

# The Electrical Load Forecasting Base on an Optimal Selection Method of Multiple Models in DSM

<http://dx.doi.org/10.3991/ijoe.v11i8.4882>

Guilin Zheng, Li Zhang\*  
Wuhan University, Wuhan, China

**Abstract**—Electrical load forecasting plays a key role in energy scheduling and planning. It is a challenge to predict electric load accurately due to the versatility of electrical loads and the vast number of users in DSM of low-voltage side. Most of electrical load forecasting research focused on single model prediction or combination model prediction, which cannot get the optimal performance for some cases. Therefore, how to gather maximum optimal information from various different models is a key point in load forecasting and analysis. In this paper, an optimal selection method of multiple models for electrical load forecasting is studied. This method overcomes the shortcoming of unitary model, such as the instability and poor accuracy in some cases. To evaluate the forecast performance, a practical case is studied based on the intelligent electricity management system, which is presented by Wuhan University. It can be seen that the prediction error of the forecasting models can be calculated automatically and final optimum model can be obtained by optimum seeking software platform.

**Index Terms**—Optimal selection method, Load forecasting, Demand Side Management (DSM), Low-voltage side

## I. INTRODUCTION

As we know, the user is a vital component of Demand Side Management (DSM), and it plays an important role in encouraging and facilitating the operation and management of power system. Electric energy management system is not only confined to the power transmission and distribution in high-voltage side, but also pays attention to the real-time monitoring and controlling of energy in low-voltage side. In DSM, the construction and management process of low-voltage side is much more important and high quality electrical load forecasting is urgent, because the predictive accuracy is closely related to the interests of system operation and users. Load forecasting is an important part of energy management system and power distribution system, and it is also the premise and foundation of electricity decision-making and control [1]. Many crucial decisions that have significant economic influence are made based on the forecasted electric load. Along with the increase of competition and regulation, and the changing of public utilities industry structure, public utilities industry are also thirst of electrical load forecasting techniques, in order to obtain higher accuracy. It is pointed out by Bunn and Farmer that a 1% increase in forecasting error of electricity demand signified a £10 million increase in operating costs [2]. An accurate prediction can help decision makers develop an optimal action program because electricity is hard to store. However,

the low-voltage side of DSM is affected by various unstable factors, such as technical progress, holidays, policy, law, population growth, the natural and social environment and so on [3]. Therefore, developing new forecasting methods and improving the accuracy become imperative, no matter for very short-term, short-term, midterm or long-term load forecasts.

Forecasting is an essential instrument in science research, and received much attention during the past three decades [4–6]. The main type of traditional forecasting models includes regression model [7–8], time series technology [9], multiple linear regression [10], autoregressive moving average model [11–12], the functional nonparametric model and the semi-functional partial linear model. Most researches were focused on the study of individual model, however, these models have some limitations and merits on their own. Recently, as the development of intelligence techniques, many new intelligent forecasting methods, such as expert systems [13], fuzzy inference [14], ANN (artificial neural networks) [15], the SVM (support vector machine), grey correlation contest modeling [16], etc., have been proposed for electrical load forecasting. Since the single forecasting model cannot obtain the optimal value in many cases, many researchers established the combining forecast models, which made up of several single models with different concept [17]. In this paper, an electrical load forecasting method in the context of intelligent electricity management system is studied, which overcomes the shortcomings of unitary model. By using relevant selection strategy, the best combination of single models is obtained to construct the combining forecast model, eventually improve the performance of forecasting sample. Besides, a load forecasting platform based on optimal selection method is established. Two different cases are analyzed to verify the precision of forecasting models.

## II. THE OPTIMAL SELECTION METHOD

### A. Model base

In order to realize the optimal selection method of multiple models, it is the key to build the model base which is a bridge connecting the single model or the combination model for electrical load forecasting. The model base maximizes the use of gathered information, and contains the variety of different information of single model for optimization. It means: (1) To forecast by selecting available single model. According to the evaluation indexes, such as standard deviation, goodness-of-fit, relevance and

relative error etc., to compare the prediction results of different single model respectively, the pros and cons of these models can be obtained. The model with the best performance will be used as the final optimization model. The advantage of this approach is that the optimal forecasting model will not miss, which generally is used for short-term load forecasting. (2) To forecast by combination model. The optimal combination forecast is obtained through selecting appropriate weight for individual sub-models, and the optimal forecasting result is obtained by weighted average. It is more effective to consider comprehensive information, which provides strong stable and high accuracy. Meanwhile, it has good adaptability to the vagaries of prediction indexes and commonly advisable for medium and long-term load forecasting.

To make the optimal selection process of model base more clearly, a schematic diagram of implementation process is plotted as shown in Fig.1. According to different sample data, the final optimum model can be automatically obtained by the error calculation and error comparison of single models or combination models. In model base, if the number of the single model is  $n$ , then the number of combination model is  $C_n^2$ . In this paper, the single model refers to least squares support vector machine model (LS-SVM), back propagation neural network (BPNN) and gray model GM (1, 1), hence there are three variance covariance combination models (VACO).

