The Short-term Load Forecasting of Electric System

Zhaoyuan Wang[1, a]

[1] North China Electric Power University(Baoding), Baoding 071000,China;

[a]2522635788@qq.com

Abstract. This thesis firstly analyzes some indexes as daily maximum load, daily minimum load, daily difference between peak and valley and daily load rate in two different regions. And then it makes a stepwise regression analysis on the relationship between the above indexes and various climate factors, so as to obtain equations of linear regression and regression errors in the two regions. After that, the thesis selects some key impacting indexes to analyze the influence of climate on the electric load forecasting. Finally, it makes short-term forecasting on the load of the two regions by the grey predication way, the data of which are compared with the real ones to judge the feasibility of the forecasting methods.

1.   Introduction

Short-term load forecasting, as a very important content in the running and scheduling of modern electric system, could improve the accuracy of load forecasting and is a significant means to guarantee the scientificity of the optimized decisions of electric system.

2.   An analysis on the distribution of various indexes

By disposing and analyzing the electric load values of 2014, we find out the daily maximum load and daily minimum load of each day to calculate the daily average load, based on which to work out the daily difference between peak and valley and the index of daily load rate.

We could find out by analyzing the results that, except some specific dates which are affected by some festivals, the load capacity of both the region 1 and region 2 has a periodical trend change. During the same time period, the daily maximum load capacity and daily minimum load capacity of region 2 are higher than the ones in region 1; the daily difference of peak and valley of region 2 is previously higher than region 1; and compared with region 1, the index of daily load rate of region 2 is more closer to 1, with a more obvious float.

According to the above analysis result, we calculate the variable coefficients of the daily maximum load capacity, daily minimum load capacity, daily load rate as well as the average value through the whole year in the two regions. What's more, we use MATLAB to fit the load duration curve in two regions through 2014.

After calculation, the results are shown as following:

Table 1 The variable coefficients and average values of various indexes in two regions

	Variable Coefficient			Average Value
	The daily maximum load capacity	The daily minimum load capacity	The daily load rate	The difference between peak and valley
Region 1	22.466%	23.914%	3.365%	4081.725
Region 2	20.742%	26.643%	3.501%	4456.267

Based on the above analysis, we could find that the indexes of load capacity in region 1 are more stable than region 2.

By making a linear fitting on the index curves of daily load rate in two regions all year round through Matlab, the curve fitting result of region 1 is y=-0.87x+11000 while region 2 is y=-0.92x+12000. Since |-0.87|<|-0.92|, we get to know that region 1 is comparatively more stable.

Based on these studies, the forecasting results of load in region 1 could be more accurate.

3. Text The analysis about the influence of climate of electric load

The section headings are in boldface capital and lowercase letters. Second level headings are typed as part of the succeeding paragraph (like the subsection heading of this paragraph). All manuscripts must be in English, also the table and figure texts, otherwise we cannot publish your paper. Please keep a second copy of your manuscript in your office. When receiving the paper, we assume that the corresponding authors grant us the copyright to use the paper for the book or journal in question. When receiving the paper, we assume that the corresponding authors grant us the copyright to use the paper for the book or journal in question. When receiving the paper, we assume that the corresponding authors grant us the copyright to use.

Firstly, assume that the five factors as the highest temperature, minimum temperature, average temperature, relative humidity and rainfall capacity have some certain effects on the daily maximum load, daily minimum load and daily average load through the whole year in both region 1 and region 2.

By making a stepwise linear regression on the daily maximum load, daily minimum load and daily average load of the whole year in region 1 and region 2 through the software of SPSS[1], we find that all the Sig of F value are < 0.001. Thus, the assumption of each model that the regression coefficient becoming 0 fails but each equation is significant; with the increase of regression equations, the correction of R2 rises gradually and the fitting turns to be very good. Finally, after studying the regression coefficients, we find that every regression equation is meaningful.

Therefore, the regression equation of daily maximum load capacity in region 1 could be worked out as[2]:

$$\hat{y}_{11} = 165.409x_2 - 17.627x_4 + 9.021x_5 + 5538.371 \tag{1}$$

The standard deviation of the forecasting value of $\hat{y}_{11}$ could be estimated by residual mean square: $S_{\hat{Y}} = \sqrt{1590868.149} = 1261.296$. Since climate factors are not decisive factors but only influence factors for load capacity, the value of standard deviation is in the scope of reasonable values.

