PAPER • **OPEN ACCESS**

Gray linear regression model based on adaptive particle swarm optimization power load forecasting method

To cite this article: Caijuan Qi *et al* 2019 *IOP Conf. Ser.: Earth Environ. Sci.* **218** 012144

View the article online for updates and enhancements.



IOP ebooks™

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of every title for free.

This content was downloaded from IP address 117.85.168.23 on 21/09/2019 at 13:41

Gray linear regression model based on adaptive particle swarm optimization power load forecasting method

Caijuan Qi[1], Kun Zhang[1], Shuhong Shi[2], Qian Zhang[2*]

[1] State Grid Ningxia Electric Power Eco-Tech Research Institute, Yinchuan, Ningxia Province, 750004, China

[2] Beijing Kedong Electric Power Control System CO.,LTD, Beijing, 100192, China

[*]Corresponding author's e-mail: 1250421169@qq.com

Abstract. Aiming at the trend of "S" growth in annual electricity consumption in Ningxia, the gray linear regression combined model (GLM) load forecasting method is adopted, which can improve the lack of exponential growth trend in linear regression model and the lack of linear factors in the GM(1,1) model. However, the traditional GLM model has theoretical defects in the process of solving parameters, and there are limitations in the application range. Therefore, this paper introduces an adaptive particle swarm optimization algorithm which is more efficient than the standard particle swarm optimization algorithm, and combines it with the gray linear regression combination model. The adaptive particle swarm optimization algorithm is used to solve the parameters of the GLM model, then carry out residual correction. Adaptive particle swarm optimization gray linear regression combination model (APSO-GLM) is proposed. The example analysis shows that the model overcomes the defects of the GM(1,1) model and the linear regression model. Compared with the single model, the prediction accuracy is higher and it has a wider application range than the traditional GLM model.

1. Introduction

Load forecasting is an important part of power system economic dispatch and an important module of an energy management system (EMS).Since the load forecast is based on the past and present of the power load to speculate its future value. Therefore, the object of the load forecasting work is the uncertainty event [1]. Methods of load forecasting are mainly divided into classical forecasting methods and modern forecasting methods [2-5]. Electric measurement is one of the important tasks of the power sector. Accurate load forecasting can be economically and reasonably arrange the start and stop of the internal generator set of the power grid, maintain the safety and stability of the power grid operation, reduce the unnecessary rotating reserve capacity, and arrange the unit maintenance plan reasonably, guarantee the normal production and life of society, effectively reduce the cost of power generation, and improve the economic and social benefits [6].

As the national economy enters the "new normal", as well as the adjustment of relevant national policies and the advancement of power system reform, many social and economic indicators in Ningxia are no longer exponentially growing, but have certain "S"-shaped curve characteristics [7]. At the same time, the factors affecting load changes are also complicated. These factors have strong uncertainties. It is not easy to find out all the factors. Therefore, the power load can be regarded as a gray system [8]. The gray prediction method requires less data and higher prediction accuracy [9], but the gray model is mainly applicable to a single exponential growth data sequence. The principle of linear regression prediction [10] is the continuity of things development and the correlation of things

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

causality. Considering the future as the current continuation, under the condition that the various conditions are relatively stable, the future development will be predicted, and the short-term forecast can achieve better results, but the long-term forecast is often not effective. In view of this, in order to predict the annual electricity consumption of Ningxia region with both linear trend and exponential trend, we adopts the gray linear regression combination model [11] constructed by combining the gray model GM(1,1) and the linear regression model in this paper. It can improve the lack of exponential growth trend in the linear regression model and the lack of linear factors in the GM(1,1) model, but the traditional GLM model has theoretical defects in the process of solving parameters.

Particle swarm optimization algorithm (PSO) is a cluster intelligent optimization algorithm developed in recent years. It has many advantages such as fewer parameters, faster convergence, and global search capability. It has been widely used in various fields [12]. In this paper, a more efficient adaptive particle swarm optimization algorithm is introduced. Combined with the GLM model, APSO-GLM is proposed. Then the residual correction is performed to further improve the prediction accuracy. Theoretical research and practice show that the proposed model has higher precision than a single prediction model and traditional GLM, can enhance the stability of prediction, and has higher ability to adapt to future environmental changes.

