The Characteristic Analysis and Forecasting of Mid-Long Term Load Based on Spatial Autoregressive Model

Guowei Cai, Deyou Yang, Ying Jiao, Chao Pan

Abstract--**Using spatial autoregressive model, the spatial characteristic between power demand and GDP is analyzed in this paper. And the combination model of forecasting in which the spatial characteristic between power requirement was considered is established. Simulations results show the distinct spatial dependence between power requirement and GDP, and the dependence between power requirementv and GDP is very strongly in space. The forecasting results show the error of the combination model of forecasting which was established in this paper is very small. Besides the combination model proposed in this paper is a effective forecasting method because of its strong adaptability.**

Keywords--**power demand; load characteristic; spatial autoregressive model; Moran *I***

I. Introduction

THE study on the characteristic and forecasting of mid-long term load is very important for power system. In the electric power market, do a good job on the characteristic analysis and forceasting of load, in particular the mid-long term load, is directly related to the economic interests of power grid[1].

The methods[2] which were alread proposed in the past aimed at the time correlation between power demond and its factors, and the study results showed the correlation between power demond and its factors was very strongly. But with the development of theoretical research, more and more papers noted the space dependency of variables[3,4]. The study on the spatial econometrics indicated that the economic and cultural development is closely related to the geographical proximity. The research interest of spatial autoregressive model lies in exploring whether there is the spatial correlation between the variables. The study on the spatial dependency has attracted wide attention because the reasonable and meaningful values can be calculated by employing the spatial autoregressive model.

The spatial dependency between power demand and GDP of 30 provinces in China by using the data of statistical yearbook. The results of Moran *I* shows the strongly spatial dependency between power demand and GDP of 30 provinces. The combination model with spatial correlation was established by employing the spatial autoregressive model, Grey model and BP neural networks.

II. Spatial Autoregressive Model

A. Spatial Autoregressive Model[5,6]

A general specification for a spatial regression model is the specification combining a spatially autoregressive dependent variable among the set of explanatory variables and spatially autoregressive disturbances. For a first-order process, the model is given by:

$$\begin{aligned} & y = \rho W_1 y + X\beta + \mu \\ & \mu = \lambda W_2 \mu + \varepsilon \\ & \varepsilon \sim N(0, \sigma^2, I_n) \end{aligned} \tag{1}$$

Where y is the $(n\times 1)$ vector with observations on the dependent variable, X the $(n\times k)$design matrix containing the explanatory variables, β the $(k\times 1)$ vector with parameters, ρ the scalar spatially autoregressive parameter, λ the scalar

Guowei Cai is with the School of Electrical Engineering, Northeast Dianli University, Jilin 132012, Jilin Province, China (e-mail: caigw@ mail.nedu.edu.cn)

Deyou Yang is with the School of Electrical Engineering, Northeast Dianli University, Jilin 132012, Jilin Province, China (e-mail: dy0101232@163.com)

Ying Jiao is with the School of Business Administration, Northeast Dianli University, Jilin 132012, Jilin Province, China (e-mail: jiaoy821225@163.com)

Chao Pan is with the School of Electrical Engineering, Northeast Dianli University, Jilin 132012, Jilin Province, China

spatial autoregressive disturbance parameter, and μ is an $(n\times 1)$ independently and identically distributed vector of error terms. W_1 and W_2 are the n-by-n neighborhood matrix that accounts for the spatial relationships (dependencies) among the spatial data.

From the general model in (1) we can derive special models by imposing restrictions.

Case 1:

For example, setting X=0 and W_2=0 produces a first-order spatial autoregressive model shown in (2).

$$\begin{aligned} & y=\rho W_1 y+\varepsilon \\ & \varepsilon \sim N(0,\sigma^2,I_n) \end{aligned} \tag{2}$$

This model attempts to explain variation in y as a linear combination of contiguous or neighboring units with no other explanatory variables.

