An Improved Short-Term Power Load Combined Forecasting With ARMA-GRACH-ANN- SVM Based on FHNN Similar-Day Clustering

Dongxiao Niu
School of Economics and Management, North China Electric Power University, Beijing, Chinia
niudx@126.com

Yanan Wei
School of Economics and Management, North China Electric Power University, Beijing, Chinia
weiyanan010@163.com

***Abstract*—In this paper, an efficient combined modeling based on FHNN similar-day clustering to forecast short-term power load is proposed. As the performance of individual models varies under different circumstances, the combination weights of forecast model should change with the circumstances. Here we classify historical power load into three parts including training set, validation set and test set model. Four methods, including Autoregressive Moving Average (ARMA), Generalized Autogressive Conditional Heteroscedasticity (GRACH), Artificial Neural Network (ANN) and Support Vector Machine (SVM), are selected as candidate models. For short load forecasting, the circumstance of the coming day is compared with those of past days and then clustered into the same category by Fuzzy Hopfield neural network (FHNN). The combining weights are obtained according to mean absolute percentage errors of different models. Then the combined forecasting model with ARMA-GRACH-ANN-SVM weighted by average with the weights obtained from FHNN clustering is got. A case study shows that the proposed combined model outperforms other forecast methods.**

***Index Terms*—short-term power load, combined forecasting, ARMA-GRACH-ANN-SVM, FHNN, similar days clustering**

I. INTRODUCTION

Electric power load forecasting is the foundation of planning and design and the assurance of operation efficiency and reliability of electric power system. Because of the inherent characteristics of uncertainty, randomness and nonlinear, the load forecast has always been a forefront and hot issue in academic research. The key of electric power load forecasting is prediction technology. Many of experts and scholars at home and abroad has been doing a lot of research and achieved much fruitful headway in prediction theory and method. Existing load forecasting methods can be divided into the traditional classical methods taking mathematic as theory basis and modern intelligent forecast methods [1]. The traditional forecasting methods mainly include: time series forecast method, regression forecast method, the trend extrapolation method; modern short-term prediction method: the artificial neural network, grey theory prediction method, fuzzy forecasting method, the wavelet analysis forecasting method, the support vector machine forecasting method.

Power load being obvious timeliness, Time series modeling exploits information contained in historical electricity prices to mine the linear or nonlinear map relations between the future historical loads. So Time series modeling is widely applied in load forecasting. Pappas s. Sp. put forward that autoregressive moving average (ARMA) method can make full use of power load forecast data information to improve the accuracy of prediction in [2]. Another time series modeling technique called the generalized autoregressive conditional heteroscedasticity (GARCH) model is put forward by Bollerslev in 1986. Time series modeling pays less attention to external influences, leading to undesirable forecasting of unstable prices.

Artificial neural networks (ANNs) and support vector machine (SVM) are the most popular modeling methods used in Artificial intelligence load forecasting. AI is designed to develop a curve-fitting tool to simulate the complex nonlinear relation between the electricity price and external factors. A new ANN model with a single output node structure was proposed in [3], showing that the ANN was more robust than time series in multi-step ahead forecasting.SVM and similar day method is used in [4], which shows a high forecast precision in short-term load forecast.

The traditional methods usually use single method to carry on the forecast with certain regression function to describe the prediction rules, which shows obvious shortcomings in complex condition factors. With the development of new intelligent technology, more advanced methods are continually put forward, the proposed combination forecasting (or integrated

Corresponding author: Yanan Wei. weiyanan010@163.com

doi:10.4304/jsw.8.3.716-723

forecasting) method has become the new direction of load forecasting.

Current combination forecast research has two main directions. One is to combine study optimization algorithm with traditional and single forecasting model [5], such as combined fuzzy clustering and neural network, the combination of genetic algorithm and neural network, which is usually used in short-term load forecasting. The other is to weighted array predicted results of more than one single model [6-7]. Comprehensively judge results, give different weight to each prediction model and thus gain a better integrated model. Based on more prediction results, according to certain optimal criteria, the combined model can improve the prediction precision of fitting results and achieve the purpose to describe a problem from various aspects. To make full use of information, the method optimal combines information of several single models. The difference of sorts of integrated model mainly lies in ways of weight coefficient, combination forecasting method of determining weight being a critical problem.

The traditional combined methods include: average weight combinations, optimal weight combinations and variance combination. The average weight of combination forecast method is relatively mature and commonly used which can reduce prediction risk of single forecasting method and improve prediction stability contrast optimal weight combinations and variance combination [8-9]. Without nonlinear combination of single forecasting models, it is relatively difficult and complicated to determine the weights of traditional combination model. Now some scholars proposed intelligent way of combination forecast, [10-11] put forward a combination forecast model based on neural network, using nonlinear fitting ability of neural network to combination forecast. Nonlinear fitting kinds of prediction results, without limit of weight coefficient range, intelligent combination forecast model thus greatly overcomes difficulty to determine weight coefficient [12]. But the traditional neural network itself has a certain limitations, such as structure complex, easily getting into local superior. Therefore, there is certain theoretical significance and practical value to seek for new intelligence combination methods.