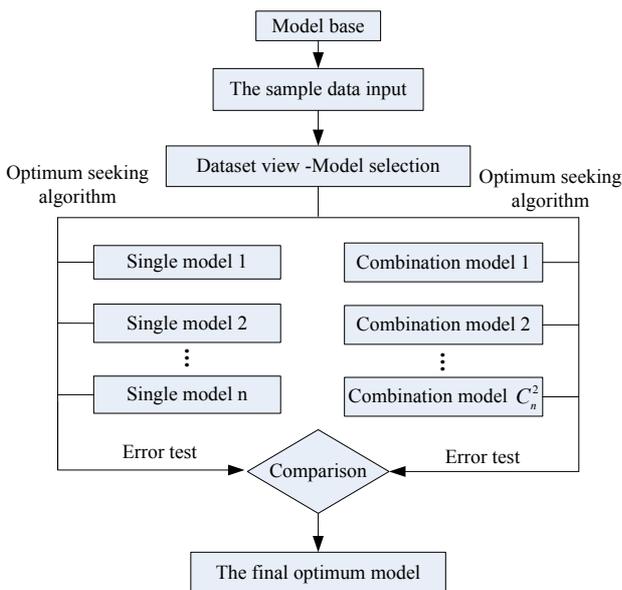


Figure 1. The schematic diagram of optimal selection process

To realize automatic optimum seeking function, the mixed programming form Visual C++ (VC) and MATLAB is adopted as a method of prediction algorithm updates without the need for debugging program in VC platform, the corresponding flow chart is shown in Fig.2. The method not only greatly reduces the workload of algorithm upgrade, but also improves the flexibility of the prediction model. The main interface of optimal selection software platform is designed as shown in Fig.3, the main functions include: (1) Historical data query; (2) Sample data extraction; (3) Invoking service of forecasting model; (4) Error evaluation; (5) Optimum seeking model. All the sample data has been run on a PC with 2GB of RAM and a 2.01-GHz-based processor, which can complete all cal-

culaton of prediction models by invoking various prediction algorithms from MATLAB software in the VC platform.

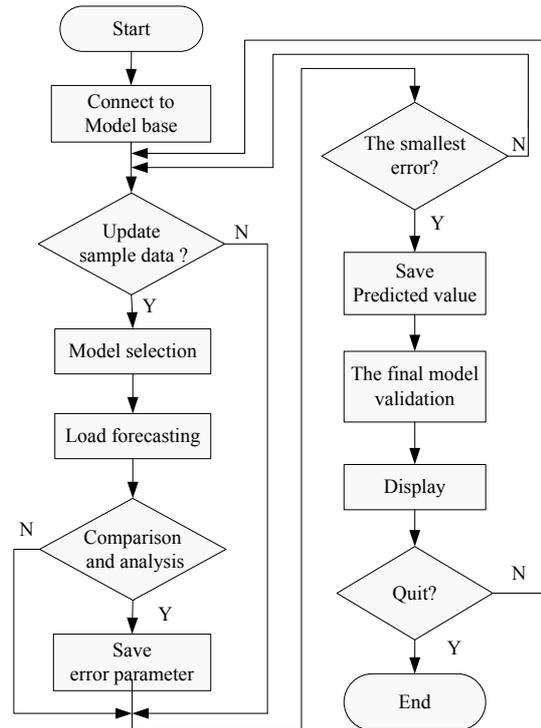


Figure 2. The flow chart of optimal selection algorithm

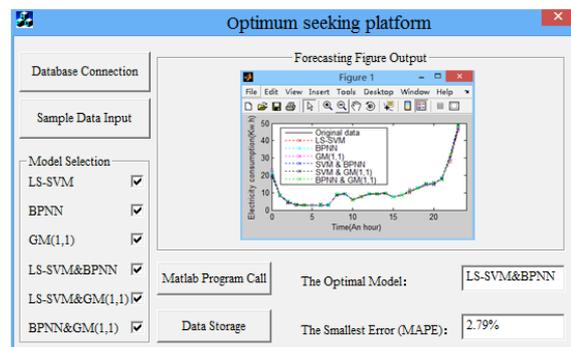


Figure 3. The main interface of optimal selection platform

## B. Review on prediction model

### 1) ILS-SVM model

LS-SVM allows the construction of an approximation by mapping the initial data set implicitly into some high dimensional feature space  $H$  through some mapping  $\phi(x)$ . This aims at constructing a simple linear approximation in the feature space  $H$  then corresponds to a nonlinear approximation in sample space. It has been developed for solving an approximation problem, and algorithm principle is as follows:

**Step 1:** Giving a training set

$$\{(x_k, y_k)\}_{k=1}^N, x_k \in R^n, y_k \in R, \text{ where } x_k \text{ is the } k\text{-th}$$

input sample,  $y_k$  is the corresponding desired output value.

**Step 2:** All can be done implicitly in  $H$  by means of the kernel function  $K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$ . Thus, LS-SVM model takes the form:

$$f(x) = w^T \phi(x) + b \quad (1)$$

Where,  $w \in R^k$  is weight vector and  $b \in R$  is deviation.

**Step 3:** In order to get  $w$  and  $b$ , the structural risk minimization principle constrained optimization problem must be solved:

$$\min_{w, b, \xi} \frac{1}{2} \|w\|^2 + \frac{\lambda}{2} \sum_{k=1}^N \xi_k^2 \quad (2)$$

$$s.t. \quad y_k = w^T \phi(x_k) + b + \xi_k, k = 1, 2, \dots, N$$

Where,  $\xi_k$  is the error variable,  $\lambda$  is the tradeoff parameter between a smoother solution and a smaller training error.

