2010 3rd International Conference on Computer and Electrical Engineering (ICCEE 2010)
IPCSIT vol. 53 (2012) © (2012) IACSIT Press, Singapore
DOI: 10.7763/IPCSIT.2012.V53.No.2.11

The Optimization of Gray Model Applied to Super Short-term Load Forecasting

Jiaxing Zhai+, Shaotao Dai and Liye Xiao

Institute of Electrical Engineering, Chinese Academy of Sciences

Beijing, P. R. China

Abstract—Gray model theory is a widely used technique in load forecasting. However, there are some problems when classic gray model is used in super short-term load forecasting because of the theory itself and external factors. To solve these problems and improve forecasting accuracy, an improved gray model is introduced in this paper to be applied on super short-term electric load forecasting. In this improved model, original data is analyzed and processed. Inner structure is also optimized and residual correction is introduced. By computing the correlation of actual data and different prediction results, the best prediction program is chosen to obtain smaller errors and better results. The forecasting algorithm can estimate model parameters, meet the requirement of dynamic power load and overcome random disturbances. Example analysis and simulation show that the forecasting error is about 2 percent when used in super short-term load forecasting. Compared with conventional gray model methods, the proposed scheme has the characteristic of simple computation, high forecasting precision and good applicability.

Keywords-super short-term electric load forecasting; gray model; parameter modification; correlation analysis; posteriori error checking; residul correction.

1. Introduction

Electric load forecasting is an important basic technology in power market. It has attracted widespread attention these years for it’s one of the key technologies to build automatic distribution network in smart grid. To a grid, the improvements of the safety of operation, economy and power quality depend on accurate electric load forecasting [1]. In recent years renewable energy and new energy get rapid development in our country’s energy structure, which contains wind energy, solar energy and so on. These clean energy have the characteristic of transient change with meteorological outside factors. So super short-term electric load forecasting should be executed to ensure the dynamic balance to maintain grid frequency, energy balance and stability between the end side with the generation side of grid. The research of super short-term load forecasting has become an important research topic in recent years.

Compared with long-term, middle-term and short-term load forecasting, super short-term load forecasting has the features of less historical load data, short prediction interval, volatile and directly affected by external environmental factors. Power load distribution is complex and can not find a good distribution discipline [2]. Thus, the traditional load forecasting theory can not achieve the required precision when applied to super short-term load forecasting.

Taking account of the problems for the traditional gray model when applied to super short-term load forecasting, this paper presents a series of improvements on the gray model method, which contains the introduction of load data processing, eliminating abnormal data to meet the requirements of model [3]. To

+ Corresponding author.
E-mail address: zhaijiaxing@mail.iee.ac.cn

optimist the internal of model, gray parameters is introduced to improve model structure. The latest data of actual load and predicted data are compared to implement the inspection and treatment on the residual data, for the compensation of predicted results.

2. Building process of gary model

The actual power system is a gray system [4-5]. Gray electric load forecasting method is the application of gray system theory, through processing and modeling of the initial data, to find the discipline of development to predict the future grid state. The gray prediction method is essentially a cumulative process. Use an exponentially curve to fit the cumulative data and get the forecasting results by a reduction process [6]. In the gray load forecasting models, GM(1,1) model is widely used.

Suppose this is the minutes-level historic load data sequence as follows:

$$x^{(0)}=[x^{(0)}(1),x^{(0)}(2),\ldots,x^{(0)}(n)] \tag{1}$$

After accumulation process, the first order accumulated sequence is generated:

$$x^{(1)}=[x^{(1)}(1),x^{(1)}(2),\ldots,x^{(1)}(n)] \tag{2}$$

in which $x^{(1)}(k)=\sum_{i=1}^{k}x^{(0)}(i)$.

