other public utilities [1]. During the past decades, various models based on mathematics or artificial intelligence [2–7] have been introduced to electric load forecasting. All these methods improve the forecasting performance to a certain degree. However, the individual methods are too simple to comprehensively deal with the forecasting problems. As Moghram and Rahman [8] stated, all the individual methods failed to yield desirable forecasting

performance. The forecasting errors will have a tremendous impact on the economic benefits. In the other words, small increase of forecasting accuracy may save millions of dollars cost which are huge to the electricity market. In order to enhance the forecasting performance, emphases have laid on the hybrid models or combined models. By integrating multiple methods, the forecasting method can process different data with different characteristics and overcome flaws existing in the individual methods. Amina et al. [9] proposed a novel fuzzy wavelet neural network model of Greek Island of Crete. By using the subtractive clustering optimized with the Expectation-Maximization algorithm, the results indicated that it provided significantly better forecasts. Xiao et al. [10] introduced a hybrid forecasting model combined with BP and GRNN. Ahead of forecasting, data preprocessing was used to improve the forecasting accuracy. A hybrid method with empirical mode decomposition (EMD), extended Kalman filter, extreme learning machine (ELM) and particle swarm optimization (PSO) was utilized by Liu et al. [11] to forecast short-term load. [12–14] proposed the examinations of the combined methods. All these hybrid and combined methods enhance the forecasting performance to a large extent. That makes them the preferred methods when considering the forecasting accuracy.

In this paper, a combined model based on generalized regression neural network (GRNN), Elman, least squares support vector machine (LSSVM) and simulated annealing (SA) is introduced. GRNN is a kind of radial basis function (RBF) network which has a strong nonlinear mapping capacity and high degree of fault tolerance. Its network structure is simple and the calculation results can achieve global convergence. Xia et al. [15] applied GRNN for short-term load forecasting and virtual instrument design. Chelgani et al. [16] used GRNN to predict microwave irradiation pretreatment and peroxyacetic acid desulfurization of coal. The results turned that GRNN was good at predicting. Li et al. [17] employed a hybrid model based on GRNN which was optimized by fruit fly optimization algorithm. Elman is a dynamic neural network which has a short-term memory. That makes it well accommodate to the time-variant characteristic. Under the influence of temperature, a wavelet Elman neural network for short-term load prediction was proposed by Kelo et al. [18]. Song [19] introduced the Elman networks on the weight convergence. Li et al. [20] utilized Chaotifying linear Elman networks. LSSVM is one of the machine learning methods which has strong nonlinear data processing ability. LSSVM is able to reduce the complexity of the calculation and improve the speed of solving. Shayeghi et al. [21] proposed a hybrid model to forecast day-ahead electricity price. Chaotic gravitational search algorithm was developed to find the optimal parameters of LSSVM. Zhang et al. [22] utilized an unbiased LSSVM with polynomial kernel. Xie et al. [23] applied clustering-LSSVM to forecast electricity price. SA is a stochastic optimization algorithm which based on Monte-Carlo iterative solution strategy. It can effectively avoid falling into local minimum and finally it approaches the global optimal. In this paper, SA not only optimizes the parameters of GRNN, but also determines the weight coefficients of the three individual methods. Hong [24] introduced seasonal SVR which is optimized by simulated annealing algorithm to forecast traffic flow. Yuan et al. [25] proposed the Cloud theory-based simulated annealing algorithm and application. Pai et al. [26] utilized support vector machine optimized by simulated annealing algorithm in electricity load forecasting. Firstly, three

individual forecasting methods are applied to forecast respectively. Then, SA is employed to determine the weight coefficients of each individual method. At last, multiplying all the three forecasting results with the corresponding weights, the final result of the combined model can be attained.

The rest of this paper is organized as follows. Section 2 introduces the theory of GRNN, Elman, LSSVM and SA which combined the proposed model. The implementation process of the proposed combined model is described in Sect. 3. In Sect. 4, a simulation of electric load forecasting of Australia electric market is shown. The contrasted results demonstrate the preponderance of the proposed method. Finally, Sect. 5 concludes the paper.

