# Study on the Relationship between Meteorological Factors and Power Load

## Chenchen,Huang

Department of Economics and Management, North China Electric Power University,

Baoding 071000, China

973118497@qq.com

**Abstract:** In this paper, a multiple linear regression model was established, and the regression equations were established for daily maximum load, daily minimum load, daily average load and each meteorological factor. Next, the parameters of the model estimation, statistical estimation and hypothesis testing. According to the results, six regressions were obtained. According to the results of F test, T test and multiple correlation coefficient R, the significant results of each equation were obtained. The significant results and fitting of the meteorological factors to the load indexes Goodness. According to the data, the regression error analysis is carried out, and the meteorological factors which can improve the precision of load forecasting are put forward.

## 1. Introduction

Short-term load forecasting is the basis of power system operation and analysis, which is of great significance for unit combination, economic dispatching and safety check. Improve the accuracy of load forecasting is an important means to ensure the scientific decision-making of power system optimization. In the modern electric power system, the load of the electric power load is various, the proportion of the load affected by meteorological conditions such as air conditioner is increasing continuously, and the influence of meteorological factors (such as temperature, humidity, rainfall) on power system load becomes more prominent. Considering the meteorological factors becomes one of the main means to improve the load forecasting precision. In this paper, the daily maximum load, daily minimum load, daily average load and the relationship between the various meteorological factors regression analysis, regression error analysis. So in many meteorological factors, to find the factors that can improve the load forecast accuracy.

## 2. Selection of Samples

In this paper, 2016 electric Cup A title, for example, to study how to find meteorological factors and the relationship between the load. The power load data (1 sampling point every 15 minutes, 96 points per day, dimension MW) for known area A and area B from January 1, 2012 to December 31, 2014 and meteorological data (daily maximum temperature , Daily minimum temperature, daily mean temperature, daily relative humidity and daily rainfall).

## 3. Multiple Linear Regression Model

The following analysis based on the maximum temperature, daily minimum temperature, daily average temperature, daily relative humidity and daily rainfall were $x_1, x_2, \cdots, x_5$. The daily maximum load, the daily minimum load, and the daily average load of area A and area B are $y_{11}, y_{12}, y_{13}$ and $y_{21}, y_{22}, y_{23}$, respectively.

The model of multivariate linear regression[1] is

$$\begin{cases} y = \beta_{i0} + \beta_{i1}x_1 + \cdots + \beta_{i5}x_5 + \varepsilon, \\ \qquad \varepsilon \sim N(0, \sigma^2), \end{cases} \tag{1}$$

The parameter $\beta_0$, $\beta_1$, $\cdots$, $\beta_5$ in the model is estimated using the least squares method, so that when the square of the error is minimized, $\hat{\beta}$ is obtained. The $\hat{\beta}$ generation back to the original model to get the estimated value of $y$, that is

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \cdots + \hat{\beta}_5 x_5. \tag{2}$$

The fitting values of the data are

$$\hat{b}_i = \hat{\beta}_0 + \hat{\beta}_1 a_{i1} + \cdots + \hat{\beta}_5 a_{i5}, i = 1, \cdots, n. \tag{3}$$

The fitting error $e = Y - \hat{Y}$ is called the residual, and $Q = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n}(b_i - \hat{b}_i)^2$ is the residual sum of squares.

## 4. Model Solving and Testing

Matlab software was used to get the result.

The regression equation for area A is

$$y_{11} = 8690.5 - 0.0577x_1 - 0.0002x_2 + 0.0052x_3 - 0.2174x_4 + 0.0072x_5$$

$$y_{12} = 4937 - 0.0202x_1 - 0.0005x_2 + 0.0034x_3 - 0.1797x_4 + 0.0057x_5$$

$$y_{13} = 7001.5 - 0.0498x_1 + 0.0040x_3 - 0.1915x_4 + 0.0064x_5$$

The regression equation for area B is

$$y_{21} = 1070.9 - 0.0233x_1 + 0.0007x_2 - 0.0001x_3 - 0.0231x_4 + 0.0012x_5$$

$$y_{22} = 6086.6 - 1729x_1 + 0.0051x_2 - 0.0010x_3 - 0.1773x_4 + 0.0089x_5$$

$$y_{23} = 8749.5 - 0.288x_1 + 0.0068x_2 - 0.0014x_3 - 0.1897x_4 + 0.0100x_5$$

The results of the model test[2] are as follows:

Table 1 Model testing results of area A

|  | $R$ | $F$ | $T_1$ | $T_2$ | $T_3$ | $T_4$ | $T_5$ |
|---|---|---|---|---|---|---|---|
| $y_{11}$ | 0.6558 | 163.9273 | -0.4871 | -0.0296 | 0.4238 | -1.8599 | 1.1272 |
| $y_{12}$ | 0.7281 | 245.1232 | 0.1064 | 0.0191 | 0.2752 | 8.6795 | 2.8677 |
| $y_{13}$ | 0.6858 | 192.8936 | -0.5627 | 0.0028 | 0.4372 | -2.1928 | 1.3273 |

Table 2 Model testing results of area B

|  | $R$ | $F$ | $T_1$ | $T_2$ | $T_3$ | $T_4$ | $T_5$ |
|---|---|---|---|---|---|---|---|
| $y_{21}$ | 0.7847 | 347.4320 | -3.5368 | 2.3873 | -0.1995 | -3.5270 | 4.7021 |
| $y_{22}$ | 0.8421 | 528.6054 | -4.4120 | 3.0524 | -0.3331 | -4.5579 | 6.0263 |
| $y_{23}$ | 0.8219 | 451.4146 | -4.3970 | 3.0342 | -0.3501 | -3.6867 | 5.1683 |

For the area A: As can be seen from the Table 1, the linear correlation of $y_{12}$ is more stronger, that is, there is a more significant linear correlation between the five meteorological factors and daily minimum load. The five meteorological factors explained 65.58% of the daily maximum load change, 72.81% of the daily minimum load change, and 68.58% of the daily average load change. The linearity of $y_{12}$ in the three equations is the most significant. Obviously, $b_4$ is the most significant factor of $y_{11}$, $y_{12}$ and $y_{13}$, that is, relative humidity contributes the most to the equation, and the

relative humidity is negatively correlated with daily maximum load and daily average load. It is worth noting that the correlation between the minimum temperature and the three indicators is almost nonexistent.

For the area B: As can be seen from the Table 2, the linear correlation of $y_{22}$ is stronger, that is, there is a more significant linear correlation between the five meteorological factors and daily minimum load. Five meteorological factors explained 78.47% of daily maximum load change, 84.21% of daily minimum load change, and 82.19% of daily average load change. The linear relationship of $y_{22}$ in the three equations is the most significant. Obviously, $b_1$ is the most significant factor of $y_{21}$ , $y_{22}$ ,and $y_{23}$, that is, rainfall contributes the most to the equation, and the relative humidity is positively correlated with daily maximum load, daily minimum load and daily average load.

The residual square sum of area B is smaller than that of area A, and the fitting effect is better.

From the above analysis, it can be seen that the load of the area B is greatly affected by meteorological factors. The daily minimum load is more correlated with the meteorological factors than the other two. The relative humidity had a greater impact on the load in the area A, and the rainfall had a greater impact on the load in the area B. Therefore, if meteorological factors are to be used to improve the accuracy of load forecasting, relative humidity and rainfall are preferentially recommended among various meteorological factors.

## 5. Conclusions

In this paper, a multiple linear regression model was established, and the regression equations were established for daily maximum load, daily minimum load, daily average load and each meteorological factor. Next, the parameters of the model estimation, statistical estimation and hypothesis testing. According to the results, we obtained six regression equations, according to the F test, T test and the results of multiple correlation coefficient R regression analysis, found that: the area B is greatly affected by meteorological factors, and the impact of rainfall on area B is bigger than other factors. The relative humidity has a greater impact on the load of the area A, and the daily minimum load is the most relevant to the meteorological factors. It can be concluded that the relative humidity and rainfall can improve the load forecasting precision.

## 6. References

[1] Bahovec, Jakšić V, Saša. Multiple linear regression.[J]. International Journal of Research in Computer Science, 2014, 66(2):45-72.
[2] Quinino R C, Reis E A, Bessegato L F. Using the coefficient of determination R 2, to test the significance of multiple linear regression[J]. Teaching Statistics, 2013, 35(2):84–88.