As the sequence $x^{(1)}(k)$ follows the exponential growth law, while the first-order differential equations is exactly the form of exponential growth. So the sequence $x^{(1)}(k)$ satisfies the following first order differential equation model, also known as bleaching differential equation. The corresponding differential equation is:

$$\frac{dx^{(1)}}{dt}+ax^{(1)}=u \tag{3}$$

Omit the derivation and the result can be written in matrix equation form as follows:

$$\begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}=\begin{bmatrix} -\frac{1}{2}[x^{(1)}(1)+x^{(1)}(2)] & 1 \\ -\frac{1}{2}[x^{(1)}(2)+x^{(1)}(3)] & 1 \\ \vdots & 1 \\ -\frac{1}{2}[x^{(1)}(n-1)+x^{(1)}(n)] & 1 \end{bmatrix}\begin{bmatrix} a \\ u \end{bmatrix} \tag{4}$$

Briefly denoted:

$$Y_n=BA, Y_n=\begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \\ x^{(0)}(n) \end{bmatrix}, B=\begin{bmatrix} -\frac{1}{2}[x^{(1)}(1)+x^{(1)}(2)] & 1 \\ -\frac{1}{2}[x^{(1)}(2)+x^{(1)}(3)] & 1 \\ & 1 \\ -\frac{1}{2}[x^{(1)}(n-1)+x^{(1)}(n)] & 1 \end{bmatrix}, A=\begin{bmatrix} a \\ u \end{bmatrix}. \tag{5}$$

There is no solution to this equation. But approximate solution could be obtained by least squares method, that is:

$$A=(B^TB)^{-1}B^TY_n=\begin{bmatrix} \hat{a} \\ \hat{u} \end{bmatrix} \tag{6}$$

Take the obtained parameters back to the original differential equation. The discrete form of prediction model could be generated, by the following formula:

$$x^{(1)}(k+1)=[x^{(1)}(1)-\frac{u}{a}]e^{-\hat{a}k}+\frac{\hat{u}}{\hat{a}}(k=0,1,2,\ldots) \tag{7}$$

This formula is called the time response function of the GM(1,1) model.

3. Feasibility Analysis and Data Processing

Not all the GM(1,1) models are valid when applied to do prediction. If the model parameters are unreasonable, it will lead to deformity of the GM(1,1) model, which may cause large load forecasting errors and even can not meet the requirements [7].

For a given sequence, the possibility of establishing a higher prediction accuracy GM(1,1) model is generally judged by the interval of sequence ratio belonged.

The feasibility theorem of gray model:

For the original series:

$$x^{(0)} = [x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n)],\ x^{(0)}(k) \in x^{(0)}$$

k=1,2,3， *…..* , *n*, define $\sigma^{(0)}(k)$ the sequence ratio:

$$\sigma^{(0)}(k) = \frac{x^{(0)}(k-1)}{x^{(0)}(k)}, k \geq 2 \tag{8}$$

when $\sigma^{(0)}(k) \in (0.1352, 7.389)$ satisfied, non-deformed model could be established with the sequence [8].

But this interval is just a basic requirement for modeling, rather than a practical one. To establish a more satisfied and effective GM(1,1) model, the sequence ratio should be in the sub-interval close to 1. According to the feasibility criteria, the following requirement should be attained:

$$\sigma^{(0)}(k) \in (e^{\frac{-2}{n+1}}, e^{\frac{2}{n+1}}) \tag{9}$$

If sequence ratio is unqualified, data transformation should be processed to make data accommodate in the area. So the principle of gray model data processing is: After treatment, the sequence ratio should be as close as possible to 1, that is, the ratio bias should be as small as possible.

Root measures are taken in this paper. Make *x* the original sequence， y_m the *m*-th root sequence：

$$\begin{cases} x = (x(1), x(2), \ldots., x(n)) \\ y_m = (\sqrt[m]{x(1)}, \sqrt[m]{x(2)}, \ldots., \sqrt[m]{x(n)}), m = 2, 3, \ldots \end{cases} \tag{10}$$

Make $\delta_x(k)$ and $\delta_m(k)$ sequence ratio deviation,

$$\begin{cases} \delta_x(k) = \dfrac{\Delta_x(k)}{x(k)} \\ \delta_m(k) = \dfrac{\Delta_m(k)}{\sqrt[m]{x(k)}} \end{cases} \tag{11}$$

To any specified positive real number ξ, $\delta_m(\max) < \xi$ could be attained by selecting the appropriate *m*.

