Power Investment Prediction Based on the Improved GM(1,1) Model

Shuang-qing LIN¹, Rui WANG¹, Ming-wei LI¹, Jian-jia ZHOU¹, Yan DENG¹, Tao LUO¹ and Ji-yang ZHANG²,*¹
Sichuan Electric Power Company, State Grid Corporation of China, China
²School of Automation Engineering, University of Electronic Science and Technology of China, China
*Corresponding author

Keywords: Power investment prediction, GM(1,1) Model, Improved GM(1,1) model, Optimization of background value.

Abstract. Power investment prediction plays an important part in fund management of power company. The traditional GM(1,1) model has been widely used in prediction. However, as the background value of traditional model has large error, using this model to predict power investment is unsuitable. In this paper, an improved GM(1,1) model is proposed. The improved model has better prediction accuracy in power investment prediction by using cubic spline interpolation to optimize background value. Case study using the power investment data of 6 cities from 2005 to 2010 demonstrates the feasibility and validity of the improved model.

Introduction

With the rapid development of economy in China, the demand for electricity has increased. Because the scale of power grid construction is expanding, power investment is increasing. In the first six months of 2016, the power investment of State Grid Corporation of China is ¥210.1 billion. Therefore, how to manage and distribute power investment scientifically is an important issue that the power company need to solve. Forecasting power investment accurately is an effective way to solve this problem, which can achieve the maximum use of power investment.

One of the models which has been mainly used for power investment prediction is the gray GM(1,1) model. It is proposed by Prof. Deng of HUST and widely used to analyze uncertain information gray systems. As this model requires low amount of data and can obtain high forecasting accuracy, it is suitable for prediction in industry and economy [1]. However, this model also has some defects. As for the power investment data, since the data fluctuates greatly, the prediction effect of the GM(1,1) model is not good. Therefore, this model needs to be improved to get higher prediction accuracy.

In this paper, an improved gray GM(1,1) model is presented, which is based on the optimization of background value. The solution to optimize the background value is cubic spline interpolation. In this way, the improved model has less error, higher prediction accuracy and easier calculation for power investment than before.

The Traditional Gray GM(1,1) Model

The traditional gray GM(1,1) model has three operations: (1) accumulated generating operation (AGO), (2) gray modeling, (3) inverse accumulated generating operation (IAGO) [2]. The modeling steps of the traditional GM(1,1) model are as follows.

(1) Suppose \( X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n)\} \) is the original data, then the AGO sequence \( X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \ldots, x^{(1)}(n)\} \) can be calculated.
\[ x^{(i)}(j) = \sum_{i=1}^{j} x^{(0)}(i), \text{ } j = 1, 2, \ldots, n \]  

(1)

(2) The background value sequence \( Z^{(i)} = [z^{(i)}(2), \ldots, z^{(i)}(n)] \) is constructed by \( X^{(i)} \).

\[ z^{(i)}(k+1) = 0.5 \left[ x^{(i)}(k+1) + x^{(i)}(k) \right], k = 1, 2, 3, \ldots, n-1 \]  

(2)

(3) The following one-order linear differential equation can be built by \( X^{(i)} \).

\[ \frac{dx^{(i)}(t)}{dt} + ax^{(i)}(t) = b \]  

where \( a \) is called the development coefficient and \( b \) is called the gray input [3].

(4) The Gray theory employs the Least squares (LS) to obtain the approximation of \( a \) and \( b \).

\[
A = \begin{pmatrix} a \\ b \end{pmatrix} = (B^T B)^{-1} B^T Y, \quad Y = \begin{pmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{pmatrix}, \quad B = \begin{pmatrix} -\frac{1}{2} \left[ x^{(0)}(2) + x^{(0)}(1) \right] \\ -\frac{1}{2} \left[ x^{(0)}(3) + x^{(0)}(2) \right] \\ \vdots \\ -\frac{1}{2} \left[ x^{(0)}(n) + x^{(0)}(n-1) \right] \end{pmatrix}
\]  

(4)

According to Eq.4, the predicted value of original time-series data can be obtained through IAGO.

\[ x^{(0)}(k+1) = x^{(i)}(k+1) - x^{(i)}(k) = (1 - e^a)(x^{(0)}(1) - \frac{b}{a})e^{-ak}, \text{ } k = 1, 2, \ldots \]  

(5)

The Improved Gray GM(1,1) Model

Analysis of Error Reason for the Traditional GM(1,1) Model

Through the analysis of the traditional GM(1,1) model, the results can be seen as follows. The prediction value of GM(1,1) model is calculated by parameter \( a \) and \( b \), that means the prediction accuracy is determined by \( a \) and \( b \). From Eq.4 we can see that \( a \) and \( b \) depend on original data and background value. As the original data sequence is invariant, whether the equation of background value is logical or not affects the prediction accuracy directly.

