The Load Forecasting Model Based on Bayes-GRNN

Yanmei Li
School of Business and Administration, North China Electric Power University, Baoding 071003, China
Email: liyanmei28@yahoo.com.cn

Jingmin Wang
School of Business and Administration, North China Electric Power University, Baoding 071003, China

***Abstract*—Comparison with the classical BP neural network, the generalized regression neural network requires not periodic training process but a smoothing parameter. The model has steady and fast speed, and meanwhile, the connection weight of different neurons is not necessary to be adjusted in the training process. The paper establishes the index system of GRNN forecasting model, and then uses Bayes theory to reduce them, which will be inputting variables of GRNN model. The method is testified to get higher speed and accuracy by simulation of actual data and comparison to classical BP neural network.**

***Index Terms*—Bayes, load forecasting, generalized regression neural network**

I. Introduction

With the introduction of the concept of smart grid, a large number of distributed power grid to run, it have brought new challenges to the distribution network planning, construction and operations. Since a large number of users install DER to provide electricity, making the distribution network planners difficult to accurately predict the load growth, thereby affecting the plan's rationality. It's been a long time that many electric operators commit themselves to the investigation of electric system load forecasting technique and have obtain some production to a certainty, for example time sequence, regression analysis, gray theory, artificial neural network and so on. The artificial neural network has powerful collateral disposal mechanism, imminent capability of arbitrary function, learning capability and self-organization and self-adaptation capability. And it is capable of considering the impact of variable factors such as weather, temperature and so on. So it has been widely applied in the field of load forecasting and decision-making. Comparison with the classical BP neural network, the generalized regression neural network requires not periodic training process but a smoothing parameter. The model has steady and high speed, and meanwhile, the connection weights of different neurons are not necessary to be adjusted in the training process. And since change factor affecting load is numerous, if history load , temperature information , weather information , date type , week type etc., these different , are difficult to ascertain that because of field time factor reason, self fit in with chooses the all unable entering characteristic real time field forecasting method therefore, various but can only be based on experience in advance selected. Bayes decision theory concept the main body of a book is made use of and method entering characteristic in choosing short-term load forecasting.

II. Bayes Theory

The good according to that Bayes viewpoint , essential points work out one decision-making , the information ought to make use of possessions to be able to gain, include sample book information and all information first in sampling and come from experience , consciousness , the subjective knowledge judging, these subjective knowledge same be valuable knowledge wealth, the middle responding to the formal introduction arrive at the deduction counting and making policy goes to, but this exactly is that classics statistics has not given think. In classical statistics deduction, only, admit and make use of sample information, but non-recognition, or make use of the subjective judgment and consciousness. Bayes theory exactly is to hope that the judgment and intuition lead subjectivity into the basis arriving at the analysis process counting deduction and building information thereby in decision analysis, inferring and making policy synthetically formally.

Bayes decision theory concept and method are used for fields such as engineering, management science already by broad field, adaptively selects input features step with Bayes method as follows:.

A Ascertains a priori probability distribution.

The priori probability $P(\overline{\omega_j})$ represents the estimation to the probability distribution of the variable $\overline{\omega_j}$, it has reflected a priori knowledge to the variable, has included practical experience and subjective judgment etc. And that a priori probability scatters in load forecasting , is what be remained to be chosen influencing factor and their probability aggregation , restricting condition is whose necessary probability greater than zero, whose

doi:10.4304/jsw.7.6.1273-1280

combination is 1. The load at the same time, forecasting (namely "big close small distant" distance according to load forecasting nearest segment of period affects maximal) and "identical date load type similarity" characteristic, the priori probability distribution is

$$\begin{cases} P_{per} = \alpha P_{dis} \\ P_{equ} = \beta P_{div} \\ P_{per} = \gamma P_{equ} \\ \sum P_{per} + \sum P_{dis} + \sum P_{equ} + \sum P_{div} = 1 \\ P_{per} > 0, P_{dis} > 0, P_{equ} > 0, P_{div} > 0 \end{cases}$$

In the formula, P_{per} and P_{dis} represent the distance from influencing factor of forecasted point far or close; and respectively, P_{equ} and P_{div} represent the priori probability of influencing factor that having identical or different date type; Take α, β, γ value range being $(1, \infty)$, may look at concrete conditions but fix. When $\alpha = 1, \beta = 1, \gamma = 1$, the priori probability distribution is a uniform one.