In this paper, a novel combined forecasting is proposed to combine several load forecasts that have been validated. We classify the circumstances into several kinds by a cluster algorithm, FHNN, then analyze the forecasting abilities of different models under different circumstances, and finally use the information on the forecasting abilities to determine the weights for future use.

The rest of the paper is organized as follows. In Section 2, we briefly introduce FHNN, including the source and calculation basis, and the way to combine several forecast methods with FHNN. In Section 3, four signal power load forecasting methods are introduced including ARMA, GRACH, ANN and SVM. According to analysis actual power load, experiments using the proposed technique are described and compared with other four individual methods in Section 4. Finally, a discussion and conclusions are presented in Section 5.

II. COMBINED FORECAST BASED ON FHNN SIMILAR-DAY CLUSTERING

In this section, a brief introduction to FHNN and four signal power load forecasting methods is given, following which the approaches to combined modeling with FHNN similar-day clustering are presented.

A. Fundamentals of FHNN

The Hopfield neural network with simple architecture and parallel potential has been applied in many fields [13-14]. Fuzzy Hopfield neural network (FHNN) is raised to unsupervised clustering [15]. A topological structure of FHNN is shown as in Figure 1.

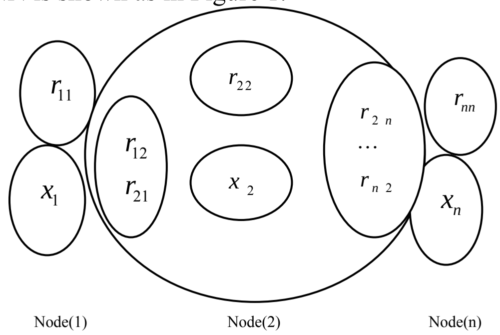

Figure 1. Topological structure of FHNN.

Where node(i), $1 \le i \le n$ represents the i[th] neuron of FHNN; r_{ij} represents the linking degree from node(i) to node(j).

The set of neurons in FHNN corresponds to the domain $X = \{x_1, x_2, \cdots, x_n\}$, and the fuzzy relation $R = (r_{ij})_{n\times n}$ is regarded as the weighting matrix between neurons in FHNN. Then FHNN is a mathematical model executing the computation of fuzzy logic.

A fuzzy Hopfield network FHNN with n order is defined as 5-tuple

$$FHNN = (NS, R, \wedge, O, Oper) \tag{1}$$

Where the set of neurons is $NS = \{node(i) | 1 \le i \le n\}$; the weighting matrix between neurons is $R = (r_{ij})_{n\times n}, r_{ij} \in (0,1]$; $\wedge = (\wedge_1, \wedge_2, \cdots, \wedge_n)^T, \wedge_i \in (0,1]$ for threshold vector; the output vector of neurons is $O = (o_1, o_2, \cdots o_n)^T, o_i \in (0,1]$, $O(t)$ is the state vector in time t of neurons in FHNN; A complete parallel computational mode is used in FHNN, where Oper

$$O(t+1) = g(R * O(t) - \wedge), i.e.,$$

$$o_j(t+1) = \mathrm{sgn}(\bigwedge_{i=1}^{r}(r_{ij} \vee o_i(t)) - \lambda_j) \tag{2}$$

Let $R = (r_{ij})_{n\times n}$ be a fuzzy relation on *X*. If R satisfies the following conditions: Anti-reflexivity: for any i, $1 \le i \le n, r_{ii} = 0$; Symmetric: $R^T = R$;

Transmission: $R \subseteq R * R$, then R is called a fuzzy distance relation on X.

R is a fuzzy equivalent relation on X if R^c is a fuzzy distance relation on *X*. If R is a fuzzy relation on X with anti-reflexivity and symmetric, where $|x| = n$, then for any $k \ge n$, R^k is a fuzzy distance relation on X, $R^k = R * R^{k-1}$. For any $s \in \{0,1,\cdots\}$, $X = \{x_1, x_2, \cdots, x_n\}$, R is a fuzzy relation on X, then for any λ, $0 \le \lambda \le 1$, $(R^{s+1})_\lambda = (R*(R^s)_\lambda)_\lambda$. Let $n-rank$ fuzzy Hopfield network FHNN=$(NS, R, \wedge, O, Oper)$, $X = \{x_1, x_2, \cdots, x_n\}$, R is a fuzzy relation on X with anti-reflexivity and symmetric, $\lambda_i = \lambda, \lambda_i \in \wedge, 1 \le i \le n, \lambda \in (0,1)$, $O(0) = e_k$ is an initial state of FHNN, then FHNN can converge to a stable state $\bar{O} \in O_s$, and the elements of $\bar{O}$ whose values are zero are clustered into one class.