**Step 4:** To solve the optimization problem in formula 2, the Lagrange Function is introduced as follows:

$$L(w, b, \xi, \alpha) = \frac{1}{2} \|w\|^2 + \frac{\lambda}{2} \sum_{k=1}^N \xi_k^2 - \sum_{k=1}^N \alpha_k [w^T \phi(x_k) + b + \xi_k - y_k] \quad (3)$$

Where,  $\alpha_k$  is the k-th Lagrange multiplier. The  $\alpha$  and  $b$  can be obtained by formula 4.

$$\begin{bmatrix} 0 & \mathbf{e}^T \\ \mathbf{e} & \Omega + \frac{\mathbf{I}}{\lambda} \end{bmatrix}_{N \times N} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \quad (4)$$

Where,  $\mathbf{e} = (1, 1, \dots, 1)^T_{1 \times N}$ ,  $\mathbf{I} = \begin{bmatrix} 1 & \mathbf{L} & 1 \\ \mathbf{M} & \mathbf{K} & \mathbf{M} \\ 1 & \mathbf{L} & 1 \end{bmatrix}_{N \times N}$  and

$\Omega = (k(x_i, x_j))_{N \times N}$ .

**Step 5:** Finally, the nonlinear approximation function of the training sets is obtained as follow:

$$f(x) = \sum_{k=1}^N \alpha_k K(x_k, x) + b \quad (5)$$

## 2) BPNN model

BP (Back Propagation neural network) model is applied widely in load forecasting field. It is characterized by self-learning, self-organizing, adaptive, and has the advantages of simple structure and strong operability, which can handle with nonlinear data and establish the corresponding model. The three-layer neural network model is used for forecasting of daily electricity consumption in this paper. Every layer has several nodes, and each node in the network is a neuron whose function is to calculate

the inner product of the input vector and weight vector by a nonlinear transfer function to obtain a scalar result.

Because there are many influence factors in electricity consumption, such as temperature, holidays and people type etc. Here temperature is mainly considered. Due to the dimensions of temperature and electricity consumption are different in the process of learning neural network which can affect learning effect, At first, data normalization is performed, and then start the training.

### a) Data normalization

**Step 1:** The historical data that comes from smart electricity management system can be normalized to  $[-1, 1]$  by the normalization function, which is expressed as:

$$S_t = \frac{S - \frac{1}{2}(S_{\max} - S_{\min})}{\frac{1}{2}(S_{\max} - S_{\min})} \quad (6)$$

**Step 2:** The historical data can be normalized to  $[0, 1]$  by the normalized function again, which can be written as:

$$S_t = \frac{(S - S_{\min})}{(S_{\max} - S_{\min})} \quad (7)$$

**Step 3:** The historical data can be normalized to  $[0.1, 0.9]$  by the normalized function again, which can be expressed as:

$$S_t = 0.8 \frac{(S - S_{\min})}{(S_{\max} - S_{\min})} + 0.1 \quad (8)$$

**Step 4:** The results obtained by BP model should be substituted in an anti-normalization function to get the final forecasting result, which is written as:

$$S = 1.25S_t - 0.1(S_{\max} - S_{\min}) + S_{\min} \quad (9)$$

### b) Training and learning

The three-layer neural network structure is chosen. It contains input layer, hidden layer and output layer. According to the data normalization result, the number of neurons in input layer, hidden layer and output layer are selected to be 4, 12, and 1 respectively. The output is the electricity consumption. Variable Set: Transfer Function are 'tansig' and 'purelin', Training Function is 'traingdx', Learning Function is 'learnqdm', learning speed is 'lr' (lr=0.05), target error is 'goal'(goal=0.01), the maximum training steps is 20000. The training error curve is shown in Fig. 4.

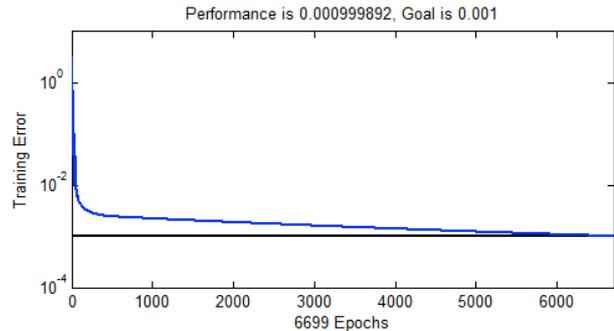


Figure 4. The training error curve comes from BPNN

3) GM (1, 1) model

Grey system theory is developed to study problems of "small samples and poor information". It looks for realistic patterns based on modeling a few available data and has no special requirements and restrictions on data sequence. Grey prediction is an important component of grey system theory, and is widely used for load forecasting. The basic modeling procedure is as follows:

**Step 1:** Giving a historical data sequence:  $x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$ , the new data sequence is generated by 1-AGO (Accumulated Generation Operation):

$$x^{(1)} = (x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)) \quad (10)$$

$$\text{Where, } x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i), k = 1, 2, \dots, n.$$

**Step 2:** Base on show exponential growth law of the sequence of  $x^{(1)}$ , the linear first-order differential equation is established as shown in formula 11.

$$\frac{dx^{(1)}}{dt} + \beta x^{(1)} = \mu \quad (11)$$

Where,  $\beta$  and  $\mu$  are undetermined coefficients.