B Ascertains a likelihood function.

The likelihood function has been a condition probability essentially it has reflected sample information, whose function value has been called likelihood rates, has been. $P(\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_N | \bar{\omega}_j)$. During the period of load forecasting, if the input variable is selected as $\bar{\omega}_j$, we may make use of N samples $\bar{x}_n (n = 1, \cdots, N)$ to calculate the Error E according to the follow formula.

$$E_i = \frac{1}{n} \sum_{j=1}^{n} R_{valid\_j} [f_{train}]$$

$$R_{valid\_j} [f_{train}] = \frac{1}{N_{valid}} \sum_{\bar{x}_j \in Z_{valid}} \frac{\left[y_j - f_{train}(\bar{x}_j) \right]}{y_j} \times 100\%$$

$R_{valid\_j}[f_{train}]$ is relative mean error in style, from training collection to get regression function f_{train} relativity on effective set; N_{valid} be an effective set Z_{valid} all together sample book number; y_j be that actual load value; $f_{train}(\bar{x}_j)$ be forecasting load value

C Calculates posterior probability. The formula of posterior probability

$$P(\bar{\omega}_j | \bar{x}_1, \bar{x}_2, \cdots, \bar{x}_N)$$

Among them, a priori probability and likelihood function action input vectors, a posteriori probability is output vector. Because a posteriori probability has synthesized a priori knowledge and sample book information, ultimate being to decide which group of influencing factors to choose being input vector is standard. A posteriori probability is increasingly big, the probability that input vector is pitched on is increasingly big.

D Chooses the input vector $\bar{\omega}_j$.

According to the actual characteristic being unlike area load, choose M may affect bigger factor composition to load waiting for choosing influencing factor, again out of, the random chooses different influencing factor combination , forms input vector $\bar{\omega}_j$.

Owing to Bayes theory, the logic block diagram of input variable adaptively selected will be show:

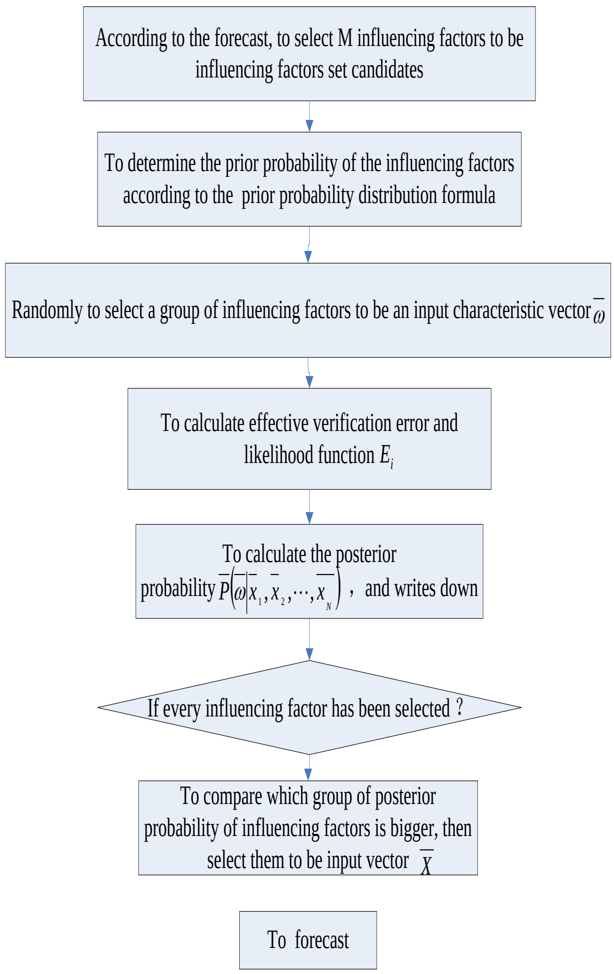

Figure 1. Logic diagram of Bayes theory

III. GRNN NEURAL NETWORK

GRNN network is a new type of neural network proposed by Donald Specht in The Lockheed Paio Alto Research Laboratory. It is built on the basis of mathematical statistics, which can approximate intimate relations according to the sample data and mainly be used for system model and forecasting. It is different from traditional neural network. The model has a clear theoretical basis and it is a neural network established on the basis of mathematical statistics.