4. Optimization of Gray Model when Applied to Super Short-Term Load Forecasting

4.1. Parameter Modification

Through large amounts of forecasting simulation, the use of GM(1,1) model to predict will have poor prediction accuracy when the data development is fast. This paper analyzed the method of the mechanism and application conditions from the theory.

The exact formula of background value for GM(1,1) should be:

$$z^{(1)}(k+1) = \alpha x^{(1)}(k) + (1-\alpha) x^{(1)}(k+1) \tag{12}$$

The following formula indicated the relation between *a* and α:

$$\alpha = \frac{1}{a} - \frac{1}{e^a - 1} \tag{13}$$

TABLE I. RELATIONSHIP BETWEEN α AND a

a	0.001	0.01	0.1	0.3	0.5	1.0
α	0.4998	0.4992	0.4916	0.4750	0.4585	0.4180
a	-0.001	-0.01	-0.1	-0.3	-0.5	-1.0
α	0.5001	0.5008	0.5083	0.5250	0.5414	0.5820

Regardless of the value of *a*, the traditional GM(1,1) model calculates the background value from:

$$z^{(1)}(k+1) = \frac{1}{2} x^{(1)}(k) + \frac{1}{2} x^{(1)}(k+1) \tag{14}$$

Therefore, according to different values of *a*, selecting a different background value α to address the issue of prediction accuracy when $|a|$ is large [10].

Using the conclusion above, this paper proposed α parameter modification to improve the model, algorithm as follows:

- Step.1:

First take GM(1,1) modeling process and forecasting operations. After the 1-AGO accumulating, the first-order accumulated sequence is generated. The mean value for the sequence is:

$$z^{(1)}(k+1)=\alpha x^{(1)}(k)+(1-\alpha)x^{(1)}(k+1) \tag{15}$$

In the first calculation, it could simply suppose $\alpha=0.5$.

- Step.2:

Get the parameters into the formula:

$$\alpha=\frac{1}{a}-\frac{1}{e^{a}-1} \tag{16}$$

Compare $\hat{\alpha}(m+1)$ with former one.

If $\left|\hat{\alpha}(m+1)-\hat{\alpha}(m)\right|>\varepsilon$, the forecast accuracy could be greatly improved, in which ε is a given small positive integer. Return to Step.1 and use $\hat{\alpha}(m+1)$ instead to calculate background values. Take GM(1,1) modeling and forecasting operations again.

When $\left|\hat{\alpha}(m+1)-\hat{\alpha}(m)\right|<\varepsilon$ satisfied, iteration is over and turns to Step.3.

- Step.3:

Do regressive reduction and get the predictive values and the operation results.

4.2. External Optimization

According to the load variation analysis of electric system, load data within one day there will be one or more general changes in peak and valley periods. But general GM(1,1) model is an exponential growth prediction model. When the load changes in a smooth rise or decline, load forecast will be more accurate. When load changes acutely forecast results may turn bad. So a sub-optimal portfolio is designed to choose the best prediction program to improve prediction accuracy.

Correlation analysis is a method to determine the associated extent between system factors. The basic idea of correlation is based on the similarity between curves, mainly the comparison between geometric shapes. This method can be used to compare the concentration between predicted curve of several forecasting model and the fitting degree of corresponding actual load curve. With greater correlation, it indicates the corresponding prediction model more excellent and smaller error.