Eq.6 can be obtained by integrating Eq.3 in interval \([k, k+1]\).

\[
\int_{k}^{k+1} \frac{dx^{(i)}(t)}{dt} dt + ax^{(i)}(t) dt = b \\
\Rightarrow x^{(i)}(k+1) - x^{(i)}(k) + a \int_{k}^{k+1} x^{(i)}(t) dt = b
\]  

(6)

Since \( x^{(0)}(k+1) = x^{(i)}(k+1) - x^{(i)}(k) \), Eq.7 is equal to Eq.6

\[ x^{(0)}(k+1) + a \int_{k}^{k+1} x^{(i)}(t) dt = b \]  

(7)

Eq.8 can be obtained by Eq.3.
\[ x^{(0)}(k+1) + az^{(1)}(k+1) = b \quad (8) \]

From Eq.7 and Eq.8, we can see the background value \( z^{(1)}(k+1) = \int_{k}^{k+1} x^{(1)}(t) \, dt \). As Figure1 shows, because the original data is a discrete time series, the traditional model calculates approximate value of \( \int_{k}^{k+1} x^{(1)}(t) \, dt \) by trapezoidal formula. So \( z^{(1)}(k+1) = 0.5 \left[ x^{(1)}(k+1) + x^{(1)}(k) \right] \) is obtained as Eq.2 shows. Because the area formed by the curve in \([k, k+1]\) is always smaller than the area of the trapezoid, the approximate value of the background value is always larger than the true value. The shadow part in Figure1 is the error reason for the traditional GM(1,1) model [4].

![Figure 1. Prediction error source diagram of the traditional GM(1,1) model.](image)

When the sequence data changes sharply, there is a large error between the true value and the calculated value of \( z^{(1)}(k+1) \). Because of the rapid changes in power investment data, we need improve the traditional GM(1,1) model to get higher prediction accuracy.

**The Improved GM(1,1) Model based on the Optimization of Background Value**

The solution to improve GM(1,1) model is ameliorating the way to calculate background value. Improving the accuracy of \( z^{(1)}(k+1) \) need reduce the area of the shadow part in Figure1. As Figure2 shows, taking the midpoint of the curve in the interval \([k, k+1]\) and getting the approximation of the background value by the area of two trapezoids. The shadow area is smaller than that in the traditional GM(1,1) model. That means the error between approximate value and actual value of background value is getting smaller and the prediction accuracy of the GM(1,1) model will increase.

![Figure 2. Diagram of background value optimization.](image)

The approximate value of the optimized background value can be calculated through Figure2.

\[
 z^{(1)} = \int_{k}^{k+0.5} x^{(1)}(t) \, dt + \int_{k}^{k+0.5} x^{(1)}(t) \, dt \approx 0.5x^{(1)}(k+\frac{1}{4}) + 0.5x^{(1)}(k+\frac{3}{4}) \quad (9)
\]

Because \( x^{(1)}(t) \) is discrete time series data, the value in the interval \([k, k+1]\) need to be calculated. Cubic spline interpolation is selected to calculate these values. It is commonly used to obtain a smooth curve from discrete data points. This fitting curve can represent the trend of data \( x^{(1)}(t) \).
Therefore, \( x^{(i)}(k + \frac{1}{4}) \) and \( x^{(i)}(k + \frac{3}{4}) \) can be obtained by cubic spline interpolation.

The improved GM(1,1) model uses Eq.9 instead of Eq.2 to calculate the background value \( Z^{(i)} \). The rest of the modeling steps is the same as the traditional GM(1,1) model.

**Example Application**

In order to prove that the improved gray GM(1,1) model has better prediction accuracy, the power investment data of some cities in China is selected to analyze the prediction model (Table 1).

**Table 1. Power investment data [¥100M].**

| city | 2005   | 2006   | 2007   | 2008   | 2009   | 2010   |
|------|--------|--------|--------|--------|--------|--------|
| city A | 0.9100 | 1.8900 | 2.300  | 3.7200 | 8.2600 | 12.7700|
| city B | 0.3180 | 1.2900 | 2.7100 | 1.8600 | 1.2150 | 2.3180 |
| city C | 0.6490 | 2.0827 | 1.7414 | 1.8971 | 1.7601 | 1.8623 |
| city D | 2.3690 | 3.9137 | 2.9063 | 2.0259 | 4.4292 | 3.3095 |
| city E | 1.8800 | 2.1800 | 1.2300 | 1.5200 | 2.8600 | 4.1000 |
| city F | 0.0200 | 0.2000 | 1.4340 | 1.6821 | 2.3900 | 4.1103 |

The data selected in Table 1 is the power investment data of 2005-2010 in 6 cities. Because the amount of time series data in every city is limited, this paper constructs prediction models based on the traditional GM(1,1) model and the improved GM(1,1) model by using every city's power investment data of 2005-2009. According to prediction models, the power investment forecasting value of every city in 2010 can be obtained. By comparing the real value and the forecasting value of every city in 2010, the prediction effects of two prediction models can be displayed.