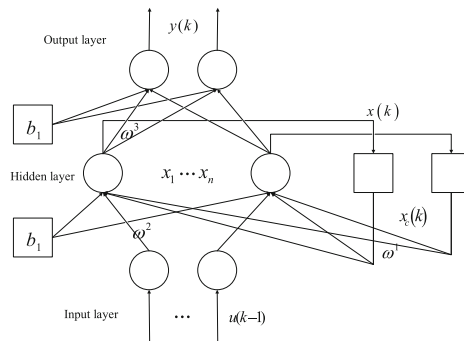
## 2 Methodologies

All the individual methods are introduced in this section, including generalized regression neural network, elman, least squares support vector machine and simulated annealing algorithm.

### 2.1 Elman Neural Network

The Elman neural network(ElmanNN) first proposed by Elman in 1990 [27], which structure includes the input layer, a particular context nodes input layer, the hidden layer(middle layer) and the output layer, is a feed-forward network with local feedback. The linear or nonlinear function is applied for the transfer function of Elman. The connection of each layer is similar to the feed-forward network and the context layer that can be seen as a one-step delay operator is used to record before moment output values of hidden layer. The framework of ElmanNN is shown in the Fig. 1 and its state space is expressed as:

$$y(k) = g(\omega^3 x(k)) \tag{1}$$



**Fig. 1.** The framework of a feedback ElmanNN with three layer structure.

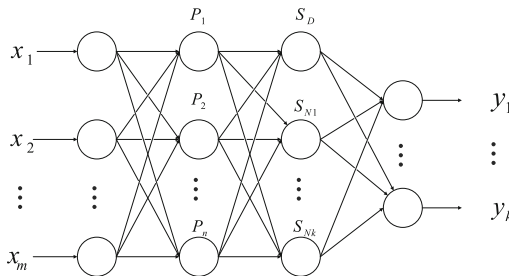
$$x(k) = f(\omega^3 x_c(k) + \omega^2(u(k-1))) \tag{2}$$

$$x_c(k) = x(k-1) \tag{3}$$

In the Fig. 1,  $y$  is  $m$ -dimension output vector;  $x$  is  $n$ -dimension hidden layer unit vector;  $u$  express  $r$ -dimension input vector;  $x_c$  is  $n$ -dimension feedback state vector;  $\omega^3$  is the weight from hidden layer to output layer;  $\omega^2$  is the weight from input layer to hidden layer;  $\omega^1$  is the weight from context layer to hidden layer;  $b_1$  and  $b_2$  are the threshold value of hidden layer and output layer, respectively;  $g(*)$  and  $f(*)$  are the transfer function of the output neurons and of hidden neurons, respectively.

### 2.2 General Regression Neural Network

The general regression neural network (GRNN) proposed by Donald F. Specht in 1991 [28] is a variation of radial basis neural networks which is designed for function approximation and regression. A GRNN consisted of four layers: the input layer, the pattern layer, the summation layer, and the output layer is a novel effective feed-forward neural network model with the standardization of the dot product weight function. Radial Basis Function (RBF) is used as transformation from input layer to hidden layer, while, a special linear transformation is used from hidden layer to output layer. In addition, the network can avoid the impact on the prediction result caused by human subjective assumptions as much as possible, because of that in GRNN there is only one artificially adjust parameters named smooth factor and network learning are all depended on the data sample. Figure 2 shows the structure of GRNN.



**Fig. 2.** The structure of the general regression neural network

The first layer is input layer which number of neurons is equal with the number of the input parameters. The second layer is RBF hidden layer which number of neurons is the training sample number. The Gaussian function is normally used to be its transfer function, which contain smooth factor. The smaller the smooth factor, the stronger the function approximation ability. The third layer was simple linear output layer.

### 2.3 Least Squares Support Vector Machine

Least squares support vector machine (LSSVM) which is developed based on the statistic theory is a new kind of machine learning techniques. Suyken et al. [29] proposed LSSVM by changing the constraint condition and risk function of SVM. The quadratic programming problem can be directly converted into linear equations. That means LSSVM has the ability to accelerate the solving speed and reduce the use of computing resources which can improve the convergence speed.