**Step 3:** Solving  $\beta$  and  $\mu$  by least square method.

Set:

$$A = \begin{bmatrix} \beta \\ \mu \end{bmatrix} = (B^T B)^{-1} B^T Y_n \quad (12)$$

Where,  $B^T$  is data matrix,  $Y_n$  is data column, as shown in formula 13.

$$B = \begin{bmatrix} -\frac{1}{2} [x^{(1)}(1) + x^{(1)}(2)] & 1 \\ -\frac{1}{2} [x^{(1)}(2) + x^{(1)}(3)] & 1 \\ \dots & \dots \\ -\frac{1}{2} [x^{(1)}(n-1) + x^{(1)}(n)] & 1 \end{bmatrix}, Y_n = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \dots \\ x^{(0)}(n) \end{bmatrix} \quad (13)$$

**Step 4:** GM (1, 1) prediction model could be obtained as expression in formula 14.

$$x^{(0)}(k+1) = \left[ x^{(0)}(1) - \frac{\mu}{\beta} \right] e^{-\beta k} + \frac{\mu}{\beta}, \quad (k = 1, 2, \dots, n) \quad (14)$$

**Step 5:** GM (1, 1) prediction model of original sequence  $x^{(0)}$  can be generated by 1-IAGO (Inverse Accumulated Generating) as shown in formula 15.

$$x^{(0)}(k+1) = (1 - e^{-\beta}) \left[ x^{(0)}(1) - \frac{\mu}{\beta} \right] e^{-\beta k} + \frac{\mu}{\beta}, \quad (k = 1, 2, \dots, n) \quad (15)$$

**Step 6:** In accordance with formula 15 a new sequence of prediction data is obtained. That is:

$$x^{(0)}(k) = \begin{bmatrix} x^{(0)}(1) & x^{(0)}(2) & \dots & x^{(0)}(n) \end{bmatrix} \quad (16)$$

The residual test is written in formula 17.

$$e^{(0)}(k) = x^{(0)}(k) - \hat{x}^{(0)}(k) \quad (17)$$

According to formula 17, a set of residual error sequence is generated as follows:

$$e^{(0)}(k) = (e^{(0)}(1), e^{(0)}(2), \dots, e^{(0)}(n)) \quad (18)$$

The GM (1, 1) Model base on residual error correction is established in accordance with formula 18, it can be written in formula 19.

$$\hat{x}^{(0)}(k+1) = (1 - e^{-\beta}) \left[ e^{(0)}(1) - \frac{\mu}{\alpha} \right] e^{-\beta k} \quad (19)$$

According to formula 15 and formula 19, the final residual error correction gray prediction model is expressed in formula 20, and also is written in formula 21.

$$\hat{x}^{(0)}(k+1) = x^{(0)}(k+1) - x^{(0)}(k) + e^{(0)}(k+1) \quad (20)$$

$$\hat{x}^{(0)}(k+1) = (1 - e^{-\beta}) \left[ x^{(0)}(1) - \frac{\mu}{\beta} \right] e^{-\beta k} + \delta(k-i)(1 - e^{-\beta}) \left[ e^{(0)}(1) - \frac{\mu}{\alpha} \right] e^{-\beta k} \quad (21)$$

In formula 21,  $\delta(k-i) = \begin{cases} 1, & k \geq i \\ 0, & k < i \end{cases}, i = n - k, k = 1, 2, \dots, n.$

4) VACO combination model

The combination forecasting model improves prediction accuracy and out-of-sample performance, at the same time, overcomes the limitation of the single model through a strategy of selecting the single best forecasting model and appropriate combination of weights. The variance covariance method (VACO) [18] is used for weights determination of different single models, the in which combination forecasting model takes the form:

$$F = \psi_1 f_1 + \psi_2 f_2 \quad (22)$$

Where  $f_1, f_2$  represents two unbiased prediction value of single forecasting model transforming the actual data to its forecasting data;  $\psi_1, \psi_2$  is corresponding weight coefficient and  $\psi_1 + \psi_2 = 1$ .

Assuming prediction error of  $f_1, f_2$  and  $F$  is  $e_1, e_2$  and  $e_F$  respectively,  $F$  is also unbiased and its error is:

$$e_F = \psi_1 e_1 + \psi_2 e_2 \quad (23)$$

Setting  $Var(e_1) = \sigma_{11}, Var(e_2) = \sigma_{22}, cov(e_1, e_2) = \sigma_{12}$  the combination forecasting model with less error receives more weight and vice versa, that is:

$$\psi_1 = \frac{\sigma_{22} - \sigma_{12}}{\sigma_{11} + \sigma_{22} - 2\sigma_{12}}, \psi_2 = \frac{\sigma_{11} - \sigma_{12}}{\sigma_{11} + \sigma_{22} - 2\sigma_{12}} \quad (24)$$

Due to  $e_1$  and  $e_2$  mutual independence,  $\sigma_{12}=0$ , the formula 24 can be given by:

$$\psi_1 = \frac{\sigma_{22}}{\sigma_{11} + \sigma_{22}}, \psi_2 = \frac{\sigma_{11}}{\sigma_{11} + \sigma_{22}} \quad (25)$$

### III. THE CASE STUDY

The analyzed sample data is assembled on the “Cloud” server in Information Center from management system of student dormitory of Wuhan University. There are three buildings in the Information Engineering District, named IED NO.2, IED NO.5, and IED NO.6 .Two buildings are undergraduate apartments in Medical Science District namely MSD NO. 1 and MSD.NO. 2 respectively. The data under study for this paper is collected via the IoT DSM system from all 5 buildings in a time domain from Sep.1, 2014 to Jan.31, 2015.The total electricity consumption status of the five buildings in five different months are shown in Fig.5.