The predicted curves are denoted:

$$x_i=\{x_i(1),x_i(2),...,x_i(n)\,|\,i=1,2,...,m\}$$

The actual value curve x_0 is denoted:

$$x_0=\{x_0(1),x_0(2),...,x_0(n)\}$$

The actual value curve often takes a set of data closest to the prediction day or the actual known curve. $\varepsilon_i(k)$ is the association coefficient between x_0 and x_i on point k. r_i is the correlation of x_i and x_0. That is:

$$\varepsilon_i(k)=\frac{\min\limits_i\min\limits_k|x_0(k)-x_i(k)|+\rho\max\limits_i\max\limits_k|x_0(k)-x_i(k)|}{|x_0(k)-x_i(k)|+\rho\max\limits_i\max\limits_k|x_0(k)-x_i(k)|} \tag{17}$$

Calculate the correlation of integrated points. Then the correlation of predicted curves x_i and actual curve x_0 could be drawn from:

$$r_i=\frac{1}{n\sum\limits_{k=1}^{n}\varepsilon_i(k)} \tag{18}$$

This paper compares the data for each day within a week. Through comparing one day with the day before that day and the previous week of the same type, it is found that the changes are similar around load peak and valley. Therefore, the peak load and the valley location could be used to do segmentation. The specific section could be divided in accordance as below[9]:

1) Early morning sessions: From 0:00 to the first small peak load. In this section electricity load is generally the lowest through out the day.

2) Morning sessions: From the first peak load to the valley load. After the first peak the electric load may continue for some time and then begin to decrease. Then the first valley appears. The electric load in this section has a relatively strong correlation of the working day, weather and temperature.

3) Afternoon sessions: From the valley load to the evening peak load. The afternoon sessions load have a big relationship of the size of the regional power grid, power region, season, temperature, weather and working days.

4) Evening sessions: From evening peak load to 23:00. There is a relatively strong correlation between evening electric load, weather and temperature.

Compare every correlation coefficient of corresponding prediction program and chose the maximum one:

$$\bar{\varepsilon}_i = \frac{\sum_{k=t_1+1}^{t_2} \varepsilon_i(k)}{t_2 - t_1}, i = 1, 2, 3 \tag{19}$$

in which t_1 and t_2 is the segmentation time。

4.3. Posteriori Error Checking

Posteriori error checking is based on the statistical case between predicted and actual values to calibration, according to the size of the absolute value of residuals, the residual point's inspection possibility and the forecast error variance. The specific methods are as follows

The historical load sequence:

$$x^{(0)} = \{x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n)\}$$

The predictive load sequence:

$$\hat{x}^{(0)} = \{\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), ..., \hat{x}^{(0)}(n)\}$$

The residual at time *k*:

$$\varepsilon_k = x^{(0)}(k) - \hat{x}^{(0)}(k), (k = 1, 2, ..., n) \tag{20}$$

The variance of historic load sequence (actual values) is:

$$S_1^2 = \frac{1}{n}\sum_{k=1}^{n}(x^{(0)}(k) - \bar{x})^2 \tag{21}$$

The variance of residual is:

$$S_2^2 = \frac{1}{m}\sum_{k=1}^{m}(\varepsilon(k) - \bar{\varepsilon})^2 \tag{22}$$

According to this posteriori error ratio *C* and small error probability *P* could be obtained from:

$$C = \frac{S_2}{S_1}$$

$$P = P\left\{\left|\varepsilon(k) - \bar{\varepsilon} < 0.674 S_1\right|\right\} \tag{23}$$

The value of *C* indicates the dispersion between forecast value and actual value. The value of *P* represents the probability that the difference between the average residual and the residual value less than the given value $0.674S_1$. The smaller of *C*, the better the result will be, which indicates that the difference between the predicted values and value is not discrete. Bigger *P* is better because it indicates that more points are in the interval $0.674S_1$.

Index *C* and *P* could be used to assess the accuracy of prediction model, as shown in the following table.

