Available online at www.sciencedirect.com

SciVerse ScienceDirect

Energy Procedia 16 (2012) 1852 – 1859

Energy Procedia

2012 International Conference on Future Engery, Environment, and Materials

Power System Load Forecasting Based on Fuzzy Clustering and Gray Target Theory

HAO Jing[1], LIU Dawei[2], LI Zhenxin[3], CHEN Zilai[3], KONG Lingguo[3]

[1]College of Electrical Engineering, Northeast Dianli University, Jilin, China
[2]Electric Power Dispatching and Communication Center of ChangChun Power Supply Corporation Changchun, China
[3]College of Electrical Engineering, Northeast Dianli University, Jilin, China

Abstract

Accuracy of power system load forecasting is affected directly by existing considerable uncertainties, which are accompanied with some correlation among load factors. And this relationship between variables can be eliminated with fuzzy clustering based on historical load in a large number of observations. Carrying out the load factors on the reduction tactics, the impact factors have been classified. Then the target theory is adopted under the standard model of predictors. Then the weight coefficients between the factors category and the predictor are determined in the ways of calculating contributing degrees to various components indicators. For verification, an actual data is provided from a power grid and sorted into four types. Contributing degrees for predicting of the 4 factors are found. It is shown by result analysis that the combined method takes advantages of accuracy and efficiency in prediction.

Keywords-fuzzy clustering; grey target theory; contributing degrees; weight coefficient

1. Introduction

Power system load forecasting is important and complex for the power enterprise for many factors affect the load forecast as uncertain and strongly random trend. There is correlation between certain factors. It is difficult to identify and unify modeling to determine methods to improve accuracy of load forecasting. Therefore, load forecasting has been a research focus at domestic and abroad. Many methods were currently used for load forecasting. Early in time series, the method is represented by regression analysis which is difficult to reflect the dynamic and linear relationship between the load and other industry power, particularly on the complex power load forecasting with less effective. In recent years, some artificial intelligence method, such as principal component analysis, expert systems, fuzzy logic method, the gray prediction method, wavelet analysis and artificial neural networks and other new

 doi:10.1016/j.egypro.2012.01.284

methods have been widely used. The relevance between various factors to some extent be eliminated by Principal component analysis, but there is no clear limits in determining the cumulative contribution of the standards. A strong learning ability and generalization ability existing in neural network method, any nonlinear mapping can be realized by neural network method which is especially actively. However, the neural network training and learning have been slow by information system in a large power system and easy to fall into local optimum, thus the right result is missed. It is hard to improve real-time data processing and shorten the training time, so the problems can not be solved by the single method alone. In dealing with complicated historical data existing in the information, and strong correlation between the variables of the problem, the method is even more obscure. Fuzzy clustering is an approach with the fuzzy and the correlation between variables were eliminated and the influencing factors were classified, finally, the factors which have Strong correlation were classified as a category [1]. Load forecasting by fuzzy clustering method which weight coefficients between the factors and prediction were determined by membership in the past [3]. In this article, the author proposed a load forecasting method which has combined the fuzzy clustering theory with the gray target theory. Gray target theory as a kind of algorithm which based on fuzzy mathematics. Contributing degrees which belongs to each factor were calculated by grey target theory in predicting and weight coefficient between forecast and the factors were found by calculating its average value. The relevance and uncertainty of influencing factors can be reduced by this algorithm. The data from 1989 to 2003 of a network was analyzed and 5 years future data is predicted by combining fuzzy clustering with gray target method, the results showed that the combined method of prediction is better.

2. Theoretical basis

2.1 Fuzzy Clustering

Fuzzy Cluster Analysis is a mathematical method which base on the different characteristics，degree of closeness and similarity between objective things and classify objective things by establishing fuzzy similarity relations [1].

Assume there is a sample and every sample has *m* index (variable) by measured. The observation data X_{ij} (i=1,2,…,n, j=1,2,…,m) was obtained and observed data matrix was denoted：

$$X=(x_{ij})=\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n1} & \cdots & x_{mm} \end{bmatrix}$$

Firstly, the distance between samples and between category and category are defined, in this paper, Euclidean distance was used. The distance between i samples and j samples, is known as the Euclidean distance and Denoted by *dij*.

$$d_{ij}=\left[\sum_{k=1}^{p}\left(x_{ik}-x_{jk}\right)^2\right]^{1/2}$$

The new category is synthesized of the nearest two categories and the new distance between the new class and the other classes is calculated. Then the categories are reduced for a category every time in accordance with the guidelines of minimum distance criteria, up until all the samples are combined into a category. Specific Cluster steps are shown in Figure 1.

