The Application of Support Vector Machine in Load Forecasting

Wenqing Zhao
School of Business Administration, North China Electric Power University, Beijing, China
Email: jbzwq@126.com

Fei Wang
School of Business Administration, NCEPU, Beijing, China
Email: laishengyuan_wf@126.com

Dongxiao Niu
School of Business Administration, NCEPU, Beijing, China
Email: niud@126.com

Abstract **The forecasting to mid-long term load is important because it can provide important evidence to the power planning. Traditional forecast techniques apply a single forecaster to carry out the task. However, this forecaster might not be the best for all situations or databases. A combinational model on the basis of Support Vector Machine (SVM) theory is proposed in this paper. During the process of the forecast, several single forecasting methods such as trend prediction model, exponent model, non-linear regression model, improved grey predictive model and improved grey verhulst predictive model, are used to form a model group, and then the fitted results by different traditional predictive models in time sequence act as the input of the support vector machine regression (SVMR) model, then by relative SVMR approach based on known input and output samples, we can obtain the test model. In the paper, the procedure of the combinational prediction on transformer faults based on SVMR is discussed in details. The example on load data has proven that the proposed model can give good results on both the fitting to the known data in time sequence and the extrapolation to the data to be predicted. Moreover, compared with other predictive approaches, both single model and other combinational model, the proposed combinational forecasting model has higher prediction accuracy.**

Index Terms **Support Vector Machine, Mid-long term load forecasting, Combinational forecasting**

I. INTRODUCTION

The forecasting to mid-long term load can provide important evidence to the power planning. The accurate load forecasting can improve the economics and reliability of power system operation. The load is related with a variety of factors, there is a complex nonlinear relationship between the load and the factors.

Traditional forecasting methods, such as linear regression, time series method and the elastic coefficient method [1], establish mathematical expression directly to describe the relationship between relevant factors and load. New forecasting methods, such as neural networks [2] and improved fuzzy clustering algorithm [3], model and analyze to the sequences of load data. Each of which has remarkable superiority and unavoidable weakness. Traditional forecast techniques apply a single forecaster to carry out the task. However, this forecaster might not be the best for all situations or databases. In order to utilize the superiority of various forecast effectively, J. N. Bates and C. W. J. Granger proposed the theory of combining prediction [9]. Because it can improve the prediction accuracy effectively, it has been a focus issue in predict research at home and abroad and is used widely. The basic principle of combining prediction is that utilize the information provided by various forecasting methods and get the combination forecasting model by the way of appropriate weighted average. The theory points out that, as to a forecast method with the worse performance, if it has independent system information and is combined with the others approach with high performance, the combinational prediction system outperforms each its constituent forecasters consequently. In the essence, if these different forecasters are complementary, the combinational of their decisions will lead to improved accuracy. At present, there are many different combining methods, such as the methods based on rough sets [4], neural network [5, 6], associated optimization [7], mutual information [8], etc.

As a new machine learning algorithm, support vector machine (SVM) can solve some practical issues, such as small sample, nonlinear, high dimension and local minimum points, etc. SVM has been used in the research

Project supported by National Natural Foundation of China (70671039, 61074078) and National Natural Foundation of Hebei Province (E2009001392) and the Fundamental Research Funds for the Central Universities (09QG33)

doi:10.4304/jcp.7.7.1615-1622

of power load prediction widely [10-13]. A new load forecasting method has been proposed in document [10]. The document [11] improved support vector machine prediction model based on artificial immune algorithm. The document [12] do the decompose processing to the load data by using wavelet transform, then get the load at each time of the day through the support vector machine with different characteristics, and finally superimposed all of the load to get the final forecasts. The document [13] did the research of short-term load forecasting based on wavelet transform and support vector machine.