TABLE II. FORECASTING EVALUATION

Accuracy	P	C	Accuracy	P	C
Excellent	>0.95	<0.35	Inadequate	>0.7	<0.45
Qualified	>0.8	<0.5	Unqualified	<=0.7	>=0.65

4.4. Correction System

Using the latest available information in load forecast is one of the keys to improve super short-term load forecasting accuracy. When the weather suddenly changes or for other reasons load demand changes suddenly, using previous method to forecast load will have a large deviation with actual values. This occasion will make the predicted value consecutive smaller or larger than the actual situation. Since then the actual value and the predictive value of the load curve shape is similar to the changes for the forecast value and actual value of the increase or decrease with almost the same tendency. Load in the latest available information and it could modify the forecast value. Amended as follows:

$$R(t) = \frac{1}{m}\sum_{j=1}^{m}(l(t-j) - \hat{x}(t-j))$$
$$\hat{x}_c = \hat{x} + R(t) \tag{24}$$

In which: *R(t)* is the error correction; $l(t-j)$ is the actual load value *j*-hour before the forecast time t; $\hat{x}(t-j)$ is the gray model prediction value *j*-hour before the forecast time t; $\hat{x}$ is the gray model prediction value of time t; $\hat{x}_c$ is the corrected prediction value.

Basis to determine whether an amendment is needed as follows:

$$\frac{l(t-j) - \hat{x}(t-j)}{l(t-j)} > P \tag{25}$$

When the number of points meeting the formula above is more then *q*(*0*<*q*<*m*), the predicted values should be amended.

5. Computation and Simulation

Based on the data of electric load data per 15mins from May 15 to June 14, 2010 of Nanjing, China, the above improved gray model is used to forecast the load values of June 5, 2010.

First of all, data entry program should be chosen. There are mainly three kind of different raw data series for use:

1) Option 1: The electric load data of the same moment five days before.

2) Option 2: The electric load data of the same day and moment five weeks before.

3) Option 3: The electric load data of three hours before forecasting moment.

Calculate the correlation of different options:

TABLE III. CORRELATION OF THREE OPTIONS

Options	(0-10)	(11-14)	(14-19)	(20-23)
Pro.1	0.842	0.786	0.690	0.732
Pro.2	0.763	0.824	0.826	0.715
Pro.3	0.815	0.709	0.671	0.784
Decision	P.1	P.2	P.2	P.3

The following figure is about the load forecast results and forecast errors of June 5.

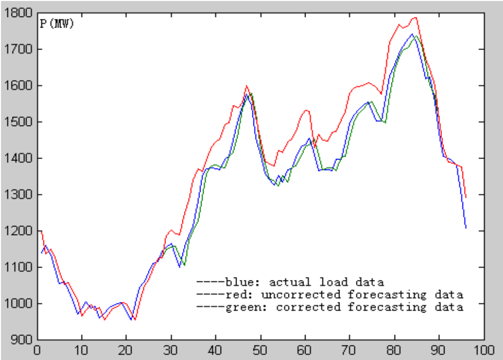

Fig.1. Load forecasting results

The prediction curve is close to the actual one. But with the mutations of the weather that day, the load change is not stable and prediction error is quite big. So the introduction of amendments to the system is necessary because the prediction three-hour average error is more than 2%.