2. Gray linear regression combined model

Set the original data sequence $x^{(0)}$

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

Generate a sequence of first order accumulation with 1-AGO

$$x^{(1)}(t) = \sum_{i=1}^{t} x^{(0)}(i) \tag{2}$$

Then: $\hat{x}^{(1)}(t) = C_1 e^{vt} + C_2 t + C_3$ is the GLM model. Parameters v, C_1, C_2 and C_3 are undetermined.

In order to obtain the above parameters, set the parameter sequence

$$Z(t) = x^{(1)}(t+1) - x^{(1)}(t) = C_1 e^{vt}(e^{v} - 1) + C_2 \qquad (t = 1,2,\cdots,n-1) \tag{3}$$

$$Y_m(t) = Z(t+m) - Z(t) = C_1 e^{vt}(e^{vm} - 1)(e^{v} - 1) \qquad (m = 1,2,\cdots,n-3; t = 1,2,\cdots,n-m-2) \tag{4}$$

Then

$$v_m(t) = \ln \frac{Y_m(t+1)}{Y_m(t)} \tag{5}$$

Considering that m takes different values, the final fitting parameters $v_m(t)$ will also be different, so we can take m=1,2,…,n-3, $v_m(t)$ is calculated separately, and the mean value of all the obtained $v_m(t)$ is taken as the fitting value of the identified parameter v.

$$v = \frac{2\sum_{m=1}^{n-3}\sum_{t=1}^{n-m-3} v_m(t)}{(n-2)(n-3)} \tag{6}$$

Make $L(t) = e^{vt}$, then $\hat{x}^{(1)}(t) = C_1 L(t) + C_2 t + C_3$.

The estimated values of C_1, C_2 and C_3 can be obtained by using the least square method

$$x^{(1)} = \begin{bmatrix} x^{(1)}(1) \\ x^{(1)}(2) \\ \vdots \\ x^{(1)}(n) \end{bmatrix}, \quad C = \begin{bmatrix} C_1 \\ C_2 \\ C_3 \end{bmatrix}, \quad A = \begin{bmatrix} L(t) & 1 & 1 \\ L(t) & 2 & 1 \\ \vdots & \vdots & \vdots \\ L(t) & n & 1 \end{bmatrix}$$

$$C = (A'A)^{-1} A' x^{(1)} \tag{7}$$

The predicted value of the generated sequence is $\hat{x}^{(1)}(t) = C_1 e^{vt} + C_2 t + C_3$

The prediction data can be obtained by reduction and reduction.

$$\hat{x}^{(0)}(t+1)=\hat{x}^{(1)}(t+1)-\hat{x}^{(1)}(t) \quad (t=1,2,\cdots,n) \tag{8}$$

3. Adaptive particle swarm optimization algorithm

The particle swarm optimization algorithm is a global optimization evolution algorithm proposed by Kennedy and Eberhart, which are derived from the simulation of bird predation behavior [13]. The specific process is as follows: Suppose that in a *D*-dimension target search space, there are *N* particles forming a community, where the *i*th particle position and velocity are respectively

$$Y_i=(y_{i1},y_{i2},\cdots,y_{iD}), \quad i=1,2,\cdots,N \tag{9}$$

$$V_i=(v_{i1},v_{i2},\cdots,v_{iD}), \quad i=1,2,\cdots,N \tag{10}$$

The optimal position so far searched by the first particle is called individual extremum, and the optimal position so far searched by the whole particles swarm is called global extremum, respectively

$$P_{best}=(p_{i1},p_{i2},\cdots p_{iD}) \tag{11}$$

$$g_{best}=(p_{g1},p_{g2},\cdots,p_{gD}) \tag{12}$$

When these two extremums are found, the particles update their speed and position according to the following formula:

$$v_{id}=wv_{id}+c_1r_1(p_{id}-y_{id})+c_2r_2(p_{gd}-y_{id}) \tag{13}$$

$$y_{id}=y_{id}+y_{id} \tag{14}$$

Where w is the inertia factor, c_1 and c_2 are learning factors, usually c_1=c_2=2, r_1 and r_2 are uniform random numbers in the range [0,1], v_{id} is the velocity of the particle. In the standard particle swarm algorithm, the inertia factor w is generally set to 1. Since the inertia factor is a variable that affects the current particle velocity, a larger value is advantageous for global search, and a smaller value is advantageous for local search, in order to better balance the algorithm. This paper proposes a linearly decreasing inertia weighting factor, as shown by equation (16).