The advantage is fast learning. The network finally converges to the most optimal cluster sample regression plane. Once the learning samples are determined, the corresponding network structure and connection weights between neurons also will be determined. Network training process is actually the process of determining smoothing parameter. Even scarcer in the sample data, the effect is also very good, the network can handle data uncertainty. General regression neural network is characterized by a few manual adjustments of the parameters, which only need to adjust parameter of σ .The learning of the network all relays on data samples, and it is fast to learn. The characteristics of the decision avoided the network to maximize the subjective impact of assumptions on the predicted results. It is an ideal means and tools for surface fitting and modeling.

As shown in Figure 2, GRNN structure has four neurons, including the input layer, pattern layer, summation layer and output layer. Corresponding network input is $X = [X_1, X_2, \cdots X_m]^T$ and the output is $Y = [Y_1, Y_2, \cdots, Y_l]^T$

The number of neurons is equal to the number of training samples m in the input layer. The distribution of neuronal cells is simple, directly passing input variables to hidden layer.

The number of neurons is equal to the number of training samples n in the pattern layer, the neurons corresponding to different samples. The transfer function of neurons i in pattern layer is

$$P_i = \exp\left[(x - x_i)^T (x - x_i) / 2\sigma^2\right] \quad i = 1,2,\cdots,n$$

X is the network input. Xi is study sample corresponded to neuron i. δ is smoothing parameter, that is, neuron i's output is exponential of Euclid square between the input variable X and corresponding sample Xi.

$$D^2 = (X - X_i)^T (X - X_i)$$

Summation layer consists of two types of neurons. One neuron model for all the output neurons arithmetic sum and the pattern layer's neurons and neuronal connections of the right value is 1.The transfer function is

$$S_D = \sum_{i=1}^{n} p_i \quad i = 1,2,\cdots,l$$

The other neuron model is for all the output neurons weighted sum. Neuron i in the mode layer sum neuron j in the summation layer .Connection weights between neurons is the first j elements Y_{ij} in the output of the first i samples Yi The transfer function of summation neurons is

$$S_{nj} = \sum_{j=1}^{m} Y_{ij} P_i \quad j = 1,2\cdots,l$$

The number of Neurons in the output layer is equal to learning sample dimension L in the output vector and the summation layer neuron output will be divided, that is:

$$y_J = S_{NJ} / S_D$$
$$J = 1,2,\cdots,l$$

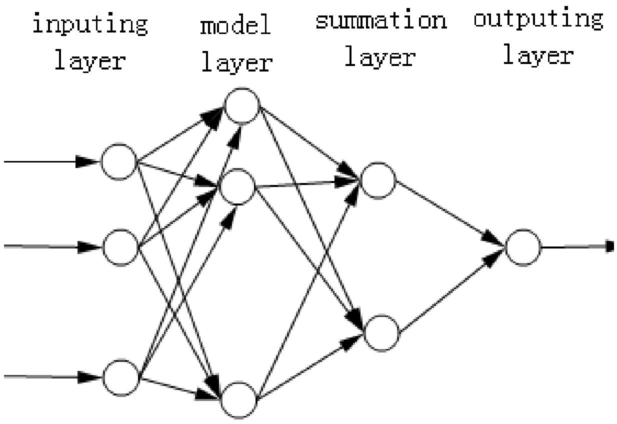

Figure 2. Structure of GRNN.

Generalized regression neural network theory is based on nonlinear (nuclear) regression analysis. Set the joint probability density function of random vector x and random variable y as $f(x, y)$, x values as x_0, the return value of y on x_0 is:

$$\hat{y}(x_0) = \frac{\int_{-\infty}^{\infty} y f(x_0, y) dy}{\int_{-\infty}^{\infty} f(x_0, y) dy} \quad . \tag{1}$$

Use Parzen nonparametric estimation; estimate the density function with the sample data set $\{x_i, y_i\}_{im1}^{n}$ according to Eq. (1):

$$f(x_0, y) = \frac{1}{n(2\pi)^{\frac{p+1}{2}} \sigma_1 \sigma_2 \cdots \sigma_p \sigma_y} \sum_{i=1}^{n} e^{-d(x_0, x_i)} e^{-d(y, y_i)} \tag{2}$$

Among it,

$$d(x_{0,} x_i) = \sum_{j=1}^{n} \left[(x_{0j} - x_{ij}) / \sigma_j\right]^2, d(y, y_i) = (y - y_i)^2$$