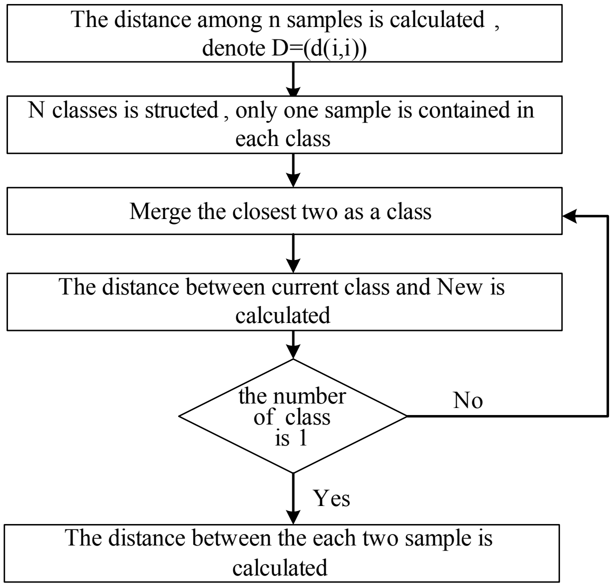

Figure 1. Cluster analysis of flow chart

2.2 Grey Target Theory

Target theory is a part of the gray theory, with basic idea that a gray target is established in the absence of standard mode [2]. The center of grey target (standard mode) is found by using target theory, and the contributing degrees are obtained by comparing original target sequence with the center of grey target. The key of target theory is that approaching degree calculation. The definition of approaching degree is that the gray correlation degree among different information space gray modes and standard modes. The geometric meaning of approaching degree is defined as the level of similarity and distance between the original target sequence and the center of a target. Since nearer the distance is, the approaching degree becomes greater.

Grey target theory which is based on the above basic principle is applied to the topic of weights coefficient, the main process as follows:

① Appropriate index is selected and standard mode is established;

② Pattern sequence is taken to unified measure trans-formation;

③The approaching degree value is calculated by formula of the approaching degree;

④ The weight of every index is obtained by combining the value of approaching degree from (3) calculated with weight strategy of approaching degree.

Definition 1 ω_i is ordered index for the multi-level sequence, P is ordered as the proposition of ω_i, $\omega_i=(\omega_i(1),\omega_i(2)\cdots,\omega_i(n))$, $\forall\ \omega_i(k)\in\omega_i$, where $k\in K=\{1,2,\cdots,n\}$, k Represent K-indexes; $i\in I=\{1,2,\cdots,m\}$, there:

① ω_i is called gray model for the P;

② If $P_k(\theta)$ is k sub-proposition of $P(\theta)$, $\omega_i(k)$ is said to the i mode k-index data;

③ $K=\{1,2,\cdots,n\}$ is index distribution;

④ $\exists\, k^*$, $s^*\in K$, $POL\omega_i(k^*)\neq POL\omega_i(s^*)$ is called multi-stage model of ω_i; all gray mode are defaulted multi-level nature when without explanation.

Definition 2 $P_k(\theta)$ is ordered k sub-proposition of $P(\theta)$, $\omega_i=(\omega_i(1), \omega_i(2),\cdots,\omega_i(n)),i\in I=\{1,2,\cdots,m\}$ as a model, $\forall\ \omega_i(k) \in\omega_i$, $k\in K=\{1,2,\cdots,n\}$, $\omega_i(k)$ subject to $P_k(\theta)$, there sequences $\omega(k)$ is composed of K index: $\omega(k)=(\omega_1(k), \omega_2(k),\cdots,\omega_m(k))$ is called k index sequence or referred to k index column.