Next, we use the same method to analyze the influence of climatic factors on the daily minimum load capacity and average daily load capacity in region 1 and on the various load capacities in region 2. The calculation results are as following:

The regression equations obtained through calculation are shown below:

$$\begin{cases} \hat{y}_{12} = 102.993x_2 - 8.538x_4 + 5.299x_5 + 2812.193 \\ \hat{y}_{13} = 134.251x_2 - 13.55x_4 + 7.266x_5 + 4341.397 \\ \hat{y}_{21} = 214.963x_3 - 18.431x_4 + 13.094x_5 + 5930.095 \\ \hat{y}_{22} = 138.968x_3 - 17.033x_4 + 9.512x_5 + 20.169x_2 + 2819.302 \\ \hat{y}_{23} = 175.11x_3 - 19.724x_4 + 9.544x_5 + 20.632x_2 + 4669.873 \end{cases} \tag{2}$$

The standard deviation of various predicted values is shown in table 2.

Table 2 The standard deviation of various predicted values

	$\hat{y}_{12}$	$\hat{y}_{13}$	$\hat{y}_{21}$	$\hat{y}_{22}$	$\hat{y}_{23}$
The standard deviation of various predicted values	775.129	1006.477	1299.412	818.203	1053.2206

By analyzing the variable coefficients of various regression equations, we summarize that the greater the values are, the greater the impact on the dependent variables. Thus, we could conclude the impact factors of climate is the daily maximum load, daily minimum load and average load.

Thus, if using climatic factors to improve the accuracy of load forecasting, the region 1 will give high priority to the lowest temperature in various climatic factors, while the region 2 to the average temperature.

4. The electric short-term load forecasting based on grey models

According to the known composite data, we establish the grey GM (1, 1) prediction model[3] to predict the electric load of the two regions from January 11th to January 17th, 2015, totally 7 days. The modeling steps are as following:

Assume the original time series of electric load is as below:

$$X^{(0)} = X^{(0)}(1), X^{(0)}(2), \cdots, X^{(0)}(n) \tag{3}$$

And then , the ascending series after accumulation would be:

$$X^{(1)} = X^{(1)}(1), X^{(1)}(2), \cdots, X^{(1)}(n) \tag{4}$$

Among them, each value in the $X^{(1)}$ series is the sum of the corresponding ordinals in the series of $X^{(0)}$. According to the grey model theory, the differential equation could be built as:

$$\frac{dX^{(1)}}{dt} + aX^{(1)} = u \tag{5}$$

Among them, the "u" and the "a" in the model are the parameters to be evaluated.

After solving the above differential equation, the formula could be reached as:

$$X(k+1)^{(1)} = [X(1)^{(0)} - u/a]e^{-ak} + u/a \tag{6}$$

By calculating the above formula, we will get the $X^{(1)}$ series, which will generate another formula as below after accumulation:

$$X(k)^{(0)} = X(k)^{(1)} - X(k-1)^{(1)} \tag{7}$$

By continuous subtraction, the data could be restored, producing predicted values of electric load.

The next step is to conduct deviation value tests on stepwise ratio. At first, we calculate the $\lambda(k)$ according to the reference data as $x^{(0)}(k-1), x^{(0)}(k)$, and then, work out the corresponding stepwise ratio deviations based on the coefficient "a", $\rho(k) = 1 - (\frac{1-0.5a}{1+0.5a})\lambda(k)$. After calculation, we find that $\rho(k)$ <0.1 is suitable for most of the data, which is considered to reach a higher requirement; the remaining data is shown as 0.1< $\rho(k)$ <0.2, meeting the general requirement. It is visible that the grey model is a good way to predict the electric load, which fits the facts a lot. So, we use the grey model to predict the electric load at each time during 11th January, 2015 to 17th January, 2015[4].The forecasting results are shown as figure1 and figure 2.
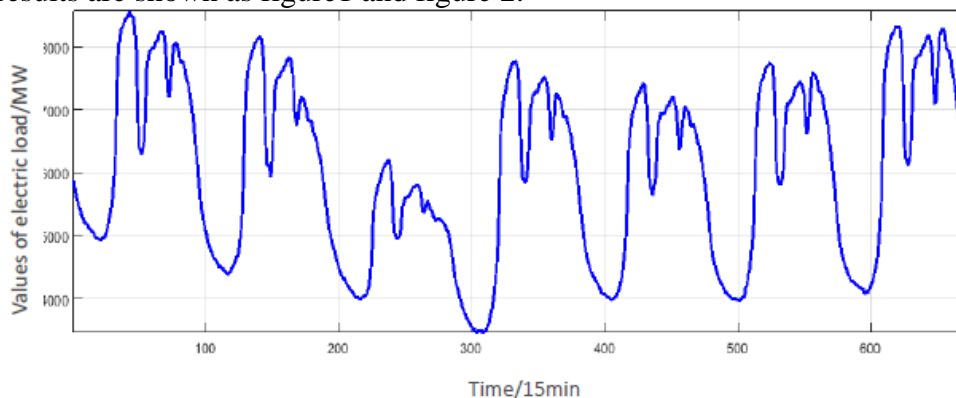