For a given sample  $\{x_i, y_i\}$ , where  $x_i \in R^n$  is the output vector,  $i = 1, 2, \dots, l$ ,  $y_i \in R$  is the output. By using the structural risk minimization principle, the optimization problem is expressed as the following constrained optimization problem:

$$\min: J(\omega, e) = \frac{1}{2} \|\omega\|^2 + \frac{1}{2} C \sum_{i=1}^N e_i^2 \tag{4}$$

$$s.t.: y_i = \omega^T \phi(x_i) + b + e_i \quad i = 1, 2, \dots, N \tag{5}$$

Where  $e_i$  denotes as the errors,  $C$  is the penalty coefficient which is used to control the degree of punishment beyond the errors of the sample.

The optimization problem can be solved by introducing the Lagrange function:

$$L(\omega, b, e, \alpha) = J(\omega, e) - \sum_{i=1}^N \alpha_i [\omega^T \phi(x) + b + e_i - y_i] \tag{6}$$

Where  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_N)^T$  is the Lagrangian multiplier.

After solving the linear equation, the final expression is as follows:

$$f(x) = \sum_{i=1}^N \alpha_i k(x_i, x_j) + b \tag{7}$$

$k(x_i, x_j)$  is the inner product kernel function. In this paper, the radial basis function (RBF) kernel function is chose because it is the most effective one to deal with the nonlinear regression problems.

### 2.4 Simulated Annealing Algorithm

Simulated annealing (SA) is a kind of widely used stochastic optimization algorithm which is based on Monte Carlo iterative solution strategy. By simulating the annealing process of solid material in physics, problems similar to the NP complex can be solved.

Starting from the random feasible solution, for a given initial control parameters, SA keeps carrying out the iterative process of “generate data processing—judgment—accept/abandonment”. That is to say when the control parameter of temperature  $T$  stays the same, the relatively optimal solution of the combinatorial optimization can be obtained by repeating the Metropolis algorithm. And then reduce the control value of  $T$ , repeat the

Metropolis algorithm under different  $T$ , when  $T$  approaches zero, the final overall optimal solution of the combinatorial optimization problem can be acquired. Four steps of SA are as follows:

1. Initialization: choose the initial temperature and the length of Markov chain which stands for the number of the iteration of Metropolis algorithm.
2. Generate new state: reduce the control temperature according to its attenuation function. Every time the temperature reduces, there exists a random disturbance which generate a new state.
3. Generate new solution: determine whether to accept the new generated state as the new solution according to the accept function.
4. Obtain the optimal solution: the optimal solution will be procured until meeting the stop criterion.

### 3 The Proposed Method

This paper proposed a new combined predict model based on several methods of power load forecasting.

The details of the combined model are shown as follows:

Step 1. Import the training data. Then build the input set and the output set of the training set and the input set and the output set of the testing set.

Step 2. Build Elman model according to the input set and the output set. Then set the scope of the number of the intermediate layer nodes and to find the optimum value by circulation. Finally, use the training set to train ElmanNN model and get the forecast result set  $\hat{Y}_1$ .

Step 3. Build the GRNN model according to the input set and the output set. Then choose the smooth factor parameter by using the simulated annealing algorithm. When the simulated annealing algorithm was used, set the parameter of the simulated annealing algorithm. Finally, use the training set to train GRNN model and get the forecast result set  $\hat{Y}_2$ .

Step 4. Build LSSVM model according to the input set and the output set. Then use the training set to train GRNN model and get the forecast result set  $\hat{Y}_3$ .

Step 5. Build combined forecast model. Get the linear weighted combination of the above three models.

$$\hat{Y} = \alpha_1 \hat{Y}_1 + \alpha_2 \hat{Y}_2 + \alpha_3 \hat{Y}_3 \quad (8)$$

Where,  $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$  are weighted coefficients.

Then, use the simulated annealing algorithm to optimize the weighted coefficients. When the simulated annealing algorithm was used, set the parameter of the simulated annealing algorithm.

Step 6. Testing the trained combined model. Apply the combined model got by step 5 to forecasting.

Step 7. Output the predicted results and calculate the accuracy.