Wherein, n is the sample size; p is the dimension of x ; σ is the width coefficient of Gaussian function, called the smoothing parameter here. Substitute Eq.(2) in Eq.(1), and exchange the order of integral and additive, then:

$$\hat{y}(x_0)=\frac{\sum_{i=1}^{n}\left(e^{-d(x_0,x_i)}\int_{-\infty}^{+\infty}ye^{-d(y,y_i)}dy\right)}{\sum_{i=1}^{n}\left(e^{-d(x_0,x_i)}\int_{-\infty}^{+\infty}e^{-d(y,y_i)}dy\right)} \quad (3)$$

As $\int ze^{-x^y}dz=0$, calculation result two integrals of Eq.(3)

$$\hat{y}(x_0)=\frac{\sum y_i e^{-d(x_{0,}x_i)}}{\sum_{i=1}^{n}e^{-d(x_{0,}x_i)}} \quad (4)$$

It can be seen that the prediction value $\hat{y}(x_0)$ in Eq.(4) is equal to weighted sum of the value of dependent variables y_i of all samples.

A Smoothing parameter optimization

In the training process, learning algorithm of Generalized regression neural network adjust the smoothing parameter σ instead of adjusting the connection weights between neurons, to adjust the transfer function of each unit in model layer so as to get the best regression estimation.

(1) Estimate smoothing parameter σ with ordinary differential evolution algorithm (DE)

Parameter estimation is to solve the value of σ when

$$d(x_0,x_i)=\sum_{j=1}^{p}\left[(x_{0j}-x_{ij})/w_j\right]^2$$ or

$d(y,y_i)=(y-y_i)^2$ is minimum value.

DE method steps: firstly determine the domain of σ as Ω, and width range of the i-th component as $h(i)$, take σ as individual, $q(\sigma)$ as fitness function, then perform the following steps:

(A) Select population size N, Weighting factor F= [0, 2], maximum evolving algebra M and hybridization rate C∈ [0, 1].

(B) Generate initial population σ^0 :{ σ_i^0 (i=1 , 2, …,N)}, set evolving algebra G=0.

(C) Calculate the fitness $P(\sigma_i^G)$ of each individual and the best individual σ_b^G of G-generation.

(D) Perform step (E) to step (G) with σ_i^G (i=1, 2, …, N) to generate the G+1-th generation population.

(E) Mutation operation, temporary variation individual $\hat{\sigma}_i^{G+1}$ is defined in Eq.(5), wherein, 1≤j, k, l≤N, and i, j, k, l differ from each other, F will control the degree of variation of differential item $\sigma_b^G+\sigma_i^G-\sigma_k^G-\sigma_l^G$ against σ_i^G.

$$\hat{\sigma}_i^{G+1}=\sigma_i^G+F(\sigma_b^G+\sigma_i^G-\sigma_k^G-\sigma_l^G) \quad (5)$$

(F) Hybridization operation: according to Eq.(6), create a hybrid between σ_i^{G+1} and this generation individuals to generate the next generation individuals, wherein, $\hat{\sigma}_{lj}^G$ is the j-th gene of the i-th individual of the G-th generation, C is the hybridization rate, random number originates from [0, 1] uniformly.

$$\sigma_{ij}^{G+1}=\begin{cases}\sigma_{ij}^G & \text{Random}\ number > C\\ \hat{\sigma}_{ij}^{G+1} & others\end{cases} \quad (6)$$

(G) Selecting operation: according to Eq.(7), filial generation individuals compete with the parent to choose the next generation individuals.

$$\sigma_i^{G+1}=\begin{cases}\sigma_i^{G+1} & P(\sigma_i^{G+1})\le P(\sigma_i^G)\\ \sigma_i^G & others\end{cases} \quad (7)$$

(H) G+1→G.

(I) If G exceeds the maximum evolving algebra M, or if the best fitness value difference between the G-th generation and the G+k-th generation is not greater than esp., go to step (J); otherwise return to Step (C). Where k is non-negative integer, which can be set by the user according to accuracy requirement, esp. =10^{-8} .

(J) Take individual σ_b^G with the best fitness value as parameter estimation value in the last generation population.