B. Approaches to Combined Forecasts by FHNN

Suppose that there are m individual forecasts by model set $\{M_i\}$; Combine them to obtain a more reliable prediction. We use external information, such as historical loads, to cluster the circumstances. The weights of individual models for combination are determined according to their historical performances under similar circumstances.

The hourly load varies from hour to hour; therefore, 24 models were built to forecast the hourly load in this study. The process of modeling for different hours is the same. We will take one hour as an example to discuss the modeling approach in the following.

Two indexes, relative error e_{ij} and mean absolute percentage error E, are used to evaluate the performances of the models.

$$e_{ij} = \frac{\hat{y}_{ij} - y_{ij}}{y_{ij}} \tag{3}$$

Where $\hat{y}_{ij}$ is the forecast load for the ith hour of the jth day and y_{ij} is the actual load of the same hour.

$$E = \frac{1}{n} \sum_{j=1}^{n} \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100 \ \% \tag{4}$$

Where n is the number of load to be forecasted.

As discussed previously, existing load forecasting methods can be divided into the traditional classical methods taking mathematic as theory basis and modern intelligent forecast methods. Combined with characteristics of electric power load, we choose ARMA and GRACH as classic prediction methods, ANN and SVM as intelligent forecast methods.

Figure 2 shows the approaches to combined forecasting by FHNN. Firstly, classify data into three parts (training set, validation set and test set), model and analyze four single models including ARMA, GRACH, ANN, and SVM. Secondly, the combining weights of different models are obtained according to their historical performance on similar days based on FHNN. At last, we build the combined forecasting model and contrast the results with single models. Specific steps are as follows.

Step 1: Initiation. Analyzing original data and classifying them into training set, validation set and test set;

Step 2: Selecting individual candidate models;

Step 3: Training, Validation and Test. Training the candidate models with the training set data; Inputting the validation set and test set to the trained models to obtain validation error e_{ij} and forecasting load p_i;

Step 4: Clustering and calculating the combining weights. Cluster the data into different categories of similar days with FHNN. For any hour of the next day in which the load to be forecasted, we search for a similar circumstance among the historical data and then calculate E of the similar days found to obtain combining weights w_i.

$$w_i = \frac{|1/E_i|}{\sum_{i=1}^{m} |1/E_i|} \tag{5}$$

Where E_i is the mean absolute percentage error of similar days determined by the ith model, i=1,2,…,m.

Step 5: Combine forecasts. Combining forecasting with different candidate models based on the weights determined in Step 4. Contrast the results with single models.

Step 6: End.

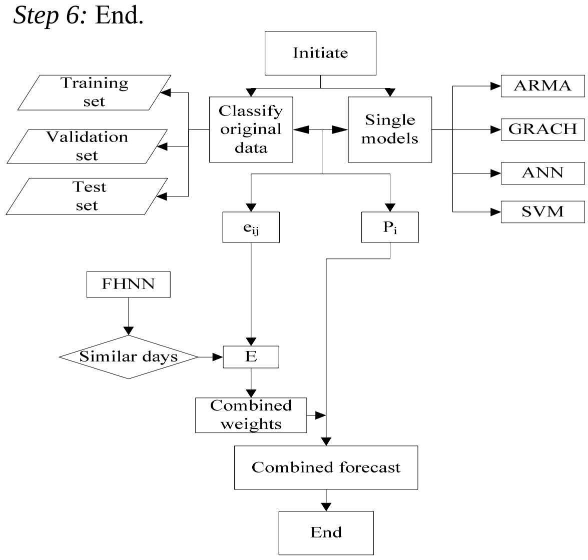

Figure 2. Flowchart of the combined forecasts using FHNN.

III. ANALYSIS ON SINGLE PREDICTION MODEL

As discussed previously, existing load forecasting methods can be divided into the traditional classical methods taking mathematic as theory basis and modern intelligent forecast methods. Combined with characteristics of electric power load, we choose ARMA and GRACH as classic prediction methods, ANN and SVM as intelligent forecast methods.

A. ARMA

ARMA is a kind of common random time-series model, found by Box and Jenkins, also called B-J method, being a kind of short-term prediction method with high precision. Time series model ARMA usually consists of three steps: the recognition model, model parameter estimation, diagnosis and inspection.

Given a time series, the autoregressive moving average (ARMA) model is a very useful for predicting future values in time series where there are both an autoregressive (AR) part and a moving average (MA) part. The model is usually then referred to as the

ARMA(R, M) model where R is the order of the first part and M is the order of the second part. The following ARMA(R, M) model contains the AR(R) and MA (M) models:

$$X_t = c + \varepsilon_t + \sum_{i=1}^{R} \varphi_i X_{t-i} + \sum_{j=1}^{M} \theta_j \varepsilon_{t-j} \tag{6}$$

B. GRACH

Bollerslev's Generalized Autogressive Conditional Heteroscedasticity [GARCH (p, q)] specification (1986) generalizes the volatility forecasting model by allowing the current conditional variance to depend on the first p past conditional variances as well as the q past squared innovations. That is,

$$\sigma_t^2 = L + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 + \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j}^2 \tag{7}$$

Where L denotes the long-run volatility, σ_t^2 denote the conditional variance.