## 4 Simulation

The main purpose of this section was put forward the simulation of the proposed combined model.

### 4.1 The Performance Indexes

In this simulation, three generally adopted error indexes, the mean square predict error (MSE), the mean absolute error (MAE) and the mean absolute percent error (MAPE), was used given as follows:

$$MSE = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2 \quad (9)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - \hat{x}_i| \quad (10)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{x_i - \hat{x}_i}{x_i} \right| \times 100\% \quad (11)$$

Where,  $x_i$  is the real data of samples;  $\hat{x}_i$  is the predictive values of samples, and  $n$  is the number of samples.

### 4.2 The Process of Simulation

Original data used in this paper come from Queensland(QLD) state of Australia's electricity market, which gathered once every half hour from 0:30 on October 16, 2014 to 0:00 on October 1, 2015, a total of 16800 data items. 16464 data items gathered from 0:30 on October 16, 2014 to 0:00 on September 24, 2015 as the training set, and the remaining 336 data as the test set. The training data of the model was divided into seven groups, and every group were established data combined model. Combined Model is that using the data (48 items) form the previous week Monday to forecast the following Monday's data (48 items), Tuesday forecast Tuesday and soon.

Firstly, Seven ElmanNN models consisting of an inputting layer, an outputting layer, a hidden layer and a recurrent layer was adopted to forecast the whole week electrical load data. Set the number of input layer nodes and output layer nodes of each ElmanNN model were all 48. To achieve better prediction results, the node number of the hidden layer was set from 10 to 20. In this simulation, depending on the MAPE value, the number of hidden layer nodes with the minimum value of MAPE were selected. The optimal value of intermediate layer nodes of each ElmanNN model are shown in the Table 1. We can easily observe that the optimal value of middle layer nodes were 15 for the Sunday and Monday model, while that were 16 for Saturday and Tuesday model. For Wednesday, Thursday and Friday model that were 11, 17 and 12, respectively.

**Table 1.** The optimal value of intermediate layer nodes of each ElmanNN model.

Time	Thursday	Friday	Saturday	Sunday	Monday	Tuesday	Wednesday
BestHidnum	17	12	16	15	15	16	11
MinMAPE(%)	2.3326	3.1158	1.6596	1.7408	2.1731	1.274	1.8265

Secondly, seven GRNN models consisting of an inputting layer, a pattern layer, a summation layer and an outputting layer was adopted to forecast the whole week electrical load data. Set the input layer nodes was 48 and the number of output layer nodes was 48. The simulated annealing algorithm with the default parameter and 30 annealing chain length was used to find the optimal smoothing parameter.

Thirdly, seven LSSVM models consisting of 48 inputs, 48 outputs and RBF kernel function was adopted to forecast the whole week electrical load data.

Finally, the above three models were combined by linear weighted combination. Combination forecasting model that obtain the combination forecast model in the form of the appropriate weighted average can improve the accuracy of the prediction and reliability. The most concern of combined forecast model is how to calculate the weighted average coefficient to make the combination forecast model more effectively and to improve the prediction precision. Therefore, the simulated annealing algorithm was adopted to optimize the weights. First of all, the initial value of three combination weighted parameters were set to 0.33, 0.33 and 0.33. Then the simulated annealing algorithm with the default parameter and 30 annealing chain length was adopted to optimize the combination weights. The optimized combination weights by the simulated annealing algorithm are given in Table 2, in which  $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$  are the combination weights of ElmanNN model, GRNN model and LSSVM model, respectively. It can be clearly seen that on Saturday and Sunday the second parameter  $\alpha_2$  are 0.9956 and 0.9958, respectively, that point out that the results of those two days' data set largely depend on the GRNN model. While, the weights of Friday and Monday shows that the LSSVM model with the weights  $\alpha_3$  of 0.9956 and 0.9858 is extremely important for the combined model, and, the weights  $\alpha_1$  of Thursday shows that the ElmanNN model is a very significantly role. In addition, for the Tuesday and Wednesday, the combined model is affected by the joint ElmanNN model and GRNN model.