$$w(k)=w_{\max}-\frac{(w_{\max}-w_{\min})\times k}{k_{\max}} \tag{15}$$

Where, $w_{\max}$ is the inertia weight maximum, generally set to 0.9; $w_{\min}$ is the inertia weight minimum, generally set to 0.4; k is the current iteration number; $k_{\max}$ is the maximum number of iterations.

4. Gray linear regression combined model based on APSO optimization

The use of GLM model for power load forecasting requires less load data, and does not need to consider the load distribution rule and trend. However, when the model predicts a faster-growing load, the prediction accuracy deteriorates. The reason is that in the process of solving the parameter v, the average value of v_m at m different values is simply taken.

In order to avoid errors caused by improper parameter values, this paper uses the APSO algorithm to solve the parameters v, C_1, C_2, C_3, so that accurate prediction can be made even when the load increases rapidly. According to the basic steps of the APSO model described above, set particle is 4-dimension, and its position is [v, C_1, C_2, C_3]. The APSO algorithm is used to solve the parameters v, C_1, C_2, and C_3. The basic steps are as follows:

Step 1: Initialization: The initial position and velocity of particles are set according to formulas (9) and (10).

Step 2: The fitness value of each particle is calculated: The individual optimal position of the particle to the current position P_{best} and the global optimal position g_{best} are set to the position of the best particle in the initial particle population, and calculate the adaptation of each particle through the objective function f. The objective function $f(Y_i)$ is

$$f(Y_i)=\frac{1}{2}\sum_{t=1}^{n}[x^{(0)}(t)-\hat{x}^{(0)}(t)]^2 \tag{16}$$

Where, $x^{(0)}(t)$ and $\hat{x}^{(0)}(t)$ are the kth actual and predicted values of the load respectively.

Step 3: Determine whether the convergence criterion of the algorithm is satisfied. If yes, execute step 6, end the iteration, and output the values of C_1, C_2, C_3, and v; otherwise, execute step 4.

Step 4: Use the formulas (13) and (14) to update the velocity and position of each particle. If the particle fitness is better than the corresponding fitness of g_{best}, g_{best} is set to the new position.

Step 5: Determine whether convergence criterion of the algorithm is satisfied. If yes, execute step 6, and output values of C_1, C_2, C_3, and v; otherwise, return to step 4 to iterate and continue to optimize.

Step 6: Output the global optimal position g_{best}, and obtain the global optimal solution of the gray linear regression combined model parameters C_1, C_2, C_3 and v, and the algorithm runs.

5. Case analysis

In order to verify the validity and practicability of APSO-GLM proposed, annual electricity consumption in Ningxia region from 2011 to 2017 was used as sample data, and the original data is shown in table 1. It can be seen from table 1 that annual electricity consumption in Ningxia region shows an increasing trend, but there is obvious volatility. Linear regression model, GM(1,1) model, GLM model and APSO-GLM model proposed in this paper are used for fitting, and the model with the highest fitting precision is used to predict the annual electricity consumption from 2018-2020.

Table 1. Annual electricity consumption data of Ningxia (unit: TWH).

Year number	1	2	3	4	5	6	7
Electricity consumption	724.5389	741.7916	811.1798	848.7546	878.3284	886.911	978.3029

5.1. Linear regression model prediction

The linear regression model is an equation of the form $\hat{x}(t)=at+b$. The values of the parameters a and b of the linear regression model solved by the least squares method are: a=39.95, b=678.73. The equation is as follows:

$$\hat{x}(t) = 39.95t + 678.73 \tag{17}$$

Get the predicted value

$$\hat{x} = (718.685,\ 758.638,\ 798.591,\ 838.544,\ 878.497,\ 918.450,\ 958.402)$$