Compared with the short-term electric load data, the long-term data has the characteristics of small samples. SVM has the advantages to the small samples that other models can not be compared with, and the SVM regression method has good capability of fitting and extrapolation. In this paper, we propose a combinational forecasting model mainly based on SVM and several single forecasting methods such as trend prediction model, exponent model, non-linear regression model, improved grey predictive model and improved grey verhulst predictive model to carry out data forecasting.

II. The Possibility of SVM to Implement the Combination Forecast

A. Support Vector Machine

Support Vector Machine (SVM) is a class of supervised learning algorithms first introduced by Vapnik[13]. To the nonlinear regression problem, SVM do nonlinear transformation by defining the appropriate kernel function to transform the input space into a high dimensional space, and then in the new space use the linear regression to find a support vector which is the nearest to the optimal classification plane. And then construct the optimal classification plane based on the training sample on the hyper-plane that parallel to the optimal classification plane. By this way, it can increase the generalization ability to the test samples. The generalization ability of SVM is much better than neural network model and fuzzy model. SVM include the classification SVM to solve classification problems and the regression SVM for function approximation. At present, SVM has been extended to solve nonlinear regression estimation problems in image processing [15], data analysis [16], classification [17], fault diagnosis [18], etc.

In practice, in order to construct a nonlinear regression support vector machine model, we need to determine the learning parameter C (Where C is the punishment parameter, which is considered to specify the trade-off between the empirical risk and the model flatness, and its default value is 1), g (Where g is the parameter in the kernel function, and its default value is $1/k$, here k is the number of input data's properties), and the parameter P (Where P is to set ε in the loss function of v-SVR). In this work, radial basis function (RBF) [19] $K(x_i, x_j) = \exp(-g \parallel x_i - x_j \parallel^2)$ is used in the SVM, so that we need to determine the parameter g.

B. Thoughts of Combination Forecasting and The Possibility of Using SVM to Achieve It

The combination forecasting method was proposed by Bates and Granger[9], its essence is to utilize the single model of information and obtain the combination forecasting model through the appropriate weighted form. The various forecasting methods proposed by the scholars at home and abroad use the minimum largest absolute error as the optimal combination of criteria to calculate the weight coefficient vector prediction in the practical applications and theoretical study. Document [20] established a optimal combination forecasting model which had the minimum square error, use the nature of the forecasting absolute error information matrix to determine the non-inferior combination forecasting and conditions. Document [21] further studied the model of the sum of squared errors of non-negative constraints and optimal combination forecast. Document [22] proposed a prediction based on the effective degree of combination forecasting model, given the linear planning method and at the same time studied the quality of the combination forecasting model based on the forecasting effectiveness. Document [23] proposed a correlation index based on the optimal combination forecasting model.

The model proposed above is mostly based on a certain kind of optimization criteria to find the right of a fixed weight, in this paper points out that the weight combination forecasting should change with the forecast steps, that is it shout be the variable weight. Long-term electric load data has the characteristics of small samples and the support vector machine is suitable for the modeling of small samples and poor information, and the fitting result and generalization ability are good. Linear combination as a variety convex combination of individual prediction method, try not to miss the original data information as the guiding principle, is a optimization of a single prediction.

But this is only within the convex optimization, limited to the prediction accuracy. In contrast, non-linear overcomes the drawbacks, it is more to predicted value of the individual forecasts as a guide, the non-linear machine learning, which are in order to achieve the purpose of extrapolation. The principle of nonlinear combining prediction based on SVM regression is that the results of m forecasters $f_i(i = 1,2,...m)$ are used as the input vectors of a support vector machine (SVM) and the corresponding actual values used as the output of the SVM. Using the predictive value to do the nonlinear machine learning effectively, and thus construct a nonlinear relationship between the predictive value and the actual value $\hat{y} = f(f_1, f_2, \ldots f_m)$. In some measure, the measure of $\hat{y}$ is the superior than the $f_i(i = 1,2,...m)$, here f is nonlinear function. This combination function has a better predictive performance to the non-linear object.