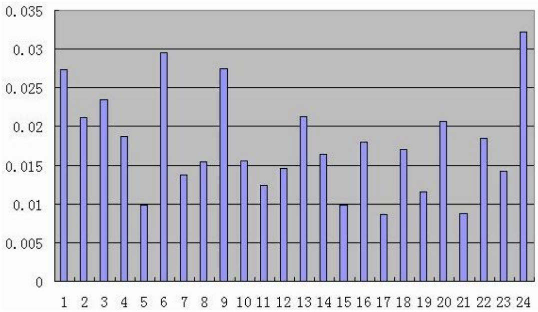

Fig 2. Forecasting errors of gray model corrected

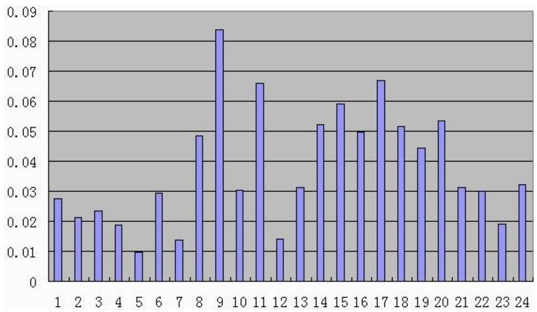

Fig.3. Forecasting error of gray model uncorrected

TABLE IV. RESULT COMPARISON

Conclusion	*S*	ε	*C*	*P*	Appraisal
uncorrected	40.30	44.71	0.20	0.89	Qualified
corrected	29.11	-0.57	0.13	0.97	Excellent

General GM(1,1) model of daily average error for a week is about 3.4% and 2.7% for the improved one. But with weather mutations (large scale rainfall on June 5) forecasting error may be quite big using improved gray model. After the introduction of amendments, system error decreased from 4.68% to 1.89% on average. Forecast accuracy has been greatly improved.

6. Conclusion

This paper introduced an improved gray model system with high prediction accuracy when applied to super short-term load forecasting based on internal and external improvements and residual correction. The simulation shows that the improved mode has high prediction accuracy with overall smooth cases. The residual correction also could greatly improve the prediction accuracy when part of forecast region is susceptible to the mutation of external factors. Compared with conventional grey model methods, the proposed scheme has the characteristic of simple computation, high forecasting precision and good applicability. Therefore, it is appropriate as an aid tool to solve the super short-term forecasting problems in power system.

7. Acknowledgment

This work is supported by Knowledge Innovation Program of the Chinese Academy of Science (KGCX2-YW-399+2) and the State-owned Assets Supervision and Administration Commission of Baiyin City, Gansu Province, China.

8. References

[1] A.S. Debs, “Modern Power System Control and Operation”, ISBN: 0-89838-265-3. Kluwer Academic, Boston, 1988.

[2] Senjyu T, Mandal P, Uezato K, Funabashi T, “Next Day Load Curve Forecasting Using HybridCorrection Method”. IEEE Transactions on Power Systems, Volume99, Issue99, 2004.

[3] Drezga I, Rahman S, “Input Variable Selection for ANN-based Short-term Load Forecasting. IEEE Trans on Power Systems”, 1998.

[4] Deng Julong, Xiao Xinping, “A new modified GM(1,1) model: grey optimization model”. Journal of Systems Engineering and Electronics, 2001(2):1-8.

[5] Dipti Srinivasan, Swee Sien Tan, C.S.chang, “Parallel Neural Network-Fuzzy Expert System Strategy For Short-term Load Forecasting: System Implementation And Performance Evaluation”, IEEE Trans on Power System,2006, 144(3): PP1100-1106.

[6] Y. Dai, J.D. McCalley, V.Vittal, “Stochastic Load Model Identification and Its Possible Applications”. North America Power System Symposium, University of Wyoming at Laramie, 1997, 505~512.

[7] Marin, F.J., Garcia-Lagos, F. Joya, G.Sandoval, F, “Global model for short-term load forecasting using artificial neural networks”, Generation, Transmission and Distribution IEE Proceedings, Volume 149, Issue 2, March 2002.

[8] Senjyu T., Takara, H., Uezato K., Funabashi T. “ONE-hour-ahead load forecasting using neural network. IEEE Transactions on Power Systems”, Volume17, Issue1, Feb.2002.

[9] Gwo-Ching Liao, Ta-Peng Tsao, “Integrating evolving fuzzy neural networks and tabu search for short term load forecasting”, Transmission and DIStribution Conference and Exposition,2003 IEEE PES, Volume2,7-12 Sept.2003.

[10] Drezga I, Rahman S, “Input Variable Selection for ANN-based Short-term Load Forecasting”, IEEE Trans on Power Systems,1998.