By accounting for the information in the lag(s) of the conditional variance in addition to the lagged t-i terms, the GARCH model reduces the number of parameters required. In most cases, one lag for each variable is sufficient. The GARCH (1, 1) model is given by: $\sigma_t^2 = L + \beta_1 \sigma_{t-1}^2 + \alpha_1 \varepsilon_{t-1}^2$. GARCH can successfully capture thick tailed returns and volatility clustering. It can also be modified to allow for several other stylized facts of asset returns.

C. ANN

The BP neural network is a kind of typical feed forward network. It has three or more layers, all neurons between different layers realizing of complete connection, and no connection in one layer. It is mainly composed by the input, hidden and output layer, each layer achieving complete connection. The input signal inputs from the input layer node, in turn passes through each hidden nodes, and then spreads to the output node.

The BP neural network learning process is composed of information positive dissemination and error back-propagation. In positive spread process, input data from the input layer is handled in hidden layer to output layer, and each layer neuron state affects only the next layer neurons state. If the expected output is not got in the output layer, reverse to back-propagation. At this time, error signal spreads to input from the output layer and adjusts weights between different layers and bias value between each neuron. With error signal smaller, after repeated iteration, the error is less than the allowable value and network training ends.

For any neuron *i*, its input and output relationship may be described as:

$$Y_i = f(\sum_{j=1}^{N} w_{ij} x_j + \theta_i) \tag{8}$$

Where, x_j is the first *j* input, Y_i is the output of neurons, w_{ij} is weights all connected to first *i* neurons, is the threshold value of neurons. *f(x)* is transfer function, generally taking Sigmoid function, such as: $f(x) = \frac{1}{1+e^{-x}}$.

D. SVM

Proposed by V.N.Vapnik in 1995, SVM is a kind of machine learning method on the basis of statistical learning theory.

The least square support vector machine (LSSVM) [16-17] is a kind of improvement method to standard support vector machines (SVM) proposed by Suykens. Through the least square value function and equality constraint, the quadratic programming problem of standard SVM is changed into linear problem, which speeds up training speed and improves convergence accuracy.

The regression function of LSSVM is set to:

$$f(x) = W^T X + b \tag{9}$$

To get w, b, problem is changed into:

$$\begin{cases} \min J(w,\zeta) = \frac{1}{2}\|w\|^2 + \frac{1}{2}\gamma \sum_{i=1}^{l} \zeta_i^2 \\ s.t. y_i = W_T X_i + b + \zeta_i, i = 1,2,\cdots,l \end{cases} \tag{10}$$

Where, l is the number of training sample set $\{(x_1, y_1), \cdots, (x_l, y_l)\}$, $X_i \in R^n$, $y_i \in R$. r is regular factor, w is weight vector. ζ_i is error, $\zeta = [\zeta_1, \zeta_2, \cdots, \zeta_l]^T$; b is constant; Lagrange function is:

$$L(W,b,\zeta,\alpha) = J(W,\zeta) - \sum_{i=1}^{l} \alpha_i (y_i - W^T X_i - b - \zeta_i), i = 1,2,\cdots,l \tag{11}$$

Where l is Lagrange multiplier.

After a series of reduction, this paper studies the regression function:

$$f(X) = \sum_{i=1}^{l} \alpha_i k(X_i, X) + b \tag{12}$$

Where K is kernel function of symmetrical function meeting Mercer conditions, $K(x_i, x_j) = \varphi(x_i)^T \varphi(x_j)$.

Ⅳ. CASE STUDY

A. Raw Data Analysis

Take power load data in some province, China in 2009 as original data; figure 3 shows 8760 hours of load data one year. As can be seen from the figure, electric power load of one year is greatly volatility and regularity. The load one year is between 100MW and 900MW, with the higher load in winter and lower in summer. Here we apply the proposed method to forecast the hourly load in August 2009. The data of the previous two months served as a training set and validation set, respectively.

At the same the loads on different hours show great variability. Therefore, in this study, 24 individual models were built to forecast the load for the 24 different hours of a day. Thus, a daily load ratio $r_{i,j}$ reflecting the load fluctuation on different types of days is proposed to remove the variations between the different weekdays, as demonstrated by Equation 13.

$$r_{i,j} = \frac{\sum_k p_{i,j,k}}{30}, \quad i=1,2,\ldots,7; \ j=1,2,\ldots,24; k=1,2,\ldots,30 \tag{13}$$

Where p_{ijk} is the jth hourly load on the it^h day of the kth week in 2009. Because the loads in August are to be forecasted, the loads over the previous 22 weeks (before August 1) are used to calculate r_{ij} .