Based on the characteristics of combining prediction and long-term power load, it is reasonable to construct the combinational forecasting for mid-long term load based on Support Vector Machine.

III. Combinational prediction based on SVM

A. Five Single Forecasting Methods

Under certain conditions, there is a clear trend in the change of power load, once found the trend we can follow it to judge the future of the load, this is the principle of the trend extrapolation forecast. Through the research of predictive models, we found that when there is fewer modeling data, exponent model can strengthen the trend of the original sequence significantly and reduce the influence of the randomness of the original sequence on the results[24], it also has the features that simple and smaller calculate amount. In addition, in many practical problems, the relationship between Independent and dependent variables is non-linear, so that non-linear regression model can solve such problems effectively. There also have many factors to impact the load, such as weather, regional economic activities, etc. these factors are difficult to know exactly. The changes of the load are gray and uncertainty.

So that, in this paper, trend prediction model, exponent model, non-linear regression model, improved grey predictive model and improved grey verhulst predictive model are used to build the first prediction model group.

① Trend prediction model carries out fitting curve using $y = a_0 + a_1 x + a_2 x^2 + ... a_n x^n$, $a_0, a_1, ... a_n$ are obtained using cumulative prediction method[25].

② Exponent model builds forecasting model using function $y = be^{ax}$. To $y = be^{ax}$, both sides are adopted by natural logarithm, then a and b are obtained using polynomial fitting[19].

③ Non-linear regression model[25] uses function $\ln y = \ln a + bx$, which is the equivalent transformation form of the function $y = ae^{bx}$, then a and b are obtained using linear regression.

④To strong random original data sequences, improved Grey theory (GM (1,1))[24] model carries out data smoothing using the first-order exponent flatness operation, $S(t) = \alpha Y(t) + (1-\alpha) S(t-1), \alpha \in [0,1]$ and the background value of the predictive model is reconstructed using parameter β [24] to obtain the minimized error rate.

⑤Given an original data sequence $X^{(0)}$, $X^{(1)}$ is 1-AGO (accumulated generating operation) sequence of $X^{(0)}$, and $Z^{(1)}$ is the average value sequence of $X^{(1)}$, then $X^{(0)} + aZ^{(1)} = b(Z^{(1)})^2$ is called Verhulst model. Verhulst model carries out back ground optimization using back ground parameter β [24].

B. Combinational Prediction Model Based on SVM

J. N. Bates and C. W. J. Granger proposed the theory of combinational forecasting. Suppose there are n(≥2) prediction forecasters, y_t is the actual value, and t=1,2, ⋯ m. The nonlinear combinational principle of combined prediction is to construct a certain function relationship among all various forecast methods, namely,

$$\hat{y}_t = \sum_{i=1}^{i=n} w_{t,} f_{it} \tag{1}$$

Where $\hat{y}_t$ denotes the combined prediction value at time t, w_i is the weight value of model i, i=1,2,…n , $\sum_{i=1}^{i=n} w_i = 1$, and f_{it} is the prediction value of forecaster i at time t, i=1,2,…n, t=1,2,…m.

The input sample order is the fifth-order, and the format is $[Y_i, 1{:}Y_{Gi}, 2{:}Y_{Vi}, 3{:}Y_{Di}, 4{:}Y_{Zi}, 5{:}Y_{Fi}]$. Where Y_i is the actual value at ith time which is the output of the combinational model based on SVM, $[Y_{Gi}, Y_{Vi}, Y_{Di}, Y_{Zi}, Y_{Fi}]$ are the values predicted by improved Grey theory (GM (1,1)) model, improved Grey Verhulst model, trend prediction model, exponent model, non-linear regression model at ith time, respectively. These predictive results of five single models constitute input of the combinational model based on SVM and the corresponding actual values used as the output of the SVM. Then, the time series data is utilized to train the prediction model based on SVM. Thus, the nonlinear function relationship between the prediction value and the actual value is modeled. The combinational prediction model is shown in Figure 1.