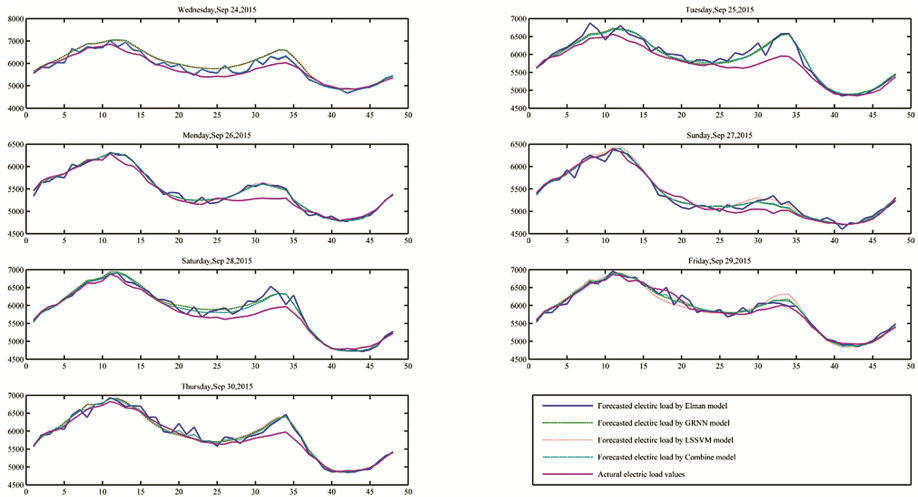
**Table 2.** The optimized combination weights by the simulated annealing algorithm

Time	Thursday	Friday	Saturday	Sunday	Monday	Tuesday	Wednesday
$\alpha_1$	0.9026	0.0022	0.0432	0.0703	0.0123	0.2950	0.3812
$\alpha_2$	0.0414	0.0022	0.8208	0.8688	0.0018	0.6612	0.5877
$\alpha_3$	0.0560	0.9956	0.1359	0.0609	0.9858	0.0438	0.0311



### 4.3 The Results and Analysis of Simulation

Figure 3 shows the actual values and forecasting values in the whole week of the four methods (Elman, GRNN, LSSVM and the combined model). As seen from Fig. 3, the curve of the combined model is more consistent with the actual data. Although at time of 12:00 to 18:00, the forecasting errors are bigger, it can still reveal that the combined model is better than the other three individual models.



**Fig. 3.** Final predicted values for each day by the four methods.

Table 3 presents the three indicators of the four forecasting methods in the form of numbers. It can be seen clearly that the average values of the combined model of the whole week are the lowest. The value of MAPE of the combined model is as low as 1.72%. MAE, MSE and MAPE of the combined model has the lowest value of all the four methods. That means the proposed combined model has the best forecasting performance. When comparing with the other three individual methods, GRNN has the largest forecasting values on Monday and Thursday, LSSVM has the largest values on Tuesday. While, Elman owns the largest forecasting data on Wednesday and Friday. The results display visually that different forecasting methods have different forecasting values which are sometimes meet the demands or sometimes unsatisfactory. Table 3 proved the high forecasting performance of the proposed combined model in another visual angle.

**Table 3.** Three statistics measures of the four forecasting methods.

Date	MSE (MW) of				MAE(MW) of				MAPE(%) of			
	Elman	GRNN	LSSVM	Combined	Elman	GRNN	LSSVM	Combined	Elman	GRNN	LSSVM	Combined
Thursday	33260.02	85072.50	76671.26	34457.04	134.57	236.91	222.62	137.50	2.33 %	4.04 %	3.81 %	2.38 %
Friday	63008.35	54911.33	49380.66	49394.17	182.83	169.15	161.86	161.90	3.12 %	2.90 %	2.78 %	2.78 %
Saturday	16255.90	12735.84	14882.34	13019.24	89.40	75.63	77.12	75.64	1.66 %	1.41 %	1.43 %	1.41 %
Sunday	14606.58	5421.87	11633.84	5934.85	91.24	53.67	76.32	56.613	1.74 %	1.03 %	1.46 %	1.09 %
Monday	36706.55	31984.93	21479.82	21567.15	125.56	138.75	107.17	107.34	2.17 %	2.37 %	1.84 %	1.85 %
Tuesday	9893.81	7118.90	18808.03	6548.34	75.815	65.29	100.14	62.18	1.27 %	1.09 %	1.67 %	1.04 %
Wednesday	22928.92	20501.05	23259.04	18394.86	108.29	94.71	102.01	88.59	1.83 %	1.59 %	1.71 %	1.50 %
Whole week	28094.3	31106.63	30873.57	21330.81	115.39	119.17	121.03	98.54	2.02 %	2.06 %	2.10 %	1.72 %