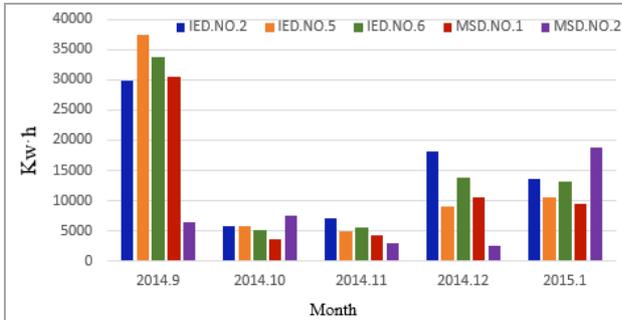


Figure 5. The total electricity consumption status form Sep.1, 2014 to Jan.31, 2015

#### A. Case data description

The effectiveness of load forecasting is determined by the precision of size. There is no kind of prediction method which can ensure that the predicted result contains no error in any scenario. To verify the feasibility of optimal selection method of multiple models and to test the performance of the optimal selection software, hourly electricity consumption data per day of IED NO.5 from Dec.15, 2014 to Jan.15, 2015 is used as original forecasting sample sets with its change trend curve shown in Fig. 6 To be sure, this choice is random .Two different case data are designed for model base in this section, they are

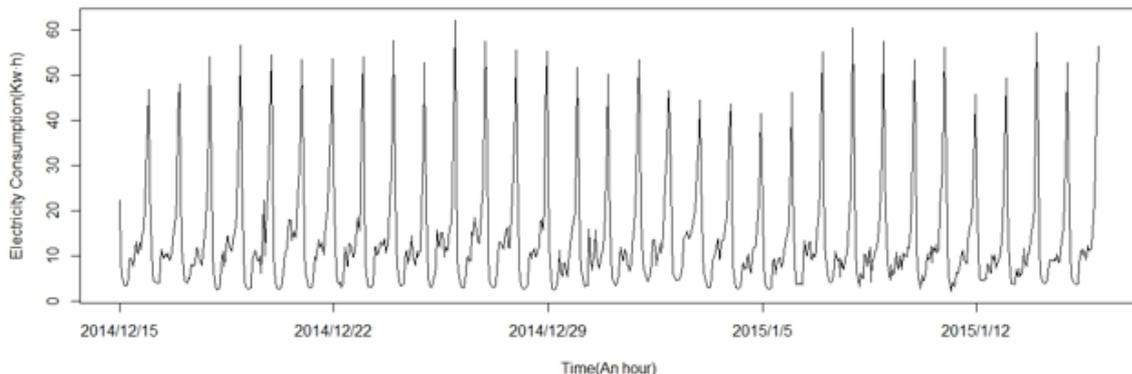


Figure 6. The original forecasting sample sets of IED NO.5 From Dec.15, 2014 to Jan.15, 2015

24 hour daily electricity consumption data and the maximum daily consumption data respectively. In IED NO.5, there are 173 rooms in which each room contains 4 persons. The fixed electrical equipment includes an air conditioner (Peak-power: 1.5KW) and 4 daylight lamps (Total power: 120W), while other electrical equipment are uncertain. For the same model and method, different training data and different checking data may lead to different forecasting results. Here we focus on the following case data.

#### Case1: The 24 hour daily electricity consumption forecasting

The 24 hour daily electricity consumption forecasting can be realized by comparison between single models and combined models based on optimal selection method. In this case, the model training data comes from original forecasting sample set (Dec.15, 2014 to Jan.4, 2015). Due to electricity usage habits differences of the students during workday and rest day, the model checking data comes from two kinds of working time, such as workday data (Jan.5, 2015, Monday, 24hour) and rest day data (Jan.10, 2015,Saturday, 24hour) are shown in Table 1, which can also be used for comparing load forecasting curve in next section.

#### Case 2: The maximum daily electricity consumption forecasting

The maximum daily electricity consumption forecasting can be done with the same method as case 1. The training data is the maximum daily consumption of original forecasting sample sets from Dec.15, 2014 to Jan. 3, 2015. The maximum daily electricity consumption data from Jan.4, 2015 to Jan.13, 2015 is chosen to test the fabricating model as shown in Table II.

#### B. Forecasting results

The electricity consumption data is randomly selected for simulation and error analysis in order to test the performance of forecasting models. In optimal selection software platform, six forecasting models are selected and the corresponding prediction errors are calculated automatically as shown in Fig.3, at the same time, the model which has the smallest error is automatically identified as the final optimum model and its result is taken as the final result. The corresponding forecasting results of 24 hour daily electricity consumption on workday (Jan.5, 2015, Monday) and rest day (Jan.10, 2015, Saturday, 24hour) are shown in fig.7and fig.8. The corresponding forecasting results of the maximum daily electricity consumption are shown in fig.9. Specifically, the prediction error analysis is presented in next section.

TABLE I. THE MODEL CHECKING DATA FOR CASE 1

Jan. 5, 2015 (Temperature : 3°C-13°C)				Jan. 10, 2015 (Temperature : 0°C-14°C)			
Time	Electricity consumption (Kw·h)	Time	Electricity consumption (Kw·h)	Time	Electricity consumption (Kw·h)	Time	Electricity consumption (Kw·h)
0:00	20.41	12:00	9.35	0:00	34.97	12:00	11.97
1:00	8.56	13:00	9.46	1:00	11.4	13:00	8.55
2:00	4.42	14:00	9.53	2:00	4.76	14:00	11.89
3:00	2.87	15:00	7.47	3:00	4.48	15:00	11.04
4:00	2.74	16:00	8.58	4:00	2.75	16:00	12.53
5:00	2.72	17:00	10.88	5:00	5.67	17:00	10.55
6:00	2.72	18:00	12.39	6:00	4.41	18:00	11.53
7:00	2.87	19:00	14.56	7:00	4.56	19:00	20.38
8:00	8.58	20:00	15.12	8:00	7.57	20:00	23.35
9:00	9.41	21:00	17.86	9:00	10.46	21:00	29.58
10:00	5.82	22:00	30.58	10:00	9.13	22:00	45.04
11:00	7.68	23:00	46.12	11:00	9.48	23:00	56.15