B Estimate the smoothing parameter σ with the modified differential evolution algorithm (MDE)

a) Maintain the diversity of species

Inbreeding will lead to degradation, it's easy to make an individual approach to a local optimum if the evolution shrinking too fast, and resulting in inbreeding of the next generation. For maintaining the diversity of species, this paper reserves the best individual and initializes another individual, which is called resetting operation, and designs the parameters to measure distribution range of offspring, as Eq.(8) shows.

$$\lambda^G = \sum_i \frac{\sum_j (\sigma_{ij}^G - \sigma_{i^*}^G)}{h_i \times N} \quad (8)$$

Among it, $\sigma_{i^*}^G = \frac{1}{N}\sum_j \sigma_{ij}^G$.

When λ^G is less than the lower limit $\lambda_{\min}$, which is calling resetting operation, what will retain the best individual σ_b^G and generate a number of individuals in its surroundings according to the normal distribution.

b) Design optimization operation

Optimization operation is introduced to use of evolutionary information, to implement deterministic optimization timely according to the trend of fitness function. Simplex method is a good optimization operation, which has excellent search ability, no need to calculate derivative, and is easy to implement. Design simplex optimization operation, which is optimizing from an individual with simplex method, and calling it according to variable frequency (known as optimization rate). Reduce optimization rate when the rate of shrinkage is fast; otherwise improve. Therefore, compare distribution range of two generations population, change the optimization rate according to Eq.(9), where the parameters are: $s_1 > s_2 > 0$.

$$P_X = \begin{cases} P_x + 0.05(1-P_x) & (\lambda^G / \lambda^{G-1}) \ge s_1 \\ P_x - 0.05P_x & s_1 > (\lambda^G / \lambda^{G-1}) \ge s_2 \\ P_x & others \end{cases} \quad (9)$$

c) Steps of MDE

set optimization target as the fitness function P(σ), forward six implementing steps of MDE is basically the same with DE, only at the A-th step lower limit $\lambda_{\min}$ and initial value p_X of optimization rate should be selected, the cycle range of the D-th step is from step(E) to step(H), and the G-th step is as follows:

(G) Simplex optimization operation: random number r originates from [0, 1] uniformly, if $r>p_x$, then go to step (H), or implement this step: take σ_i^G and two randomly selected individuals σ_j^G and σ_k^G to form the initial type of optimization with simplex optimization to get the best individual σ_{ib}^{G+1} and replace σ_i^G. Design simplex on two-dimensional subspace in the domain to ensure that is convex.

(H) Selecting operation: it's the same with the G-th step of DE.

(I) G+1→G.

(J) Resetting operation: calculate λ^G according to step (H), if $\lambda^G \le \lambda_{\min}$, then implement resetting operation.

(K) Change the optimization rate according to Eq.(9).

(L) If G exceeds the maximum evolving algebra M, or if the best fitness value difference between the G-th generation and the G+k-th generation is not greater than esp., go to step (M), wherein, k, eps. and DE are the same.

(M) Take the best individual $\sigma_b^{G_j}$ as parameter estimation value in the last generation population.

IV. CASE ANALYSIS

In this paper, Southern Hebei Network of the 2009 data were analyzed to September 10, 2009 to September 19 points for the entire sample for the study and active load to September 20 for the entire load for the test samples were forecast in September on the 21st load. 52 selected by experience on the condition variable attributes, which is 12 load data, that is, on the 10th to the 19th day the whole point of load; The remaining 40 non-load data, including weather, the date type, sunshine duration, maximum temperature, minimum temperature, average temperature, the biggest humidity, humidity, such as minimum 40 factors, including the date and type of rest days (including weekends and the statutory rest), the weather conditions on the provision of meteorological information is divided into 17 types. The neural network input vector, using the above-mentioned were all rough intensive SR algorithm primaries to the impact of load reduction factors. Table I lists a variety of reduction algorithm results. Can be seen, after attribute reduction, the input vector be simplified.

TABLE.I
RESULTS AFTER BAYES REDUCTION METHOD

method	variables	number of hidden nodes	time/min	η/(%)
before reduction	40	13	53	10.89
after reduction	29	12	12	6.13

TABLE. I has carried out a form from the function forecasting accuracy and training time the forecast to Tab. II kinds method in two aspects comparing, has forecast accuracy among them adopt average proportional error (η) to be the analytical index, whose definition as

$$\eta = \frac{\left[\sum \frac{(|P_{Ai} - P_{Fi}|)\cdot 100}{P_n}\right]}{n}$$

follows: Among them,

P_{Ai} be to forecast value, P_{Fi} be actual value, n is to forecast number of times.