In order to show the prediction accuracy of each model, absolute percentage error (APE) and mean absolute percentage error (MAPE) are adopted for performance evaluation [5].

\[
MAPE = \frac{1}{m} \sum_{i=1}^{m} APE_i, \quad APE_i = \frac{|y^{(i)}_{p} - y^{(i)}_{t}|}{y^{(i)}_{t}}
\]

(10)

Where \( m \) denotes the number of cities, \( y^{(i)}_{p} \) represents the predicted value and \( y^{(i)}_{t} \) stands for the original value.

The prediction results of 6 cities in 2010 are shown in Table 2.

**Table 2. Prediction results of 6 cities in 2010.**

| city | traditional GM(1,1) prediction data [¥100M] | improved GM(1,1) prediction data [¥100M] | original data [¥100M] | traditional GM(1,1) APE | improved GM(1,1) APE | the difference of APE between two GM(1,1) models |
|------|---------------------------------------------|------------------------------------------|-----------------------|------------------------|----------------------|-----------------------------------------------|
| A    | 11.2374                                     | 12.4127                                  | 12.7700               | 0.1200                 | 0.0280               | 0.0920                                        |
| B    | 1.5654                                      | 1.6075                                   | 2.3180                | 0.3247                 | 0.3065               | 0.0182                                        |
| C    | 1.6715                                      | 1.6728                                   | 1.8623                | 0.1025                 | 0.1018               | 0.0007                                        |
| D    | 3.5294                                      | 3.4624                                   | 3.3095                | 0.0664                 | 0.0462               | 0.0202                                        |
| E    | 2.7786                                      | 2.8506                                   | 4.1000                | 0.3223                 | 0.3047               | 0.0176                                        |
| F    | 3.9499                                      | 4.2234                                   | 4.1103                | 0.0390                 | 0.0275               | 0.0115                                        |

According to Table 2, the MAPE of each prediction model can be calculated. The MAPE are shown in Table 3.

**Table 3. MAPE of the traditional GM(1,1) model and the improved GM(1,1) model.**

|                      | traditional GM(1,1) MAPE | improved GM(1,1) MAPE | the difference of MAPE between two GM(1,1) models |
|----------------------|-------------------------|----------------------|-----------------------------------------------|
| A                    | 0.1625                  | 0.1538               | 0.0267                                        |
Figure 3 give a visual comparison between the traditional model and the improved model.

It can be seen that the MAPE of the improved model based on cubic spline interpolation is 13.58%, which is 2.67% less than that of the traditional model. Besides, the improved model has better performance in every city. The APE of the improved model is smaller than that of the traditional model in every city. Especially in city A, the prediction error of the improved model is reduced by 9.2% compared with the traditional model. Therefore, the prediction accuracy of the improved GM(1,1) model is better than the traditional GM(1,1) model.

Summary

By analyzing the traditional GM(1,1) model, this paper finds that the background value is the main source of prediction error. Therefore, an improved gray GM(1,1) model is presented. This model reduces the error of background value through a new formula based on cubic spline interpolation. The improved model is applied to power investment prediction in 6 cities. The result shows the error of the improved GM(1,1) model is smaller than that of the traditional GM(1,1) model.

References

[1] Wang, Yang, Shanlin, et al. Dynamic GM(1,1) Model Based on Cubic Spline for Electricity Consumption Prediction in Smart Grid[J]. China Communications, 2010, 7(4):83-88.
[2] Lei Y, Guo M, Zhao D, et al. Application of Grey Model GM(1,1) to Ultra Short-Term Predictions of Universal Time[J]. Artificial Satellites, 2016, 51(1):19-29.
[3] Hou Li-qiang, Yang Shan-lin, et al. Mid-term Load Forecasting Based on Buffer Operator and Modified Grey Model[J]. Journal of System Simulation, 2013(s1):1-5.
[4] Dong Ke, Lv Wen-yuan. Optimization of Background Value in GM(1,1) Based on Cubic Bezier Basis Function Interpolation[J]. Journal of Mathematics, 2017, 37(5): 1022-1028.
[5] Wang L, Fu D M. Estimation of Missing Values Using a Weighted K-Nearest Neighbors Algorithm[C]//International Conference on Environmental Science and Information Application Technology. IEEE, 2009:660-663.