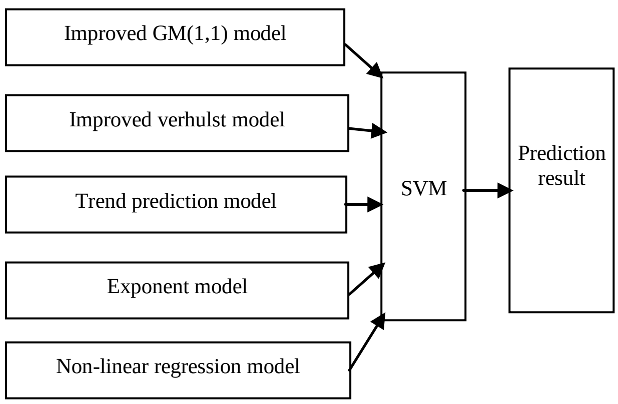

Figure 1. The combinational prediction model based on SVM

C. Error Standards

A mean absolute percentage error (MAPE) is used to evaluate the forecasting accuracy, which is as follows:

$$MAPE_{Tr} = \frac{1}{n_{Tr}} \sum_{i=1}^{n_{Tr}} \frac{\left| \hat{y}_{Tri} - y_{Tri} \right|}{y_{Tri}} \times 100\%$$

$$MAPE_{Te} = \frac{1}{n_{Te}} \sum_{i=1}^{n_{Te}} \frac{\left| \hat{y}_{Tei} - y_{Tei} \right|}{y_{Tei}} \times 100\%$$

$$MAPE = MAPE_{Tr} + MAPE_{Te} \tag{2}$$

Where $\hat{y}T_{ri}$, yT_{ri} and nT_r represent the forecasting, actual values and sample number of training set, respectively, $MAPE_{Tr}$ represents fitting error. $\hat{y}T_{ei}$, yT_{ei} and nT_e represent the forecasting, actual values and sample number of testing set, respectively, $MAPE_{Te}$ represents extrapolated error.

D. Steps of Prediction Model Based on SVM

Step 1: Training set (TR) and testing data set (TE).

The real values of a time series are divided into training set (TR) and testing data set (TE).

Step 2: Data processing.

With the raw input data set (time series T), first thing we need to do is to carry out data processing. For example, when the data value has a wide range, the normalization strategies should be built to speed up the convergence, and after normalization, the data values are between -1 and 1.

Step 3: Prediction model construction and data fitting.

1) Five single prediction models are constructed using yT_{ri} of training set (TR) and parameters of models are obtained. That is, $a_0, a_1, \ldots a_n$ which is parameter of trend prediction model are obtained, a and b which is parameter of exponent model are obtained, a and b which is parameter of non-linear regression model, α and β which is parameter of improved GM(1,1) model are obtained, a , b and β which is parameter of improved Grey Verhulst model are obtained.

2) The parameters obtained in 1) are input in five single prediction models and fitting values $\hat{y}_{Tri}$ of y_{Tri} are obtained. These fitting values $\hat{y}_{Tri}$ are training samples of SVM.

3) The SVM forecasting model is constructed using training samples obtained in 2), then the parameters of SVM (C, g and p) are obtained.

4) y_{Tri} is fitting using the SVM forecasting model in 3) and the second fitting values of time series T are obtained. At the same time, the fitting error $MAPE_{Tr}$ can be obtained through (2).

Step 4: predicting data value of testing data set (TE).

Prediction is carried out using five single models in step 3 and five prediction values (at time point t) $[Y_{Gi}, Y_{Vi}, Y_{Di}, Y_{Zi}, Y_{Fi}]$ are obtained. Then these five prediction values are inputted as SVM model and the prediction value at time point t is obtained. At the same time, the error $MAPE_{Te}$ and $MAPE$ can be obtained through (2) in following.