We can see from Tab.1, making use of Bayes reduction, the input variables reduces from 43 to 31, calculating time reduces to 13min, the average relative error decrease from 11.11% to 5.82%, not only the attribution was reduced in the maximal degree, but the computing time is shortest relatively, and the calculation error is minimum. It is shown from the training time and accuracy. The method in the paper is better than traditional BP Neural Network, so it is suitable for the forecasting samples in this area.

Normalized the value of input and output variables to [0, 1]. Regulation is one of the many ways here by the following formula:

$$\hat{x} = \frac{x - x_{\min}}{x_{\max} - x_{\min}}$$

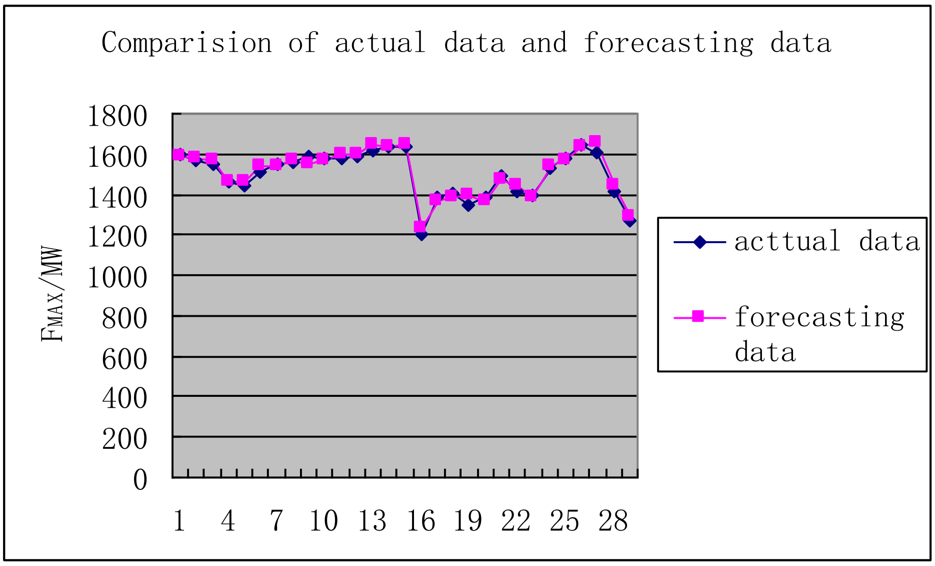

Figure 3. Comparison of actual data and forecasting data of max load everyday

TABLEI
COMPARISON RESULTS OF DIFFERENT METHOD

forecasting method	March		July	
	η/(%)	t/s	η/(%)	t/s
method in this paper	0.68	9.7	0.68	9.7
BP neural network	1.69	122.5	1.69	122.5
SVR model	1.33	47.3	1.33	47.3

CONCLUSION

This paper has been submitted one kind of short-term load forecasting method based on Bayes. On the premise of every factor data test result being indicated, being method's turn to be able to think that the forecast is connected with load in synthesis, reach higher forecast accuracy, and be that one kind of effective short period load forecasts method within shorter training time.

ACKNOWLEDGMENT

The research was supported by “the Fundamental Research Funds for the Central Universities”, Fund No: 11MR62.

REFERENCES

[1] GU Zhi—hong, NIU Dong—xiao, WANG Hui—qing, Research on application of general regression neural network in short-term load forecasting, ELECTRIC POWER, vol. 39, no. 4, pp. 11-14, 2006.
[2] PAN Rui, LIU Jun-yong, NI Ya-qi, Ultra-short term forecasting for short circuit current based on load forecasting and general regression neural network, Power System Protection and Control, vol. 38, no. 18, pp. 94-99, 2010.
[3] WU Yaohua, Monthly Load Forecasting Based on General Regression Neural Network, JOURNAL OF NORTHWEST HYDROELECTRIC POWER, vol. 23, no. 4, pp. 9-12, 2007.
[4] LI Ruqi, YANG Licheng, MO Shixun, Short term electric load forecasting based on accumulated weather and ACA-GRNN, RELAY, vol. 36, no. 4, pp. 58-62, 2008.
[5] Lin Yushu, Li Jian, Zhao Cuocai. Application of Rough Generalized Regression Neural Network in Heating Load Forecasting, Electric technology, no. 12, pp. 1-19, 2007.
[6] Chen Yaowu, Wang Leyu, Long Hongyu. Short-term load forecasting with modular neural networks. Proceedings of the CSEE, vol. 21, no. 4, pp. 79-82, 2001.