TABLE II. THE MODEL TRAINING AND CHECKING DATA FOR CASE 2

Time	Temperature	Maximum daily consumption (Kw·h)	Time	Temperature	Maximum daily consumption (Kw·h)
2014/12/15	-1°C~10°C	46.81	2014/12/30	2°C~16°C	50.11
2014/12/16	-4°C~9°C	48.01	2014/12/31	-3°C~10°C	53.22
2014/12/17	-3°C~11°C	53.98	2015/1/01	-3°C~9°C	46.57
2014/12/18	-3°C~11°C	56.52	2015/1/02	3°C~11°C	44.36
2014/12/19	0°C~13°C	54.28	2015/1/03	2°C~14°C	43.55
2014/12/20	-1°C~12°C	53.24	2015/1/04	2°C~16°C	41.35
2014/12/21	-4°C~8°C	53.59	2015/1/05	3°C~13°C	46.12
2014/12/22	-1°C~10°C	53.91	2015/1/06	-1°C~6°C	55.05
2014/12/23	0°C~11°C	57.65	2015/1/07	-3°C~9°C	60.42
2014/12/24	-1°C~13°C	52.7	2015/1/08	-1°C~12°C	57.41
2014/12/25	1°C~10°C	62.12	2015/1/09	2°C~12°C	53.34
2014/12/26	2°C~7°C	57.36	2015/1/10	0°C~14°C	56.15
2014/12/27	2°C~8°C	55.33	2015/1/11	-°C~15°C	45.71
2014/12/28	0°C~13°C	55.31	2015/1/12	4°C~11°C	49.37
2014/12/29	1°C~16°C	51.60	2015/1/13	0°C~11°C	59.16

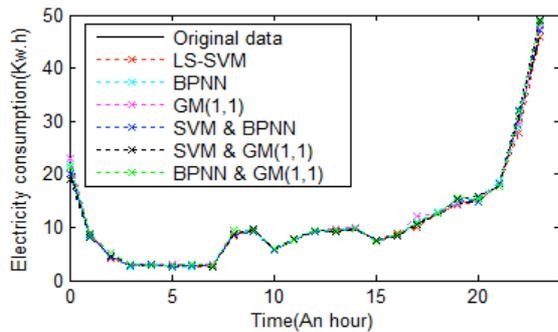


Figure 7. Prediction results of 24 hour daily electricity consumption on workday

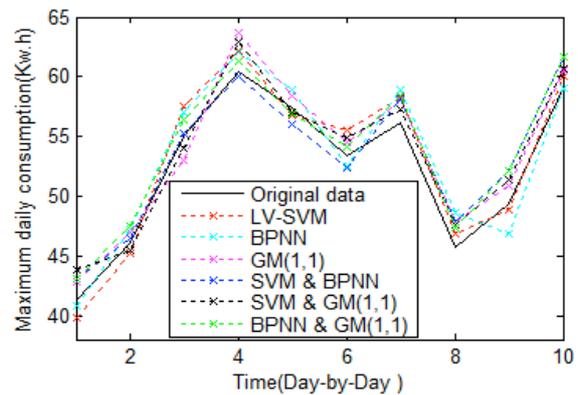


Figure 9. The prediction results of the maximum daily electricity consumption

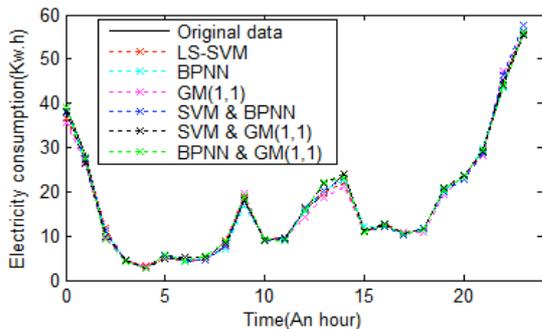


Figure 8. Prediction results of 24 hour daily electricity consumption on rest day

#### IV. ANALYSIS AND DISCUSSION

##### A. Evaluation index

Load forecasting is an estimation of the future load, thus inevitably there is deviation between actual value and predicted value, which is defined as prediction error. The evaluation and analysis of prediction error not only can provide prediction accuracy which is an important reference for decision-making, but also is useful for improving the load forecasting result. In order to test

the forecasting accuracy, the actual value and the predicted value are compared. In both cases, three evaluation indexes are employed to investigate the forecasting ability of the proposed model, which are defined in equations (26)-(28).

1) Mean absolute error

$$MAE = \frac{\sum_{i=1}^n |A_i - P_i|}{n} \quad (26)$$

2) Mean absolute percentage error

$$MAPE = \frac{\sum_{i=1}^n \left| \frac{A_i - P_i}{A_i} \right|}{n} * 100\% \quad (27)$$

3) Root mean square error

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (A_i - P_i)^2}{n}} \quad (28)$$

Where,  $A_i$  is the  $i$ -th actual value,  $P_i$  is the  $i$ -th predicted value, and  $n$  is the total number of data for prediction and comparison.