## 5 Conclusions

Interfered by various of factors, time series data has plenty of complex characteristics. Considering that an individual method cannot deal with all kinds of data, a novel combined model for STLF is presented in this paper. The proposed model combined with generalized regression neural network (GRNN), elman and least squares support vector machine (LSSVM). Optimizing by simulated annealing (SA), each individual method is assigned a weight coefficient. Multiplying all the three forecasting results by corresponding weights coefficients, the final forecasting results can be attained. In order to verify the performance of the combined model, electric load data from Queensland of Australia is utilized. The result shows that the average MAPE of the combined model is 1.72 % which is lower than the existing hybrid model named MFES proposed by Zhao et al. [30]. It had reduced MAPE of MFES by 29.04 %. The results of comparison demonstrate the excellent performance of the combined model. The reasons why the combined model has higher forecasting accuracy are listed below. Firstly, the proposed model combines two kinds of artificial neural networks which have strong forecasting performance. Between them, GRNN is optimized by SA, thus makes it more accurate. Secondly, instead of traditional average allocation method, SA is employed to determine weight coefficients of each individual model. Taking advantage of each model, the higher forecasting accuracy can be obtained. In a nutshell, the combined model outperforms other individual models. With higher accuracy, it is a promising tool in the future.

**Acknowledgments.** The authors would like to thank the Natural Science Foundation of PR of China (61073193,61300230), the Key Science and Technology Foundation of Gansu Province (1102FKDA010), the Natural Science Foundation of Gansu Province (1107RJZA188), and the Science and Technology Support Program of Gansu Province (1104GKCA037) for supporting this research.

## References

1. Masa-Bote, D., Castillo-Cagigal, M., et al.: Improving photovoltaics grid integration through short time forecasting and self-consumption. *Appl. Energy* **125**, 103–113 (2014)
2. Hong, W.C.: Electric load forecasting by seasonal recurrent SVR (support vector regression) with chaotic artificial bee colony algorithm. *Energy* **36**, 5568–5578 (2011)
3. Li, S., Wang, P., Goel, L.: Short-term load forecasting by wavelet transform and evolutionary extreme learning machine. *Electr. Power Syst Res.* **122**, 96–103 (2015)
4. Deihimi, A., Showkati, H.: Application of echo state networks in short-term electric load forecasting. *Energy* **39**, 327–340 (2012)
5. Zhang, R., Dong, Z.Y., Xu, Y., Meng, K., Wong, K.P.: Short-term load forecasting of Australian national electricity market by an ensemble model of extreme learning machine. *Gener. Transm. Distrib. IET* **7**, 391–397 (2013)
6. Kandil, N., Wamkeue, R., Saad, M., Georges, S.: An efficient approach for short term load forecasting using artificial neural networks. *Int. J. Electr. Power Energy Syst.* **28**(8), 525–530 (2006)
7. Jin, M., Zhou, X., Zhang, Z.M., Tentzeris, M.M.: Short-term power load forecasting using grey correlation contest modeling. *Expert Syst. Appl.* **39**, 773–779 (2012)