[7] Pan Feng, Cheng Haozhong, Yang Jingfei, et al, Power system short-term load forecasting based on SVM. Power system Technology, vol. 28, no. 21, pp. 39~42, 2004.

[8] WU Yaohua, A new method of monthly load forecasting, Science & Technology Information, no. 30, pp. 71-72, 2007.

[9] WU Yaohua, Study on M id—and Long-term Power Load Forecasting Based on Wavelet Soft-threshold Technology, GUANGDONG ELECTRIC POWER, vol. 20, no. 12, pp. 5-8, 2007.

[10] LIU Xueqin, WU Yaohua, CUI Baohua, Application of wavelet soft-threshold de-noising and GRNN in monthly load forecasting, Power System Protection and Control, vol. 37, no. 14, pp. 59-62, 2009.

[11] GUO Bin, MENG Ling-qi, DU Yong, MA Sheng-biao, Thickness prediction of medium plate mill based on GRNN neural network, Journal of Central South University: Science and Technology, vol. 42, pp. 960-965, 2011.

[12] ZHOU Hao，ZHENG Li—gang. FAN Jian—ren. CEN Ke—fa, Application of general regression neural network in prediction of coal ash fusion temperature, Journal of Zhejiang university (Engineering Science), vol.38, no. 11, pp. 1479-1482, 2004.

[13] [WU Yao-hua, Long term load forecasting based on GM-GRNN in power system, Relay, vol. 35, no. 6, PP. 45-48, 53, 2007.

[14] YU Jian-ming, LI Meng, SHU Fei, Application of GRNN-algorithm on Load Modeling of Power System, Proceedings of the CSU-EPSA, vol. 21, no. 1, pp. 104-107, 2009.

[15] LU Ning, ZHOU Jian-zhong, HE Yao-yao, Particle swarm optimization-based neural network model for short-term load forecasting, Relay, no. 12, pp. 65-68, 2010.

[16] LIU Xue-qin1, WU Yao-hua2, CUI Bao-hua1, Short-term load forecasting model based on extended rough set, Relay, no. 5, pp. 25-28, 38, 2010.

[17] Osman, Z.H.; Awad, M.L.; Mahmoud, T.K., Neural network based approach for short-term load forecasting, Power Systems Conference and Exposition, pp. 1-8, 2009.

[18] Senjyu, T.; Takara, H.; Uezato, K.; Funabashi, T., One-hour-ahead load forecasting using neural network, Power Systems, vol. 17, pp. 113 – 118, 2002.

[19] Wei-Min Li; Jian-Wei Liu; Jia-Jin Le; Xiang-Rong Wang, The financial time series forecasting based on proposed ARMA-GRNN model, Machine Learning and Cybernetics, vol. 4, pp. 2005 – 2009, 2005.

[20] SUN Da-shuai, MA Li-xin, WANG Shou-zheng. Design of short-term load forecast systems based on the theory of complex systems, Journal of University of Shanghai For Science and Technology, 2011, 32(1): 39-43.

[21] LIAO Li, XIN Jian-hua, ZHAI Hai-qing, WEI Zhen-hua. Short Term Load Forecasting Model and the Influencing Factors, JOURNAI OF SHANGHAI JIAOTONG UNIVERSITY, 2004, 38(9): 1544-1547.

[22] ZHU Ji-ping, DAI Jun. Optimization Selection of Correlative Factors for Long -term Load Prediction of Electric Power, Computer Simulation, 2008, 25(5): 226-229.

[23] WEI Lingyun, WU Jie, LIU Yongqiang. Long term electric load forecasting based on system dynamics, Automation of Electric Power Systems, 2000, 24 (16): 44-47.

[24] XIAO Xian-yong, GE Jia, HE De-sheng. Combination Method of Mid-Long Term Load Forecasting Based on Support Vector Machine, Proceedings of the CSU-EPSA, 2008, 20(1): 84-88.