Case 2:

Setting W_2=0 produces a mixed regressive-spatial autoregressive model shown in (3). This model is analogous to the lagged dependent variable model in time series. Here we have additional explanatory variables in the matrix X to explain variation in y over the spatial sample of observations.

$$\begin{aligned} & y=\rho W_1 y+X\beta+\varepsilon \\ & \varepsilon \sim N(0,\sigma^2,I_n) \end{aligned} \tag{3}$$

Case 3:

Letting W_1=0 results in a regression model with spatial autocorrelation in the disturbances shown in (4).

$$\begin{aligned} & y=X\beta+\mu \\ & \mu=\lambda W_2\mu+\varepsilon \\ & \varepsilon \sim N(0,\sigma^2,I_n) \end{aligned} \tag{4}$$

Case 4:

A related model known as the spatial Durbin model is shown in (5), where a “spatial lag” of the dependent variable as well as a spatial lag of the explanatory variables matrix X are added to a traditional least-squares model.

$$\begin{aligned} & y=\rho W_1 y+X\beta+\varepsilon \\ & \varepsilon \sim N(0,\sigma^2,I_n) \end{aligned} \tag{5}$$

B. Parameters Estimation for the Spatial Autoregression Model

Let $A=I-\rho W_1$ $B=I-\lambda W_2$, model (1) can be written as:

$$\begin{cases} AY-X\beta=\mu \\ B\mu=\varepsilon \end{cases} \tag{6}$$

Assuming ε is a vector of normal disturbance terms, the likelihood function shown as:

$$L(\theta)=-\frac{n}{2}\log(2\pi\sigma^2)+\log|B|+\log|A|-\frac{v^T v}{2\sigma^2} \tag{7}$$

Where $v=B(AY-X\beta)$, $|I-\rho W_1|>0$ $|I-\lambda W_2|>0$.

The parameters can be estimated by maximizing the model (7).

C. Moran's Index[7]

The best-known spatial association statistics for ordinal and interval data are Geary's c and Moran's I. These statistics are also special cases of the general cross–product statistic.

The univariate Moran's I statistic is given by:

$$I=\frac{\sum_{i=1}^{n}\sum_{j=1}^{n}W_{ij}(y_i-\bar{y})(y_j-\bar{y})}{S^2\sum_{i=1}^{n}\sum_{j=1}^{n}W_{ij}} \tag{8}$$

Where $S^2=\frac{1}{n}\sum_{i=1}^{n}(y_i-\bar{y})$, $\bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i$.

It is common practice to interpret Moran's I as a correlation coefficient, although its value is strictly speaking not restricted to the [-1,+1] interval. High positive values signal the occurrence of similar attribute values over space (either high or low values), and hence spatial clustering. Negative values indicate the joint occurrence of high and low attribute values in nearby locations. A value close to the expected value of Moran's I in the absence of spatial correlation can be taken as evidence of a random allocation of attribute values over space.

III. THE COMBINATION MODEL WITH SPATIAL CORRELATION FOR LOAD FORECASTING

A. Combination Model forLoad Forecasting

Assuming m forecasting models were used in a problem, the combination model can be shown as:

$$\begin{cases} \hat{y} = \sum_{i=1}^{m} \omega_i f_i \\ \sum_{i=1}^{m} \omega_i = 1 \\ \omega_i \geq 0 \quad i = 1, 2, \cdots\cdots m \end{cases} \tag{9}$$

Where f_i is the forecasting value of the i^{th} model.

Assuming y is the real value, the absolute error of the combination model can be written as:

$$e = \left| y - \hat{y} \right| \tag{10}$$

How to choose the weights is the key of the combination model. And the combination model aims at minimising the absolute error, so it can be changed to a optimization model:

$$\begin{cases} \min J = e \\ \hat{y} = \sum_{i=1}^{m} \omega_i f_i \\ \sum_{i=1}^{m} \omega_i = 1 \\ \omega_i \geq 0 \quad i = 1, 2, \cdots\cdots m \end{cases} \tag{11}$$

B. The Combination Model with Spatial Correlation forLoad Forecasting

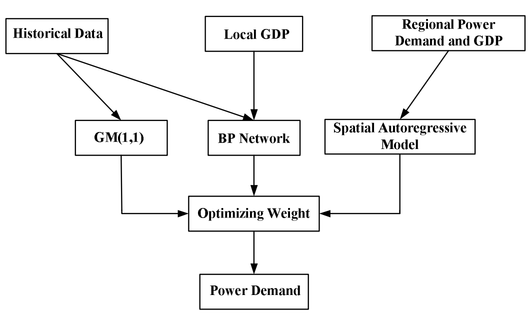


Fig.1 Combination model frame

The Grey Model[8] has the advantage in dealing with the data with tendency, and the internal pattern of power system load can be seized by employing Grey Model. The BP natural network[9] can be used to solving the nonlinear mapping. The factor impacting on the power system load can be regarded as one input of the BP natural network in forecasting. In this paper, the combination model with spatial correlation for load forecasting was established by employing the spatial autoregressive model, Grey model and BP neural networks. The combination model is shown in Fig.1.