IV. Experiments and Evaluation

A. Experiment 1

In order to verify the effectiveness of the proposed model in this paper, we use the electric load data of certain area in recent years to do some experiments. The original data is shown in Table I.

TABLE I.
THE ELECTRIC LOAD IN A CERTAIN REGION

date	load
2000	12351
2001	13087
2002	13823
2003	13692
2004	14764
2005	16706
2006	20340
2007	22167

We select the data of 2000 to 2006 for the training set and data of 2007 for the testing set. At first, we use these data of certain area in recent years to do some experiments to verify the effectiveness of the improved proposed grey models in this paper. The parameters of the improved GM (1, 1) model and the improved verhulst model are that a=0.1006,b=10250.779 and a=0.0561,b=8.7204E-06 respectively. And the parameter of the traditional GM (1,1) model is that a=0.0993,b=10068.916 and the verhulst model's is that a=0.04726, b=8.4948035E-06.

In the following tables, we use ARE, AARE, PGM, PVerhulst, TGM and TVerhulst to represent the absolute relative error, the average absolute relative error, the improved GM(1,1) model, the improved verhulst model, the traditional GM(1,1) model and the verhulst model respectively. The result is shown in Table II and Table III.

TABLE II.
THE RESULTS OF THE GREY MODELS

Real value		12351	13087	13823	13692	14764	16706	20340	22167
PGM	Fitting value	12351	12090	13386	14803	16370	18104	20021	22140
	ARE(%)	0	7.618	3.161	8.114	10.878	8.368	1.568	0.122
PVerhulst	Fitting value	12351	13043	13865	14856	16070	17590	19547	22152
	ARE(%)	0	0.336	0.304	8.501	8.846	5.292	3.899	0.068
TGM	Fitting value	12351	11875	13114	14483	15994	17664	19507	21543
	ARE(%)	0	9.261	5.129	5.777	8.331	5.734	4.095	2.815
TVerhulst	Fitting value	12351	13126	14050	15170	16554	18303	20585	23678
	ARE(%)	0	0.298	1.642	10.795	12.124	8.725	1.205	6.816

TABLE III.
THE COMPARISON BETWEEN THE MODELS' RESULT

The prediction model	AARE (%)
TGM	5.143
TVerhulst	5.305
PGM	4.979
PVerhulst	3.406

From Table II and Table III, one can easily conclude that the average absolute relative error of the traditional models is 5.143% and 5.305% respectively, and the precision of the proposed models is only 4.979% and 3.406% respectively. Obviously, precision of improved grey models are higher than the traditional models. And we also can conclude that, to the non-monotonic and swing sequence, the improved verhulst model has higher accuracy.

And then using these data to predict based on the proposed model in this paper. The forecasting steps are as follows. The first seven actual data are used to be original sequence, and the actual data of 2007 is used to make a comparison with predictive value. So $X^{(0)} =$ (12351, 13087, 13823, 13692, 14764, 16706,20340).

The prediction results are obtained by improved Grey model, improved Grey Verhulst model, trend prediction model, exponentiation model, and non-linear regression model , respectively, and they are listed as follows.

Y_G =(12351,12090,13386,14803,16370,18104,20021,22140)

Y_V =(12351,13043,13865,14856,16070,17590,19547,22152)

Y_T =(12670,12999,13874,15295,17262,19775,22834,26439)

Y_E =(11858,12759,13792,14772,15894,17102,18402,19800)

Y_N =(11858,12785,13753,14794,15915,17120,18417,19811)

Moreover, parameters of these five models are shown in Table IV.