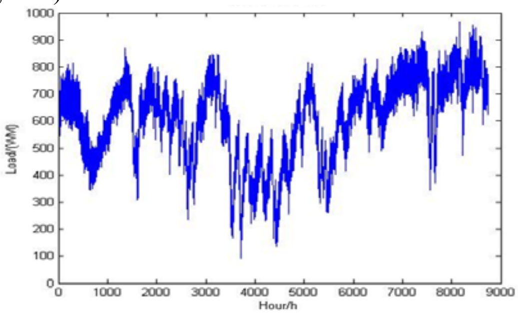

Figure 3. Hourly load line.

B. Analysis and Forecasting of Individual Models

As discussed previously, ARMA, GRACH, ANN and SVM are chosen as compared and individual models to forecast power load. We build the individual models with the training set and then obtain the validation error for the validation set as shown in Table1 and Figures A1-A4.

From these figures, we can see that the SVM performs better than the other models. Most of the errors of the SVM lie between -0.03 and 0.03, and there are fewer error spikes. Load around 7:00 and 13:00 on the 4th day are overestimated, and load around 13:00 on the 7th and 15th day underestimated.

Figure A2, which illustrates the relative errors of the validation set of ANN, shows that load from 13:00 to 19:00 on the days from the 8th to the 10th day are overestimated. Just as in the SVM, some loads on the 7th day are underestimated.

Figure A3 shows the validation errors of the GARCH model; it also shows that the worse performance occurs at the peak load time, from 13:00 to 16:00. It should be noticed that the model fails to provide an acceptable forecast. More extreme errors appear in Fig. 8 than in the Figures A1 and A2. It should also be noted that the errors within one day show great consistency, i.e., they are either below zero or beyond zero. The reason for this is that GARCH is a classic time series model, which relies heavily on historical load.

Figure A4 shows the validation errors generated by ARMA. It can be seen that more variations appear in it. The distribution of errors generated by this method is similar to that generated by the GARCH model. This is because both methods use the historical load information.

Figure A5 to A8 show the test errors using the same methods discussed above. The test performances of the models are very similar to the validation performances analyzed. In addition, the test errors of the two kinds of intelligent modeling, SVM and ANN, are similar, and the errors of the two other methods are similar as well. In general, the intelligent modeling shows better performance than the latter.

C. Combined Forecasting with ARMA-GRACH-ANN-SVM based on FHNN

Firstly, we cluster the circumstances into different categories using FHNN and then calculate the combining weights of different models according to their performance under different circumstances. The first- and second-layer dimensions of the FHNN are set to 4. By combining weights and the individual forecasts, the combination forecast load is obtained.

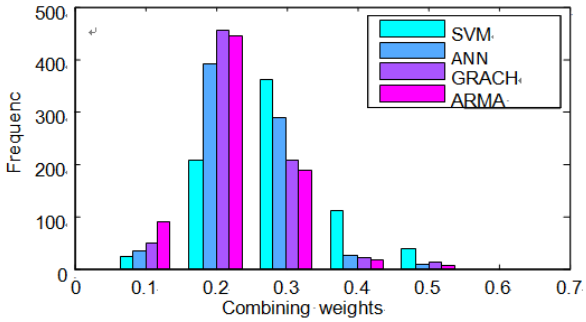

Figure 4. Combined weights of different models.

Figure 4 is a histogram of the combining weights of the four models for the 744 hourly prices in August 2009. It can be seen that the combining weights of the SVM, mostly between 0.3 and 0.6, are obviously larger than those of the other models because the average weights of the four models are 0.25. To the contrary, the combining weights of ARMA are smaller (more weights are less than 0.2) than those of other models. From the performance comparison between the models, we know that the distribution of combining weights for the different models is just as we expected

Figure 5 shows the test errors generated by FHNN combined modeling. It can be seen that Figure 5 is quite similar to Figure A5. We know that the model with the least forecasting error has the largest combination weight, which is reasonable because the SVM outperforms the other models. It should be noted that the most extreme errors occur in the afternoon and evening. On some days, the load during those periods are overestimated, such as on the 7th, 17th and 23th day; on some days, the prices are underestimated, such as on the 2nd day.

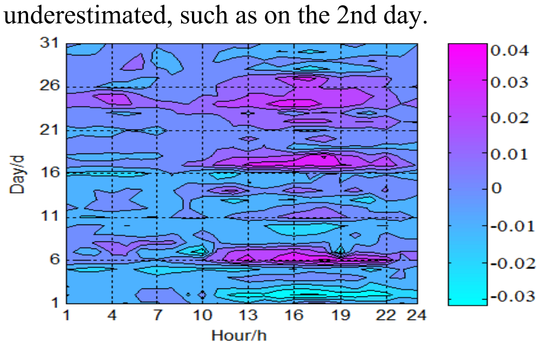

Figure 5. Test error (relative error) of combined modeling.

Table 1 is a comparison between the individual modeling and the combination forecast. Here we predict power load of 744 hours in August. To illustrate and drawing conveniently, we average load and error according to 24 hour points of 31 days. Then the average load and error are determined by different models. It can be seen that combined modeling clearly improves the forecast accuracy. The mean absolute percentage error of ARMA is 2.69%, which is the best performance and GRACH 2.10%, ANN 1.96%, SVM 1.79%. The mean absolute percentage error of combined modeling is 1.51%, much lower than that of the single modeling and even the best individual model, SVM.