1) Standard Mode Selection

Assume $\omega_i=\{\omega_i(1),\omega_i(2),\cdots,\omega_i(m)\}$ is the i-state model of index, $\omega(k)=\{\omega_1(k), \omega_2(k),\cdots,\omega_n(k)\}$ is the k index sequence, then the standard mode $\omega_0(k)=\{\omega_0(1),\omega_0(2),\cdots, \omega_0(n)\}$ is selected. Formula of $\omega_0(k)$ is expressed as equation (3):

$$\omega_0(k)=\begin{cases}\max\limits_i \omega_i(k), POL\omega(k)=POL\omega(\max)\\ \min\limits_i \omega_i(k), POL\omega(k)=POL\omega(\min)\end{cases} \tag{3}$$

Where, m is the number of the index state model, n is the number of selected state variables; The k-state volume which has a great value property is represented by formula of $POL(\omega(k))=POL(\max)$. The k-state volume which has a minimum value property is represented by using formula of $POL(\omega(k))=POL(\min)$.

2) Grey target transformation

If there is largely different in the amount of the two states data, then the role of small value will be covered by the large value. Therefore, it is necessary that different order of magnitude of data is unified into the same measure. Order T as transformation, if $T\omega_0=x_0$, $x_0=(x_0(1), x_0(2),\cdots, x_0(n))=(1,1,\cdots,1)$ there $T\omega_i(k)=x_i(k)$, T is called for the uniform measure transformation, the formula of $T\omega_i(k)$ is expressed as equation (4):

$$T\omega_i(k)=\frac{\min\{\omega_i(k),\omega_0(k)\}}{\max\{\omega_i(k),\omega_0(k)\}}=x_i(k) \tag{4}$$

$x_i(k)$ is the transformed value, and $x_i(k)\in[0,1]$. Where, the smaller $|\omega_i(k)-\omega_0(k)|$ is, the $x_i(k)$ closer to 1, the greater $|\omega_i(k)-\omega_0(k)|$ is, then $x_i(k)$ more faraway from 1.

3) Approaching degree calculation

Order $\Delta=\{\Delta_{0i}(k)|\ i\in I,k\in K, \Delta_{0i}(k)=|x_i(k)-x_0(k)|=|x_i(k)-1|,x_0(k)\in x_0,\text{make } x_0=T\omega_0\}$ and

$$\begin{cases}\Delta_{0i}(\max)=\max\limits_i\max\limits_k \Delta_{0i}(k)=\max\limits_i\max\limits_k |1-x_i(k)|\\ \Delta_{0i}(\min)=\min\limits_i\min\limits_k \Delta_{0i}(k)=\min\limits_i\min\limits_k |1-x_i(k)|\end{cases}$$

$\gamma(x_0(k), x_i(k))$ is said to coefficient for approaching degree and expressed as equation (5):

$$\gamma(x_0(k),x_i(k)) \\ =\frac{\min\limits_i\min\limits_k \Delta_{0i}(k)+\rho\max\limits_i\max\limits_k \Delta_{0i}(k)}{\Delta_{0i}(k)+0.5\max\limits_i\max\limits_k \Delta_{0i}(k)} \tag{5}$$

Where, ρ is judgment of policy-makers. Through repeated verification, in this paper, ρ=0.5.Then, $\gamma(x_0(k), x_i(k))$ is said to approaching degree of x_i and expressed as equation (6):

$$\gamma(x_0,x_i)=\frac{1}{n}\sum_{k=1}^{n}\gamma(x_0(k),x_i(k)) \tag{6}$$

3. Based on Fuzzy Clustering and Grey target theory prediction model

In this paper, prediction model is applied to Load Forecasting which combined the fuzzy clustering with the target theory. Because value of load is affected by many factors and some correlation is existing. All affecting factors are classified by fuzzy cluster analysis. The number of classification is determined according to European cluster between the category and category. The target theory is introduced and the maximum load is used as standard mode and the identified class is gray target transformed. The contribution degree between every category and the predictor is calculated.

Finally, approaching degree is taken as weight coefficients between the affecting factors and the predictor. The maximum load value for entire network is used as predictor and the eight electricity industry is used as affecting factors for the power load. There is some correlation between these electricity industries. For this reason, these eight industries are classified and the correlation is eliminated by the fuzzy clustering analysis. The number of categories is determined by Euclidean distance between the class and the class. The calculated approaching degree is taken as weight coefficients in predicting.

3.1 The actual power grid load data analysis and prediction

The maximum load value of a grid from year of 1989 to 2003 is taking as the underlying data, the historical data is shown in Table Ⅰ~ Table Ⅲ.