TABLE IV.
THE PARAMETERS OF FIVE SINGLE MODEL

model	parameter		
improved GM(1,1)	β =0.33	a=0.1006	u=10250.779
improved Verhulst	β =0.27	a=0.561	b=8.7204E-06
trend prediction model	a_0=12887	a_1=-490	a2=273
exponent model	a=0.0732	b=11021	
non-linear regression model	a=11047.9	b=0.073	

The training samples are obtained based on the above prediction results which are the results of the first forecasting model group, and the training samples are listed as follows.

$$\begin{bmatrix} 12351\ 1:12351\ 2:12351\ 3:12670\ 4:11858\ 5:11885 \\ 13087\ 1:12090\ 2:13043\ 3:12999\ 4:12759\ 5:12785 \\ 13823\ 1:13386\ 2:13865\ 3:13874\ 4:13729\ 5:13753 \\ 13692\ 1:14803\ 2:14856\ 3:15295\ 4:14772\ 5:14794 \\ 14764\ 1:16370\ 2:16070\ 3:17262\ 4:15894\ 5:15915 \\ 16706\ 1:18104\ 2:17590\ 3:19775\ 4:17102\ 5:17120 \\ 20340\ 1:20021\ 2:19547\ 3:22834\ 4:18402\ 5:18417 \\ 22167\ 1:22140\ 2:22152\ 3:26439\ 4:19800\ 5:19811 \end{bmatrix}$$

The second forecasting is carried out using SVM and the results is Y_{SVM} = (12351, 12606, 12971, 13692, 14898, 16706,19158,22112). The parameters (C， ε and

g) of SVM model are 65536, 0.0625 and 0.015625, respectively.

Moreover, we compared the proposed model with five single forecasters according error standards in section III.C, the results are given in Table V.

TABLE V.
THE COMPARISON BETWEEN THE MODELS' RESULT

model / data	Actual value	improved GM(1,1)	improved Verhulst	trend prediction	exponent model	non-linear regression	proposed model
2000	12351	12351	12351	12670	11858	11885	12351
2001	13087	12090	13043	12999	12759	12785	12606
2002	13823	13386	13865	13874	13729	13753	12971
2003	13692	14803	14856	15295	14772	14794	13692
2004	14764	16370	16070	17262	15894	15915	14898
2005	16706	18104	17590	19775	17102	17120	16706
2006	20340	20021	19547	22834	18402	18417	19158
2007	22167	22140	22152	26439	19800	19811	22112
$MAPE_{Tr}$		5.672	3.883	8.98	4.95	4.97	2.36
$MAPE_{Te}$		0.122	0.068	19.27	10.68	10.63	0.24
MAPE		6.794	3.951	28.25	15.63	15.60	2.60

From the results, it can be concluded that proposed model is generally better than the traditional algorithm. For $MAPE_{Tr}$, the proposed model is 2.36, which is the smallest in 6 models of Table V, and so does $MAPE$. Obviously, the proposed combinational model based on SVM is effective.

B. Experiment 2

Document [26] used the variable weight combination forecasting model to do predict. We used the proposed model in this paper to do predict with the same data. The result and the construct of the model are shown in Table VI . From the average absolute relative error of the two models shown in Table VI, one can easily conclude that the proposed model is generally better than the variable weight combination forecasting model in document [26].

TABLE VI.
THE COMPARISON BETWEEN THE MODELS' RESULT

Real value	Proposed model	Document model
	Fitting value	Fitting value
1015	1015.07	1036.93
1097	1097.07	1114.72
1209	1184.84	1204.06
1356	1355.94	1350.84
1490	1516.74	1495.23
1677	1676.94	1638.64
1838	1838.07	1815.34
1967	1966.94	1964.13
ARE（%）	0.478	1.072

C. Experiment 3

In order to verify the effectiveness of the proposed models, we can also use electric load data in document [27] to do some experiments. We select the data of 1986 to 1995 for the training set and to predict the load data of 1996 to 2000. The result and the construct to the other models are shown in Table VII. As shown in the table, we can get the conclusion that the proposed model is generally better than the variable weight combination forecasting models in document [27]. In Table VII, we used model 1 to model 3 to present the equal weight combination forecasting model, Variance - covariance optimal combination forecast model and grey combination forecasting model, respectively.