TABLE I.
PERFORMANCE COMPARISON OF DIFFERENT MODELS

Tiem point	Actual load	ARMA		GRACH		ANN		SVM		Combined model	
		Forecast load	Error	Forecast load	Error	Forecast load	Error	Forecast load	Error	Forecast load	Error
1	561.59	552.64	-1.59%	549.04	-2.24%	555.04	-1.17%	569.46	1.40%	562.82	0.22%
2	542.21	528.42	-2.54%	530.5	-2.16%	546.16	0.73%	539.41	-0.52%	521.89	-3.75%
3	522.87	508.49	-2.75%	514.02	-1.69%	528.21	1.02%	529.42	1.25%	525.66	0.53%
4	513.2	533.1	3.88%	525.16	2.33%	498.12	-2.94%	508.46	-0.92%	519.44	1.22%
5	503.96	518.45	2.87%	512.04	1.60%	512.05	1.60%	495.12	-1.75%	500.47	-0.69%
6	504.02	498.15	-1.16%	500.13	-0.77%	498.54	-1.09%	512.14	1.61%	509.15	1.02%
7	541.74	531.78	-1.84%	524.01	-3.27%	560.78	3.51%	557.1	2.83%	559.13	3.21%
8	524.75	542.06	3.30%	518.21	-1.25%	510.34	-2.75%	521.64	-0.59%	522.46	-0.44%
9	493.35	502.16	1.79%	489.06	-0.87%	501.78	1.71%	479.23	-2.86%	489.12	-0.86%
10	487.64	500.78	2.70%	475.01	-2.59%	480.42	-1.48%	479.34	-1.70%	492.04	0.90%
11	509.59	481.24	-5.56%	517.83	1.62%	511.42	0.36%	514.23	0.91%	511.78	0.43%
12	574.47	561.04	-2.34%	568.43	-1.05%	569.12	-0.93%	598.12	4.12%	570.46	-0.70%
13	506.2	482.64	-4.65%	494.05	-2.40%	512.41	1.23%	519.31	2.59%	512.47	1.24%
14	489.63	471.08	-3.79%	475.01	-2.99%	498.46	1.80%	499.47	2.01%	499.24	1.96%
15	503.96	494.72	-1.83%	486.24	-3.52%	516.71	2.53%	516.24	2.44%	513.12	1.82%
16	490.66	495.07	0.90%	497.05	1.30%	513.45	4.64%	500.14	1.93%	504.89	2.90%
17	500.83	518.61	3.55%	530.47	5.92%	514.34	2.70%	516.78	3.19%	521.14	4.06%
18	528.16	498.31	-5.65%	533.08	0.93%	542.04	2.63%	531.48	0.63%	526.44	-0.32%
19	548.58	530.94	-3.22%	552.04	0.63%	540.74	-1.43%	558.01	1.72%	540.97	-1.39%
20	544.79	540.91	-0.71%	530.34	-2.65%	538.1	-1.23%	550.46	1.04%	538.91	-1.08%
21	539.13	548.06	1.66%	546.31	1.33%	531.04	-1.50%	528.13	-2.04%	544.42	0.98%
22	540.76	556.03	2.82%	554.14	2.47%	547.04	1.16%	549.12	1.55%	554.23	2.49%
23	573.81	591.34	3.06%	589.16	2.68%	560.24	-2.36%	574.89	0.19%	589.46	2.73%
24	568.59	570.91	0.41%	580.18	2.04%	549.2	-3.41%	566.72	-0.33%	566.1	-0.44%
E			2.69%		2.10%		1.96%		1.79%		1.51%

V. CONCLUSION

In this paper, a new short-term power load combined forecasting method is proposed. By analyze power load features, we classify original data into three parts including training set, validation set and test set model. Then four single models, including ARMA, GRACH, ANN and SVM, are chosen as classic methods and intelligent methods. Forecast validation set and test set model on the basis of training set separately by four single models and obtain the validation error and test error. The combining weights of different models are obtained according to their historical performance on similar days based on FHNN. At last, we build the combined forecasting model with ARMA-GRACH-ANN-SVM weighted by average with the weights obtained from FHNN clustering. A case using power load data in some province, China in 2009 is analyzed. By contrast the errors between single models and combined model we can see that the proposed techniques improved the forecast significantly.

APPENDIX A FIGURES

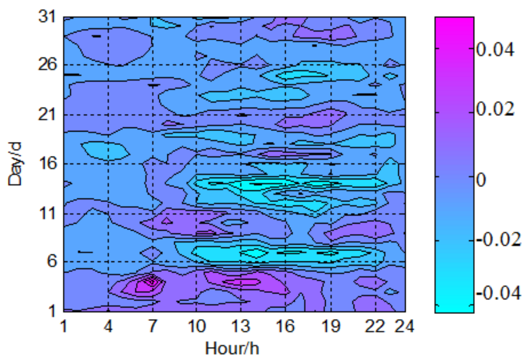

Figure A1. Validation error (relative error) of SVM.