Table 1 Electricity Industry And Maximum Load Form 1994 To 1998

UNIT:BILLION KWH TEN THOUAND KW

Year / industry	1989	1990	1991	1992	1993
agricuture	23.1	24.9	23.2	25.2	24.9
industry	594.2	613.5	655.5	705.5	747.2
geology	0.3	0.3	0.4	0.5	0.6
building	4.7	4.7	5.8	7.5	9.5
traffic	13.7	14.1	14.9	15.6	16.3
business	7.0	7.6	9.0	10.6	12.0
other	22.0	23.5	26.3	29.3	32.6
urban	48.8	64.4	68.6	81.1	93.7
Maximum load	672.8	726.8	773.8	862.9	937.4

Table 2 Electricity Industry And Maximum Load Form 1994 To 1998

UNIT:BILLION KWH TEN THOUAND KW

Year / industry	1994	1995	1996	1997	1998
agriculture	25.0	23.9	25.2	27.9	26.1
industry	765.6	789.7	824.1	839.2	819.6
geology	0.9	0.8	1.0	0.8	0.7
building	10.1	10.7	10.7	9.4	8.9
traffic	17.3	15.2	18.4	17.7	16.9
business	13.2	15.1	16.4	18.6	19.0
other	34.9	37.0	40.3	43.8	45.8
urban	104.3	118.1	101.6	146.7	151.1
Maximum load	1013.6	1061.6	1088.6	1159.9	1152.4

Table 3 lectricity Industry And Maximum Load Form 1999 To 2003

UNIT:BILLION KWH TEN THOUAND KW

Year / industry	1999	2000	2001	2002	2003
agriculture	27.1	29.6	33.1	31.2	31.7
industry	836.6	880.9	888.9	922.4	1018.2
geology	0.7	0.7	0.6	0.5	0.5
building	8.6	9.7	9.8	8.9	10.8
traffic	17.5	17.9	17.8	24.4	27.2
business	22.7	27.4	31.6	35.8	39.7
other industry	51.0	56.9	57.8	61.8	69.7
urban	151.7	162.4	172.5	179.3	193.7
Maximum load	1145.9	1168.6	1177.2	1305.2	1397.5

The factors include agriculture, industry, geology, building, traffic, Business, other and urban. These eight industries are taken as affecting factors. Because the correlation is existed between eight factors, the method of fuzzy clustering and statistical software SAS are used to clustering analysis by minimum distance method. The cluster graph is shown in figure 2.

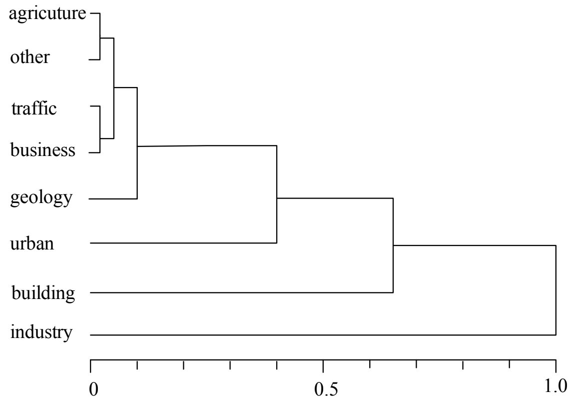

Figure 2. The result of fuzzy clustering analysis

From the figure 2, the affecting factors are obviously classified as four categories on the basis of their Euclidean distance more than 0.2. Agriculture, Geology, Traffic, Business and other industry are clustered for the first category. Urban is clustered for the second category. Building is clustered for the third category. Industry is clustered for the fourth category. These four categories are taken as affecting factors of the largest load of whole network. The predicted amount (maximum load) is taken as the standard model in the target theory and the four category of affecting factors to target transformation. Finally, target contribution degree is calculated. The results are shown in Table Ⅳ.

Table 4 Clustering Result

Category / index	The first category	The second category	The third category	The four category
Approaching degree	0.5594	0.6032	0.564	0.5219
Gray target Contribution	0.2488	0.2683	0.2509	0.2321

The Maximum load will be predicted by combining independent variable with contributing degrees in Table 1. The predicting results and the error are shown in Table Ⅴ.