TABLE VII.
THE COMPARISON BETWEEN THE MODELS' RESULT

model / date	Real value	Model 1	Model 2	Model 3	Document model	Proposed model
1996	1968	1932.43	1950.80	1935.06	1948.75	1930
1997	2061	2079.14	2091.85	2080.82	2090.11	2064
1998	2130	2210.71	2207.69	2210.08	2208.13	2191
1999	2284	2349.29	2324.62	2345.26	2332.71	2323
2000	2617	2558.29	2536.84	2555.50	2549.62	2511
MAPE		2.316	2.172	2.285	2.153	2.139

V. CONCLUSION

This study employed several prediction approaches to construct model group, then carried out prediction based on SVM, and finally, a combinational prediction model was constructed. The experimental results show that the proposed model can reduce the error rate. Moreover, by comparing with the grey prediction algorithm, we conclude that our proposed model is better and more efficient than the traditional methods Moreover, compared with other predictive approaches, both single model and other combinational model, the proposed combinational forecast model has higher forecasting accuracy. Therefore, the problem of good fitting and bad extrapolation in traditional predictive approaches is solved to some extent.

In our future research, we will improve the performance of our model. And further studies can be performed to build a complete prediction system. Another issue we will discuss separately is how to choose the parameter when doing the prediction.

ACKNOWLEDGEMENT

The financial support of the National Natural Science Foundation of China (61074078), the National Natural Foundation of Hebei Province (E2009001392) and the Fundamental Research Funds for the Central Universities (09QG33) are gratefully acknowledged.

REFERENCES

[1] LI Yumei. The utility combination forecasting method in mid-long term load forecasting [D]. Sichuan University , 2006. (In Chinese)

[2] Zeng M, Liu B H, Xu Z Y, et al. Short term load forecasting based on artificial neural network and fuzzy theory[J].Journal of Hunan University :Natural Sciences ,2008 ,35(1):58-61.(In Chinese)

[3] Chen Rouyi, Zhang Yao, Wu Zhigang, et al. Application of improving fuzzy clustering algorithm to power load forecasting [J].Proceedings of the CSU EPSA, 2005,17(3):73- 77. (In Chinese)

[4] Zhong Bo, XIAO Z. A method of combination forecast based on rough set[J]. Statistic Research Journal, 2002(11):37 -39.(In Chinese)

[5] GU X H , XING M , NIU D X. Multifactor influenced combined grey neural network models for power load forecasting[J].East China Electric Power Journal ,2006 ,34(7) : 548 - 551.(In Chinese)

[6] Zhao Haiqing. The Application to Power Load Forecasting of Optimization Combinatorial Predication Model[J].OPERATIONS RESEARCH AND MANAGEMENT SCIENCE,2005,14(1):115-118. (In Chinese)

[7] YANG S J, NIU D X. Model of forecasting power load based on optimized combination of model library[J].Journal of North China Electric Power University ,2005 ,32 (1) :42 - 44.(In Chinese)

[8] Li Chunsheng, Wang Yaonan,Chen Guanghui. Combination Forecast Model Based on Mutual Information and Its Application to Power Load Prediction[J].Journal of Hunan University(Natural Sciences) ,35(9):58-61.(In Chinese)

[9] BATES J M, GRANGER C W J. Combination of forecasts [J] .Operations Research Quarterly, 1969 ,20 (4) :451 - 468.