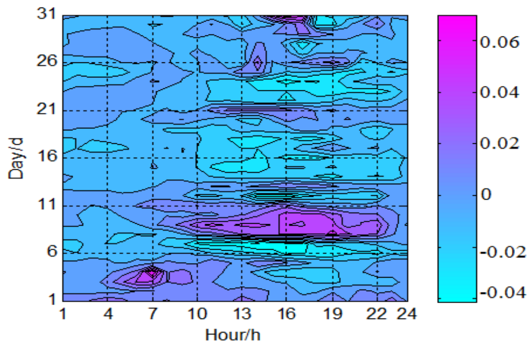

Figure A2. Validation error (relative error) of ANN.

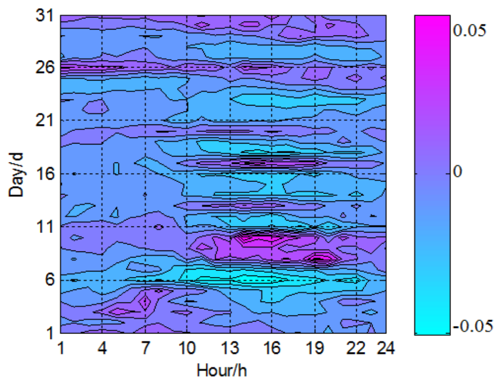

Figure A3. Validation error (relative error) of GRACH.

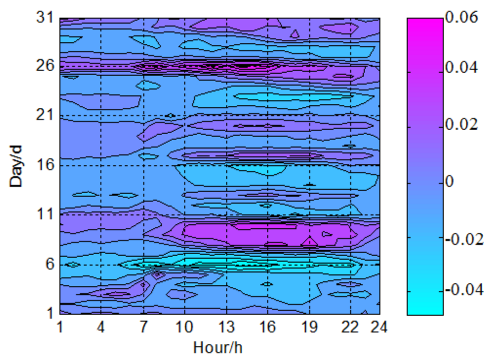

Figure A4. Validation error (relative error) of ARMA.

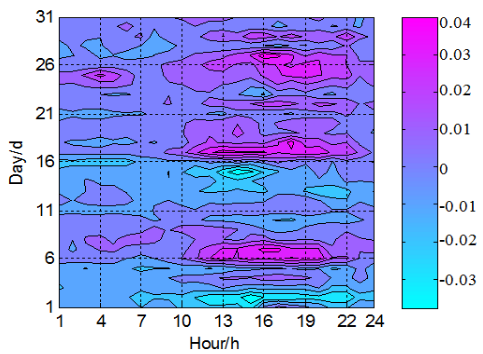

Figure A5. Test error (relative error) of SVM.

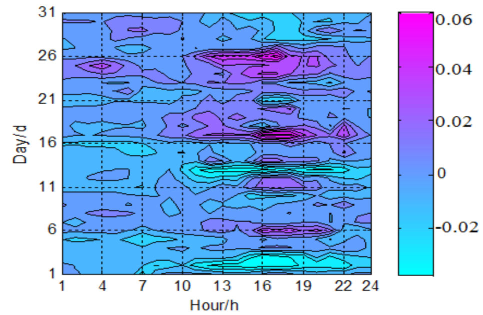

Figure A6. Test error (relative error) of ANN.

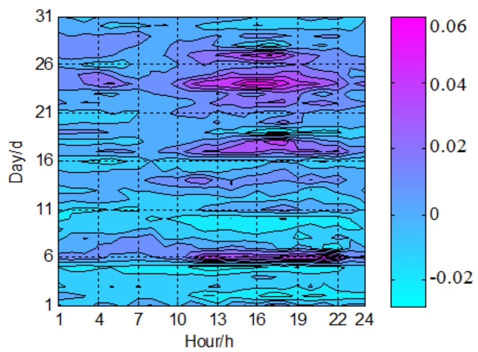

Figure A7. Test error (relative error) of GRACH.

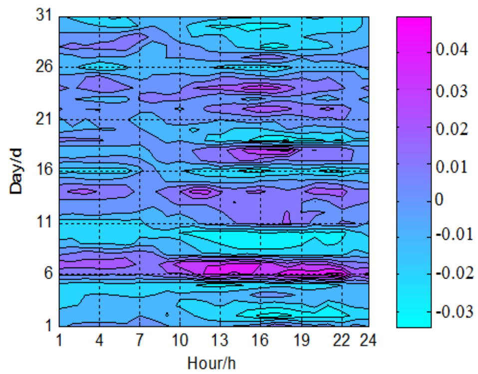

Figure A8. Test error (relative error) of ARMA.

ACKNOWLEDGMENT

This work is supported in part by The National Natural Science Foundation of China (NSFC) under Grant Nos. 71071052 and The Fundamental Research Funds for the Central Universities under Grant Nos. 12QX22.