Table 5. Comparison Of Predicting Results And The Largest Load Forecast Error Between Combination Algorithm And Single Algorithm From 2004 To 2008

Year / Load and error	Maximum load(MW)				
	2004	2005	2006	2007	2008
Ture value(ten thousand kw)	1535.7	1648.5	1802.6	2015.2	2137.7
Combination Algorithm(ten thousand kw)	1539.8	1653.2	1825.7	2066.3	2148.5
Relative Error(%)	0.27	0.285	2.946	2.534	0.483
Single Algorithm(ten thousand kw)	1480.2	1622.5	1725.5	1920.4	2016.1
Relative Error(%)	3.611	1.578	4.277	4.706	5.687

In order to better explain the precision of combination forecasting method, the prediction graph is drawing in Figure 3.

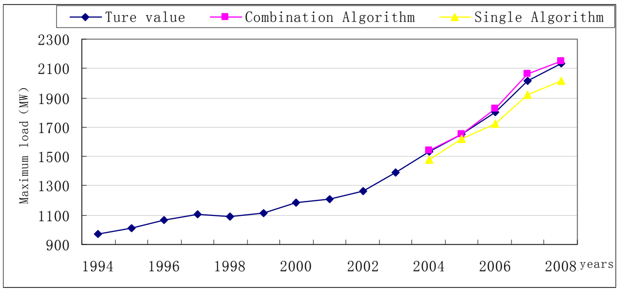

Figure 3. Comparison of two prediction methods

Form figure 3 can be clearly seen that predicting results of combination forecasting method is significantly better than a single fuzzy clustering. Moreover, the overall benefits from the power system point of view, this slightly higher estimation of the entire power system planning is better than harm.

4. Conclusion

In this article, Grey Target Theory is applied to load forecast and proposed load forecasting method which based on fuzzy cluster analysis. By cluster analysis of historical load data and Classification of affecting factors, the strong correlation between the factors is eliminated. And the predicted amount (maximum load) has been taken as the standard model in the target theory.

After the gray target transformation, contributing degrees are calculated as weighting coefficients between factors and predictor. That will not only eliminate the correlation between the independent variables, but also get rid of the past method which determined the weight coefficient by regression. The accuracy of the prediction method has been proved by maximum power grid load forecasting data that it is significantly higher than the traditional ways.

References

[1]Shyh-Jier Huang,Kuang-Rong Shih, “Application of a fuzzy model for short-term load forecast with group method of data handling enhancement[J],” Electrical Power and Energy Systems, pp. 631–638, October 2001.

[2]He xiaoqun, MultivariatStatistical Analysis[M]. Renmin University Press . Beijing China, 2008.

[3] Li jian-po, “A grey target theory based method for state assessment of power transformer[J],” vol.1 ournal of jilin University (Ennineering and Technolony Edition),pp.201~205. may 2008.

[4]Zhang yan-jie, “Short-term load forecasting method based on integrating fuzzy cluster rocongnition with statistic correlation[J]”. Power System Technology, vol. 26,pp,32~36,Nov 2002.

[5]Zhang chun-lei, “Power system short-term load forecasting based on fuzzy clustering analysis and rough sets[J],” Journal of North China Electric Power University,vol.3,pp.38~45,may 2008

[6]Yang zheng-lin, “Improved Cluster Analysis Based Ultra-short Term Load Forecasting Method[J],” Automation of Electric Power Systems,vol.29,pp.83~97. Dec 2005.

[7]Yin cheng-qun, “Load forecasting based on fuzzy clustering and grey relational analysis[J],”Journal of North China Electric Power University,Vol.33,pp.24～27,Nov 2006.

[8]Xie zhi-jian.Study of Method for Solving Weights[J]. Journal of north university of China. Vol.5,pp.313~316,may 2005.

[9]Ding Qiao-lin,PAN Xue-hua,YANG Xue-ming, “Optimal Combined Forecasting Method Used in Power Load Forecasting,” Power System Technology，Vol.32，pp.127–130，june 2008.

[10]Zhang ya-jun,Liu zhi-gang,Zhang da-bo, “A Combination Load Forecasting Model Based on Multi-Neural Networks[J],” Power System Technology, vol. 30,pp.21~24,Nov 2006.

[11]Liu yao-nian,Fu xiang-yun, “Short-Term Load Forecasting Method Based on Fuzzy Pattern Recognition and Fuzzy Cluster Theory[J],” Transactions of China Electrotechnical Society,vol.17,pp.83~95, Oct 2002.

[12]Kang chong-qing,xia qing, “Review Of Power System Load Forecasting And Its Development[J],” Automation of Electric Power Systems,vol.28,pp.1~10. Sept 2004.