8. Moghram, I., Rahman, S.: Analysis and evaluation of five short-term load forecasting techniques. *IEEE Trans. Power Syst.* **4**(4), 1484–1494 (1989)
9. Amina, M., Kodogiannis, V.S., Petrounias, I., Tomtsis, D.: A hybrid intelligent approach for the prediction of electricity consumption. *Int. J. Electr. Power Energy Syst.* **43**(1), 99–108 (2012)
10. Xiao, L., Wang, J., Yang, X., Xiao, L.: A hybrid model based on data preprocessing for electrical power forecasting. *Electr. Power Energy Syst.* **64**, 311–327 (2015)
11. Liu, N., Tang, Q., Zhang, J.H., Fan, W., Liu, J.: A hybrid forecasting model with parameter optimization for short-term load forecasting of micro-grid. *Appl. Energy* **129**, 336–345 (2014)
12. Wang, J.Z., Zhu, S.L., Zhang, W.Y., Lu, H.Y.: Combined modeling for electric load forecasting with adaptive partial swarm optimization. *Energy* **35**(4), 1671–1678 (2010)
13. Xiao, Y., Liu, J.J., et al.: A neuro-fuzzy combination model based on singular spectrum analysis for air transport demand forecasting. *J. Air Transp. Manage.* **39**, 1–11 (2014)
14. Souptick, C., Sanjay, G., Dilip, K.P.: A combined neural network and genetic algorithm based approach for optimally designed femoral implant having improved primary stability. *Appl. Soft Comput.* **38**, 296–307 (2016)
15. Xia, C.H., Lei, B.J., Wang, H.P., Li, J.N.: GRNN short-term load forecasting model and virtual instrument design. *Energy Procedia* **13**, 9150–9158 (2011)
16. Chelgani, S.C., Jorjani, E.: Microwave irradiation pretreatment and peroxyacetic acid desulfurization of coal and application of GRNN simultaneous predictor. *Fuel* **90**(11), 3156–3163 (2011)
17. Li, H.Z., Guo, S., Li, C.J., Sun, J.Q.: A hybrid annual power load forecasting model based on generalized neural network with fruit fly optimization algorithm. *Knowl. Based Syst.* **37**, 378–387 (2013)
18. Kelo, S., Dudul, S.: A wavelet Elman neural network for short-term load prediction under the influence of temperature. *Electr. Power Energy Syst.* **43**, 1063–1071 (2012)
19. Song, Q.: On the weight convergence of Elman networks. *IEEE Trans. Neural Netw.* **21**(3), 96–101 (2010)
20. Li, X., Chen, G., Chen, Z., Yuan, Z.: Chaotifying linear Elman networks. *IEEE Trans. Neural Netw.* **13**(5), 1193–1199 (2002)
21. Shayeghi, H., Ghasemi, A.: Day-ahead electricity prices forecasting by a modified CGSA technique and hybrid WT in LSSVM based scheme. *Energy Convers. Manage.* **74**, 482–491 (2013)
22. Zhang, M., Fu, L.: Unbiased least squares support vector machine with polynomial kernel. In: 8th IEEE International Conference on Signal Processing (ICSP-2006), vol. 3, Guilin, China, pp. 16–20 (2006)
23. Xie, L., Zheng, H., Zhang, L.Z.: Electricity price forecasting by clustering-LSSVM. In: Proceedings of the International Power, Engineering Conference, pp. 697–702 (2007)
24. Hong, W.C.: Traffic flow forecasting by seasonal SVR with chaotic simulated annealing algorithm. *Neurocomputing* **74**, 2096–2107 (2011)
25. Lv, P., Yuan, L., Zhang, J.: Cloud theory-based simulated annealing algorithm and application. *Eng. Appl. Artif. Intell.* **22**, 742–749 (2009)
26. Pai, P.F., Hong, W.C.: Support vector machines with simulated annealing algorithms in electricity load forecasting. *Energy Convers. Manage.* **46**(17), 2669–2688 (2005)
27. Chandra, R.: Competition and collaboration in cooperative coevolution of Elman recurrent neural networks for time-series prediction. *IEEE Trans. Neural Netw. Learn. Syst.* **26**, 3123–3136 (2015)
28. Specht, D.F.: A general regression neural network. *IEEE Trans. Neural Netw.* **2**(6), 568–576 (1991)

29. Suykens, J.A.K., Van Gestel, T., De Brabanter, J., et al.: Least Square Support Vector Machines. World Scientific, Singapore (2002)
30. An, N., Zhao, W., Wang, J., et al.: Using multi-output feedforward neural network with empirical mode decomposition based signal filtering for electricity demand forecasting. *Energy* **49**, 279–288 (2013)