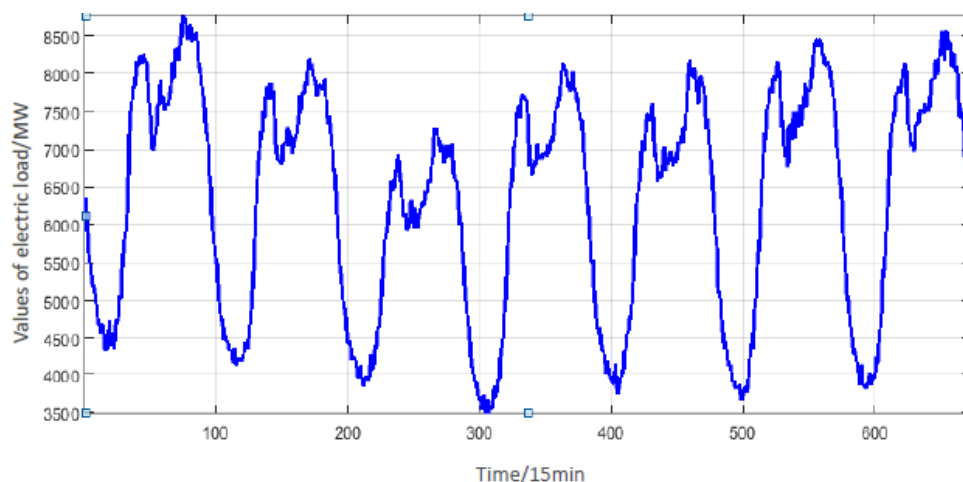Fig.1 The predicted values of electric load in region 1

Fig.2 The predicted values of electric load in region 2

Through calculation, the average value of the error percentage between the real values and the predicted values of electric load in region 1 is 8.849923%, while the average value of the error percentage between the real values and the predicted values of electric load in region 2 is 6.625552%. Therefore, we could infer that the accuracy of the average predicted results of region 1 in 2015 is 91.150077%, while the one in region 2 is 93.374448%. Thus, it is visible that the predicted results of load in region 2 are more accurate.

5. Summary

Through calculation, the average value of the error percentage between the real values and the predicted values of electric load in region 1 is 8.849923%, while the average value of the error percentage between the real values and the predicted values of electric load in region 2 is 6.625552%. Therefore, we could infer that the accuracy of the average predicted results of region 1 in 2015 is 91.150077%, while the one in region 2 is 93.374448%. Thus, it is visible that the predicted results of load in region 2 are more accurate.

References

[1]. Peng Peng, Jiahong Peng. Research on the prediction of power load based on Multiple linear regression model [J] Journal of Safety Science and Technology,2011,p.158-161

[2]. Shoukui Si, Zhaoliang Sun. Mathematical modeling algorithm and application [M] Nation Defense Industry Press, 2015,p.150-153

[3]. Zhiming Zhang. Research on short-term electrical load forcasting based on grey theory [D] Master Degree Dissertation of Hunan University,2009,p.24-27

[4]. Zhuo Jinwu, Wei Yongsheng, Qin Jian, Li Biwen. The application of MATLAB in mathematical modeling [M] Beijing University of Aeronautics and Astronautics Press, 2011,p.266-270