*B. The error analysis*

In this paper, according to the error analysis of each single model, the combination model can be constructed by selecting standard error of any two single models as the weight coefficient. There are three single models, which means are three combined models in model base. The three evaluation indexes for the total six models are calculated. The comprehensive comparative analysis of MAE, MAPE and RMSE for both case are shown in Table 3, Table 4 and Table5.

Take MAPE as an example to analyze the forecasting results .As we know, the electrical load on campus is similar to some extend due to regular timetable. For example, the student’ campus life and school activities have some kind of similarity on workday in a month statistically, which lead to original sample data of the electricity

consumption form all workdays in a month with some kind of similarity too.

When the different checking data are selected for the 24 hour daily electricity consumption testing, the results can be seen as follow: 1) In Table 3, the maximum MAPE value is 4.1046% from GM (1, 1) model and the minimum value is 2.7906% from combination model (LS&BPNN), but the MAPE value of single model (such as LS-SVM, BPNN) is lower than combination model (such as LS-SVM&GM (1, 1),BPNN&GM(1,1).The combination model (LS&BPNN) can be obtained as the final optimum model by the optimal selection software. 2) In Table 4,the changing rule of the prediction error is similar to Table 3.The maximum

MAPE value is 4.521 % from GM (1, 1) model and the minimum value is 2.2850% from combination model (LS&BPNN), the final optimal model is still combination model (LS&BPNN).However, the respective average value of three error indexes on workday is lower than on rest day owing to the students' life habit is more similar from Monday to Friday.

When the training data and checking data are updated, taking case 2 data as example, the prediction error results can be seen from the Table 5.The maximum MAPE value is 3.2322% from combination mode (BPNN&GM (1, 1)), the minimum value is 2.7073% from single model (LS-SVM) which can be displayed as the final optimal model. These show that the selecting of training samples and testing samples is a critical for performance evaluation.

As is known through the above analysis, a certain individual forecasting model or combination model is not always best in all situations, so the optimal selection method of multiple models for load forecasting is presented in this paper. This method overcomes the limitation of unitary model, which might present instability and unacceptably low forecasting accuracy in certain situation. Empirical study is given to put forward advisable plans for optimizing load adjusting meanwhile, it can provides reliable guidance for power grid operation and power construction planning, which is also important for the sustainable development of electric power industry.

TABLE III. THE ERROR INDEXES FORM SIX FORECASTING MODELS FOR 24 HOUR DAILY ELECTRICITY CONSUMPTION ON WORKDAY

Index	LS-SVM	BPNN	GM(1,1)	LS-SVM&BPNN	LS-SVM&GM(1,1)	BPNN&GM(1,1)
MAE	0.39132625	0.41916585	0.428065188	0.298994583	0.431678754	0.421941875
MAPE (%)	0.036642243	0.03491538	0.041046206	0.027906598	0.038569667	0.037583823
RMSE	0.223179664	0.26430916	0.264247881	0.089723479	0.280450592	0.259899133

TABLE IV. THE ERROR INDEXES FORM SIX FORECASTING MODELS FOR 24 HOUR DAILY ELECTRICITY CONSUMPTION ON REST DAY

Index	LS-SVM	BPNN	GM(1,1)	LS-SVM&BPNN	LS-SVM&GM(1,1)	BPNN&GM(1,1)
MAE	0.489420167	0.4822792	0.880272917	0.503593746	0.498472188	0.465676017
MAPE (%)	0.038364005	0.03503992	0.0452068	0.028500562	0.041000411	0.039414166
RMSE	0.310655277	0.28577982	0.791305197	0.200471844	0.270808022	0.292054561

TABLE V. THE ERROR INDEXES FORM SIX FORECASTING MODELS FOR THE MAXIMUM DAILY ELECTRICITY CONSUMPTION

Index	LS-SVM	BPNN	GM(1,1)	LS-SVM&BPNN	LS-SVM&GM(1,1)	BPNN&GM(1,1)
MAE	1.414468	1.583492	1.6674501	1.4716814	1.509426	1.6029194
MAPE (%)	0.027073008	0.03064844	0.03184107	0.028822663	0.029651225	0.032322679
RMSE	1.226094863	1.64964054	1.622956296	1.472803126	1.369414616	1.540708457

## V. CONCLUSIONS

Providing high-accuracy electrical load forecasts enables the reduction of risk, thereby improving the economic social benefits of Demand Side Management (DSM), which reduces generation cost, improves the security of the power system and helps administrators to develop an optimal action program. Although over a thousand scientific papers address the topic of load forecasting every year, only a few are dedicated to find a general model for load forecasting that improves the performance in different cases. In this paper, an optimal selection method for electrical load forecasting is studied on the low-voltage side of DSM. Owing to the same model and method, different training data and different checking data may lead to different forecasting results, the sample data of 24 hour daily electricity consumption and the maximum daily electricity consumption are analyzed which come from intelligent electricity management system in Wuhan University. Prediction accuracy comparison and analysis of six models are also presented. The forecasting results show that the optimal selection method overcomes the limitation of unitary model, which reduces instability and eliminates unacceptably low forecasting accuracy in some cases. In addition, this method can be utilized in other forecasting fields, such as product sales forecasting, tourism demand forecasting, early warning and flood forecasting, wind speed forecasting, and traffic flow forecasting.