[25] LI Ru-qi, CHU Jin-sheng, XIE Lin-feng, WANG Zong-yao. Application of IAFSA-RBF Neural Network to Short-Term Load Forecasting, Proceedings of the CSU-EPSA, 2011, 23(2): 142-146.

[26] ZHOU Xiao-hua, et Al. Application of Elman Neural Network to Short-Term Load Forecasting in Rural Power System, Journal of Anhui Agricultural Sciences, 2011, 39(4): 2424-2426.

[27] WANG Lin-chuan, BAI bo, YU Feng-zhen, YUAN Ming-zhe. Short-term load forecasting based on weighted least squares support vector machine within the Bayesian evidence framework, Relay, 2011, 39(7): 44-49.

[28] ZHOU Ying, YIN Bang - de, REN Ling, BIAN Xue － fen. Study of Electricity Short- term Load Forecast Based on BP Neural Network, Electrical Measurement & Instrumentation, 2011, 48(2): 68-71.

[29] SHI Hai-bo. Power Load Forecasting Based on Principal Component Analysis and Support Vector Machine, Computer Simulation, 2010, 10: 279-282.

[30] LI Xiao-bo, LUO Mei, FENG Kai. Comparison of neural network methods for short-term load forecasting, Relay, 2007, 35(6): 49-53.

[31] LI Ru-qi, YANG Li-cheng, MO Shi-xun, SU Yuan-yuan, TANG Zhuo-zhen. Short-term electric load forecasting based on accumulated weather and ACA-GRNN, Relay, 2008, 36(4): 58-62.

[32] CHEN Chen, et. Al. Agricultural Product Quality Mining based on Bayesian Classification, Journal of Anhui Agricultural Sciences, 2011, 39(12): 7448-7449.

[33] SHEN Jin-biao, LV Yue-jin. A reasoning and diagnosis model based on rough sets and bayesian networks, 2009, 34(6): 815-818.

[34] CAI Na, ZHANG Xue-feng. Attribution Reduction of Bayesian Rough Set Model, Computer Engineering, 2007, 33(24): 46-48.

[35] HUANG Chong-zhen, LIANG Jing-guo. Predicting construction quality of a marine drilling platform based on GRNN, Journal of Harbin Engineering University, 2009, 30(3): 339-343.

[36] ZHOU Shaolong, ZHOU Feng. GRNN model for prediction of port cargo throughput based on time series, Journal of Shanghai Maritime University, 2011, 32(1): 70-73.

[37] WEN Ai-hua, LI Song. Forecast of Railway Freight Volumes Based on Generalized Regression Neural Network, Rail Way Transport and Economy, 2011, 33(2): 88-91.

[38] YU Ping-fu, LU Yu-ming, WEI Li-ping, LONG Wen-qing, SU Xiao-bo. Application of General Regression Neural Network （GGRNN） on Predicting Yield of Cassava, Southwest China Journal of Agricultural Sciences, 2009, 22(6): 1709-1713.

[39] NIU Dong-xiao, QI Jian-xun. The Research of the Load Forecasting Method of the Variable Structure Neural Network Based 0n the Fuzzy Treatment, OPERATIONS RESEARCH AND M ANAGEM ENT SCIENCE, 2001, 10(2): 86-92.

[40] LI De-zhi, HE Yong-xiu, ZHANG Yu. Urban load forecast model based on land use change and load characteristics, Electric Power, 2011, 44(1): 6-10.

[41] LIN Hui, LIU Jing, HAO Zhi-feng, ZHU Feng-feng, WU Guang-chao. Short-term load forecasting for holidays based on the similar days’ load modification, Relay, 2010, 7: 47-51.

[42] ZHANG Hong-xu , YAO Jian-gang , CAO Wei , MA Gui-xia. Ultra-short Term Load Forecasting Based on Improved

Gray Model, Proceedings of the CSU-EPSA, 2009, 21(6): 74-77.
[43] WANG Hong-wei, LIN Jian-liang, Sales Prediction Based on Improved GRNN, Computer Engineering & Science, no. 1, pp. 153-155, 2010..

Yanmei Li, Lecturer and Master, works in School of Business and Administration, North China Electric Power University, Baoding, China. Her research interests inclue Modeling and model Applications.

Jingmin Wang, professor, works in School of Business and Administration, works in North China Electric Power University, Baoding, China, who's research interests include Modeling and model Applications.