Compare with the conventional combination model, the spatial correlation can be taken into account in the proposed combination model by employing the spatial autoregressive model. The weight was calculated by using the Matlab optimization toolbox.

IV. EXAMPLE

A. Spatial Correlation Analysis Between Power System Demand and GDP

The spatial correlation of economy is becoming more and more strong with high-speed development of economy and economic cooperation between every province in China. The time series analysis method was used to study the correlation between power demand and economy from the perspective of time, and did not consider the spatial correlation.

Using the data of statistical yearbook and space structure of 30 provinces in China, the spatial correlation between power demand and GDP was calculated by employing the spatially autoregressive model.

The results calculated by using the spatially autoregressive model are shown in Tab.1. R^2 is the multiple correlation coefficient, the value of R^2 shows the relevance between independent variable and dependent variable. The value in parentheses is significant test statistic.

Tab.1 Estimating results

Model	Constant	β	ρ	λ	R^2
Regression Model	208.45 (0.001)	0.082 (0.008)			0.4571
Residual Spatially Autoregressive Model	236.21 (0.00)	0.061 (0.00)		0.6224 (0.00)	0.9112
Hybrid Spatially Autoregressive Model	176.32 (0.00)	0.053 (0.001)	0.4681 (0.00)		0.8252

In Tab.1, the value of p (significant test statistic) is less than 0.5, so the estimated parameters of regression model and spatially autoregressive model are significant. The multiple correlation coefficient of spatially autoregressive model are bigger than the regression model. The value of the multiple correlation coefficient shows that the spatial correlation between power demand and GDP of 30 provinces in China is very strong.

At the same time, the Moran scatterplot which is shown in Fig.2 between power demand and GDP wase plotted.

The Fig.2 in which the data points not only were collected in the first quadrant and the third quadrant but also the line was diagonal shows the positive correlation between power demand and GDP is very strong in space. And the value of Moran I (Moran's I=0.3981) shows the positive correlation between power demand and GDP is very strong in space also.

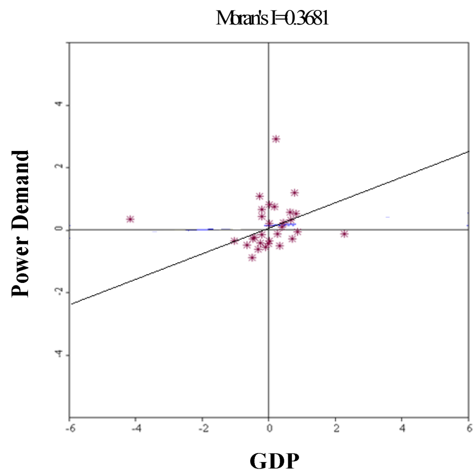


Fig.2 Moran scatterplot of power requirement and GDP

B. The Study on The Power Demand Forecasting

Tab.2 Data for forecasting

Year	2001	2002	2003	2004	2005
Power Demand (10^9kwh)	295.08	306.27	338.7	371.8	378.22

The results which was analized by employing the spatially autoregressive model shown that the power demand influenced not only by the local GDP but also by the GDP of the neighboring area. So the combination forecasting model with spatial correlation was established by employing the spatial autoregressive model, Grey model and BP neural networks.

Tab.3 Forecasting results

Method		Year	2006	2007
		Actual Value	412.26	466
Model Ⅰ	Forecasted Facts	Forecasted Value	397.83	447.83
		Error (%)	3.5	3.9
	Actul Facts	Forecasted Value	401.13	451.09
		Error (%)	2.7	3.2
Model Ⅱ	Forecasted Facts	Forecasted Value	433.7	448.84
		Error (%)	5.2	5.4
	Actul Facts	Forecasted Value	431.22	443.17
		Error (%)	4.6	4.9

Tab.3 shows the forecasted results, and the Model Ⅰ expresses the combination forecasting model proposed in this paper and Model Ⅱ expresses the combination forecasting model without spatial correlation. Fig.3 is the error char.