## REFERENCES

- [1] Önköl D, Sayim K Z, and Lawrence M, "Wisdom of group forecasts: Does role-playing play a role?" *Omega*, Vol. 40, Issue 6, pp. 693–702, Dec. 2012. <http://dx.doi.org/10.1016/j.omega.2011.01.010>
- [2] Bunn DW, Farmer ED, "Comparative models for electrical load forecasting", New York: *John Wiley and Sons*, 1985, Chap.4.
- [3] Wang Y, Wang J, Zhao G, et al., "Application of residual modification approach in seasonal ARIMA for electricity demand forecasting: a case study of China", *Energy Policy*, Vol.48, pp.284-294, Sep. 2012 <http://dx.doi.org/10.1016/j.enpol.2012.05.026>
- [4] Donate J P, Cortez P, Sánchez G G, et al., "Time series forecasting using a weighted cross-validation evolutionary artificial neural network ensemble", *Neurocomputing*, Vol. 109, No. 3, pp. 27-32, Jun. 2007.
- [5] R. Palivonaite, M.Ragulskis, "Short-term time series algebraic forecasting with internal smoothing", *Neurocomputing*, Vol. 127, No. 15, pp. 161-171, Mar. 2014. <http://dx.doi.org/10.1016/j.neucom.2013.08.025>
- [6] J. Shao, "Application of an artificial neural network to improve short-term road ice forecasts", *Expert Syst. Appl.*, Vol. 14, Issue 4, pp. 471-482, Apr. 1998. [http://dx.doi.org/10.1016/S0957-4174\(98\)00006-2](http://dx.doi.org/10.1016/S0957-4174(98)00006-2)
- [7] Sorjamaa A, Hao J, Reyhani N, et al., "Methodology for long-term prediction of time series", *Neurocomputing*, Vol. 70, Issue 16-18, pp. 2861-2869, Oct. 2007. <http://dx.doi.org/10.1016/j.neucom.2006.06.015>
- [8] Al-Hamadi H M, Soliman S A, "Long-term/mid-term electric load forecasting based on short-term correlation and annual growth", *Electr. Power Syst. Res.*, Vol. 74, Issue 3, pp. 353-361, Jun. 2005. <http://dx.doi.org/10.1016/j.epsr.2004.10.015>
- [9] Dong R, Pedrycz W, "A granular time series approach to long-term forecasting and trend forecasting", *Physica A*, Vol. 387, Issue 13, pp. 3253-3270, May. 2008.
- [10] Mohamed Z, Bodger P, "Forecasting electricity consumption in New Zealand using economic and demographic variables", *Energy*, Vol. 30, Issue 10, pp. 1833-1843, Jul. 2005. <http://dx.doi.org/10.1016/j.energy.2004.08.012>
- [11] Saab S, Badr E, Nasr G, "Univariate modeling and forecasting of energy consumption: the case of electricity in Lebanon", *Energy*, Vol. 26, Issue 1, pp. 1-14, Jan. 2001. [http://dx.doi.org/10.1016/S0360-5442\(00\)00049-9](http://dx.doi.org/10.1016/S0360-5442(00)00049-9)
- [12] Wang B, Tai N, Zhai H, et al., "A new ARMAX model based on evolutionary algorithm and particle swarm optimization for short-term load forecasting", *Electr. Power Syst. Res.*, Vol. 78, Issue 10, pp. 1679-1685, Oct. 2008. <http://dx.doi.org/10.1016/j.epsr.2008.02.009>
- [13] Rahman S, Hazim O, "Load forecasting for multiple sites: development of an expert system-based technique", *Electr. Power Syst. Res.*, Vol. 39, Issue 3, pp. 161-169, Dec. 1996. [http://dx.doi.org/10.1016/S0378-7796\(96\)01114-5](http://dx.doi.org/10.1016/S0378-7796(96)01114-5)
- [14] Mamlook R, Badran O, Abdulhadi E, "A fuzzy inference model for short-term load forecasting", *Energy Policy*, Vol. 37, Issue 4, pp. 1239-1248, Apr. 2009. <http://dx.doi.org/10.1016/j.enpol.2008.10.051>
- [15] Hsu C C, Chen C Y, "Regional load forecasting in Taiwan-applications of artificial neural networks", *Energ. Convers. Manage.*, Vol.44, Issue 12, pp. 1941-1949, Jul. 2003. [http://dx.doi.org/10.1016/S0196-8904\(02\)00225-X](http://dx.doi.org/10.1016/S0196-8904(02)00225-X)
- [16] Jin M, Zhou X, Zhang Z M, et al., "Short-term power load forecasting using grey correlation contest modeling", *Expert Syst. Appl.*, Vol.39, Issue 1, pp. 773-779, Jan. 2012. <http://dx.doi.org/10.1016/j.eswa.2011.07.072>
- [17] Lopez J A, Diebold F X, "Forecast evaluation and combination", New York: *National Bureau of Economic Research*, 1996, Chap.3.
- [18] Niu D, Cao S H, Zhao L, et al., "Electric power load forecasting technology and its application", Beijing: *China Electric Power Press*, 2009, Chap.2.

## AUTHORS

**G. L. Zheng** is with the Department of Automation, School of Power and Mechanical Engineering, Wuhan University, Wuhan 430072, China (e-mail: glzheng@whu.edu.cn).

**L. Zhang** is with the Department of Automation, School of Power and Mechanical Engineering, Wuhan University, Wuhan 430072, China (Corresponding author: Phone:+86-027-68773985; Fax:+86-027-68775892; zhangli00121@163.com).

This work was supported in part by National Natural Science Foundation of China under Grant No. 61272114, Marine Renewable Energy Special Fund Project of the State Oceanic Administration of China under Grant No.GHME2013JS01. Submitted 21 July 2015. Published as re-submitted by the authors 10 October 2015.