[10] Zhai Yongjie, Wang Jingxian, Zhou Lihui. Power system mid-term load forecasting based on fuzzy support vector machines[J].Journal of North China Electric Power University,35(12):70-73. (In Chinese)

[11] Lu Zhigang, Zhou Ling, Yang Lijun, et al. Power load forecasting based on artificial immune algorithm weighted-SVM model[J]. Relay, 2005, 33(24):42-44. (In Chinese)

[12] Zhu Zhiyong, Lin Mugang, Zhang Shengji.Short-term load forecasting based on wavelet transform and support vector machine[J].Microcomputer Applications,2005, 26(4): 440-42. (In Chinese)

[13] Liu Mengliang, Liu Xiaohua, Gao Rong. Short Term Load Forecasting Using Wavelet Transform and SVM Based on Similar-Days[J]. Transactions Of China Electro technical Society,2006,21(11):59-64. (In Chinese)

[14] Vapnik V N, Statistical learning theory[M]. New York: Wiley, 1998.

[15] Hu Zhengping, Wu Yan, Zhang Ye. A novel fast support vector machine based on support vector geometry analysis[J]. Journal of Image and Graphics, 2007, 12(1):82-86.,(In Chinese).

[16] Jiao Licheng, Bo Liefeng, Wang Ling. Fast sparse approximation for least squares support vector machine. IEEE Transactions on Neural Networks, 2007, 18(3):685-690.

[17] Doumpos Michael, Zopounidis Constantin, Golfinopoulou Vassiliki. Additive support vector machines for pattern

classification. IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics, 2007, 37(3):540-550.
[18] ZHAO Wenqing, ZHU Yongli, ZHANG Xiaoqi. Combinational Forecast for Transformer Faults Based on Support Vector Machine[J]. Proceedings of the CSEE, 2008, 28(25):15-19. (In Chinese).
[19] Lin Maoliu, Chen Chunyu, A performance comparison of SVMs based onfourier kernel and RBF kernel [J], Journal of Chongqing University of Posts and Telecommunication, 2005, 17(6)"647-650(In Chinese).
[20] Tang Xiaowo. Research on error matrix of combination prediction [J]. Journal of University of Electronic Science and Technology of China, 1992, 21(4):448-454. (In Chinese)
[21] Ma Yongkai, Tang Xiaowo, Yang Guiyuan. A study on basic theory of the optimal combination prediction method of non-negative weights [J]. Operations research and management science, 1997, 6(2):1-8. (In Chinese)
[22] Chen Huayou, Hou Dingpi. Research on superior combination forecasting model based on forecasting effective measure[J]. Journal of University of Science and Technology of China, 2002, 32(2):172-180. (In Chinese)
[23] Wang Yingming. Research on the methods of combining forecasts based on correlativity. Forecasting, 2002, 21(2):58-62. (In Chinese)
[24] Wu Lizeng. Assessing approach of transformer condition[D]. Hebei: North China Electric Power University, 2005. (In Chinese)
[25] Niu Dongxiao, Cao Shuhua, et al. Power load forecasting technology and its application[M]. Beijing: China Electric Power Press, 2009. (In Chinese).
[26] Gu Jie. Study On The Varied Weight Synthesis Model of Mid-Long Term Load Forecasting In Power System Proceedings of the CSU-EPSA, 2003,15(16):58-59. (In Chinese)
[27] Zhou Quan,Ren Haijun, et al. Variable Weight Combination Method for Mid-long Term Power Load Forecasting Based on Hierarchical Structure .Proceedings of the CSEE, 2010,30(16):50-51. (In Chinese)

Wenqing Zhao, This author was born in Shanxi province, China. She received Doctor Degree of electric power system and its automation in 2009 from North China Electric Power University.

Wenqing acted as a visiting scholar in the Deakin University, at present she is a associate professor in North China Electric Power University. Her present research interests are artificial intelligence and data mining.

Fei Wang, This author was born in Hebei province, China. She is studying Master degree of computer application in North China Electric Power University.

Her present research interests are artificial intelligence.

Dongxiao Niu, This author was born in Anhui province, China. He received Doctor Degree of business management in 1999 from North China Electric Power University.

At present he is a professor in North China Electric Power University. His present research interests are load forecasting.