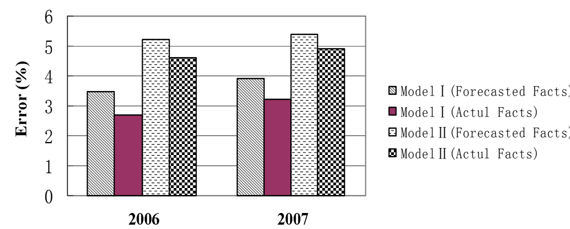


Fig.3 Error char

It is clearly seen that the accuracy of the two models with actul facts is better than the two methods with forecasted facts from Tab.3 and Fig.3. In real process, the facts also are the unknown quantity. And the accuracy of the proposed model with forecasted facts is better than the model without spatial correlation but with actul facts. The forecasted results show that the propsed model satisfies the need of procreative department and managing department.

V. CONCLUSION

In this paper, the spatial correlation between power demand and GDP was analized by employing the spatially autoregressive model. And the combination forecasting model with spatial correlation was established by employing the spatial autoregressive model, Grey model and BP neural networks. The results of example indicated:

1. The results calculated by using the spatially autoregressive model shown that the spatial correlation between power demand and GDP was very strong;
2. Moran I and Moran scatterplot shown the stronger positive spatial correlation between power demand and GDP was very strong;
3. The error of the proposed model with spatial correlation was lesser correspondingly, and the proposed forceasting model is a effective method for mid-long term load forecasting.

VI. REFERENCES

[1] Xiaodong Niu, “Power System Load Forecasting Technology and Its Applaction,” *China Electric Power Press*, 1998

[2] Chongqing Kang, "Power System Load Forecasting T," *China Electric Power Press*, 2007

[3] LI Xuying，CHEN Hongmin. Spatial Autoregression Model on the Income of Residents and City Economy. *Systems Engineering-Theroy Methodology Application*, Vol. 10 ,pp. 12-16, 2005

[4] L I Xuying. Spatial econometr ic models for data with spatial dependence. *Jouranl of Shang Hai Maritime University*, Vol..27, pp. 70-74, 2006

[5] Anselinl.Spatial Econometrics: Methods and Models.*Kluwer-Academic*, Dordrecht, 1988.

[6] Pace, R.K and R.Barry . Sparse spatial autoregressions.*S tatistics & Probability Letters,* Vol.33, pp. 291 -297, 1997

[7] Moran, P.A.P．A test for the serial dependence of residuals．*Biom etrika,*, Vol.37, pp. 178 – 181, 1950

[8] Julong Deng., "Grey System Theory," *Huazhong Unoversity of Science and Technology Press,* 2002

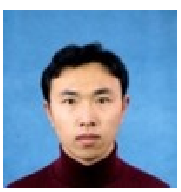
[9] Hagan M T，Demuth H B，Beale M H．Neural network design．*Beijing：China Machine Press*，2002

[10] Mohan Saini L，Kumar Soni M．Artificial neural network-based peak load forecasting using conjugate gradient methods．*IEEE Trans on Power Systems*，Vol.17, pp. 907-912, 2002

[11] Gu Jie，Shen Gang，Xu Guanghu．Study on the improved method for mid-and long-term load forecasting of power system．*Electric Power Automation Equipment*，Vol.22, pp. 1-4, 2002.

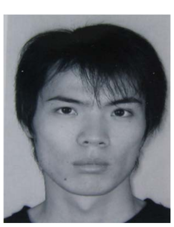
VII. BIOGRAPHIES

Guowei Cai was born in Jilin Province in Peoples' Republic of China, in 1968. He received PH.D in electrical power system and its automation from Institute of Harbin Technology in 1999. His research interests are analysis on power system transient stability and electrical power markets.



Deyou Yang was born in Heilongjiang Province in Peoples' Republic of China, in 1983. He received M.Sc. degree in electrical power system and its automation from Northeast Dianli University in 2009. His research interests are analysis on power system transient stability and power system load forecasting.

Ying JIao was born in Liaonin Province in Peoples' Republic of China, in 1982. He will receive M.Sc. degree in technical economy and management from Northeast Dianli University in 2009. His research interests is electric power market



Pan Chao was born in Henan Province in Peoples' Republic of China, in 1981. He received M.Sc. degree in electrical power system and its automation from Northeast Dianli University in 2007. His research interests are analysis on power system transient stability and faults diagnosis.