REFERENCES

[1] Niu DX, Cao SH, Lu JC, Power load forecasting technology and its application, China Electric Power Press, 2009.
[2] Pappas S. Sp., Ekonomou L., Karampelas P., etal, “Electricity demand load forecasting of the Hellenic power system using an ARMA model. Elect”, Power Sys. Research, vol.80, No.3, pp.256-264, 2009.
[3] H. Pao. “Forecasting electricity market pricing using artificial neural networks”, Energy Conv. and Manag. Vol.48, pp.907-912, 2007.
[4] Li XM, Gong DC, Li LF, “Next day load forecasting using SVM. Advances in NN”, Vol.3498, pp.634-639, 2005.

[5] JM Tang, JY Liu, K Yang, "Short-Term Load Combination Forecasting by Grey Model and Least Square Support Vector Machine", Power Syst. Techno.，Vol.33, No.3, pp.63-68, 2009.
[6] CQ Kang, Q Xia, Y Shen, "Integrated model of power system load forecasting", J of Tsinghua Univ. Vol.39, No.l, pp. 8-11, 1999.
[7] CQ Kang, X Cheng, Q Xia, "Novel Approach Considering Load-relative Factors in Short-term Load Forecasting", Elect. Power Syst. Research, Vol.70, No.2, pp. 99-107, 2004.
[8] Anandalingam G, Chen L, "Linear Combination of Forecasts: A General Bayesian Model", J of Forecasting, vol.8, No.3, pp.199-213, 1989.
[9] JD Xie, GQ Tang, GF Xu, "The application of combination forecasting method in power load forecasting", China Power, vol.31, pp. 3-5, 1998.
[10] Bunn D W, "A Bayesian Approach to the Linear Combination of Forecasts", ORQ, vol.26, No.2, pp. 325-329, 1975.
[11] KG Xie, CY Li, JQ Zhou, "Combination load forecasting model based on neural network", Chinese J of EE, vol.22, No.7, pp. 85-89, 2002.
[12] N Lu, JZ Zhou，YY He, "Simulated Annealing Theory based on Particle Swarm Optimization for Support Vector Machine Model in Short-Term Load Forecasting", International Conference on Information Engineering and Computer Science， Wuhan, China, 2009.
[13] Jolai, F., Ghanbari, "A. Integrating data transformation techniques with Hopfield neural networks for solving travelling salesman problem", Expert Syst. with Applic. Vol.37No.7, pp.5331-5335, 2010.
[14] Shen, Y. J., Wang, M. S, "Broadcast scheduling in wireless sensor networks using fuzzy Hopfield neural network", Expert Syst. with Applic. Vol.34No.2, pp.900-907, 2008.
[15] WY Hsu, "Fuzzy Hopfield neural network clustering for single-trial motor imagery EEG classification", Expert Syst. with Applic. Vol.39, No.1, pp.1055-1061, 2011.
[16] Vapnik V N，Cortes C, "Support vector networks", ML，vol.20, No.3, pp.273-297, 1995.
[17] SUYKENS J A K，GESTEL T V，BRABANTER J D, Least squares support vector machines, Singapore: World Scientific Press, 2002.

Dongxiao Niu is born in Anhui province, China and birth of October, 1962. He got the bachelor degree of applied mathematics in Wuhan University of hydraulic and electric engineering from September 1978 to July 1982, and technology economy and management professional doctorate of North China Electric Power University from August 1999 to April 2002.

His work experience is as follows. September 1982 till now, teaching in North China Electric Power University. September 2005 till now, head of School of Economics and Management, North China Electric Power University. Some publications are as follows.

[1] D.X Niu, S.H. Cao, J.C. Lu, et al, Technology and application of electric power load forecasting, Beijing, China, China Electric Power Press, 2009.

[2] D.X Niu, J.C Li, J.Y Li, et al, "Middle-long power load forecasting based on particle swarm optimization", Computers and Mathematics with Applications, 2009, Vol. 57, No. 11, pp.1883-1889, 2009,.

[3] D.X. Niu, Y.L. Wang, C.M. Duan, et al, "A New Short-term Power Load Forecasting Model Based on Chaotic Time Series and SVM", Journal of Universal Computer Science, Vol. 15, No. 13, pp.2726-2745, 2009.

His main research directions include project predict, decision theory and application, project comprehensive evaluation method and application.

Prof. Niu is high master of National Grid Economic Research Institute. He is also the editor or Board Member of "IGJEI", "East China electric Power", "Journal of North China Electric Power University", "Electric Power Automation Equipment", "Science of Science and Management of S. & T" and so on. He has awarded the first prize of National commercial science and technology progress in 2009, third prize of Scientific and Technological progress award in 2009 and so on.

Yanan Wei is born in Henan province, China and birth of October, 1987. She got the bachelor degree of management in North China Electric Power University from September 2005 to July 2009. Now she is studying in School of Economics and Management North China Electric Power University from September 2011. Her main research directions include technical economy forecast and evaluation, decision theory and application.