5.2. Gray model prediction

The original data is accumulated and generated immediately and the mean is generated. The accumulated gray model is an equation of the form $\hat{x}^{(1)}(t)=ae^{vt}+b$. The equation is as follows by using MATLAB for curve fitting:

$$\hat{x}^{(1)}(t) = 14728.1e^{0.0048t} - 14726.2 \tag{18}$$

Using the formula (8), the fitting value of GM(1,1) to the original sequence is obtained.

$$\hat{x}^{(0)}(t) = (724.539,\ 758.075,\ 795.269,\ 834.289,\ 875.222,\ 918.164,\ 963.213)$$

5.3. Grey linear regression combined model prediction

According to formula (1) ~ (6) and MATLAB, the calculated value of v is: v=0.03; thus, the gray linear regression combination model is:

$$\hat{x}^{(1)}(t) = C_1e^{vt} + C_2t + C_3 = C_1e^{0.03t} + C_2t + C_3 \tag{19}$$

According to formula (7), MATALB is used to obtain the values of parameter C as C_1=38929.08，C_2= -443.33，C_3= -38940.51

The gray linear regression combination model is obtained

$$\hat{x}^{(1)}(t) = 38929e^{0.03t} - 443.33t - 38940.51 \tag{20}$$

The predicted values of GLM model are obtained according to formula (8)

$$\hat{x}^{(0)}(t) = (716.785,\ 763.474,\ 799.793,\ 837.204,\ 875.741,\ 915.438,\ 956.330)$$

5.4. APSO-GLM Model prediction

According to the fourth section, in the APSO-GLM model, 50 particles are randomly generated, the maximum number of iterations is 200, w_{max}=1.2, w_{min}=0.2, and c_1=c_2=2, and a gray linear regression combination model is obtained. Since the linear regression model itself has systematic errors, the residual correction of this model is required.

The predicted values of APSO-GLM model are obtained

$$\hat{x}^{(0)}(t) = (724.538,\ 749.518,\ 810.365,\ 851.461,\ 893.044,\ 897.041,\ 976.927)$$

5.5. Comparative analysis of forecasting models

The predicted and actual values of the four models are fitted, and the fit is shown in the Figure 1. The prediction accuracy of the four models is tested. The error test of the model is shown in Table 2.

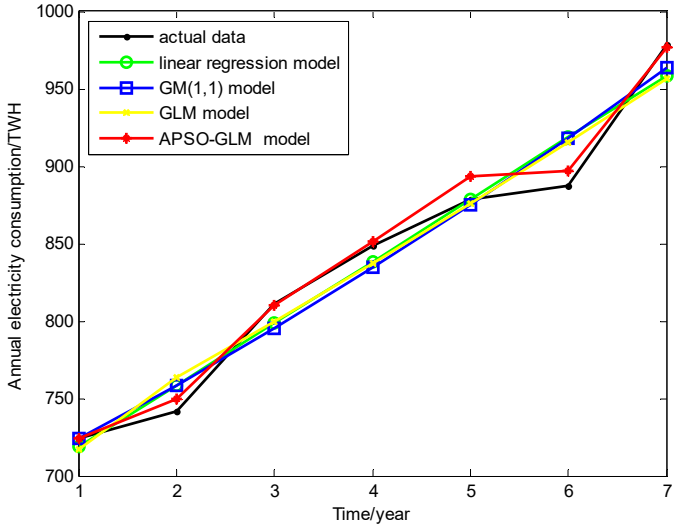


Figure 1. Comparison of predicted and actual values of each model

Table 3. Forecasting results of annual electricity consumption from 2018-2020 (unit: TWH)

Year	Annual electricity consumption
2018	1071.2
2019	1099.3
2020	1121.6

Table 2. Relative error comparison

Year	Actual value	Linear regression model		GM(1,1) model		GLM model		APSO-GLM model	
		predicted value	relative error/%	predicted value	relative error/%	predicted value	relative error/%	predicted value	relative error/%
1	724.5389	718.685	0.80790	724.539	0	716.785	1.07016	724.539	0
2	741.7916	758.638	2.27107	758.075	2.19518	763.474	2.92301	749.518	1.042
3	811.1798	798.591	1.55191	795.269	1.96138	799.793	1.40378	810.365	0.101
4	848.7546	838.544	1.20302	834.289	1.70438	837.204	1.36090	851.461	0.319
5	878.3284	878.497	1.91638	875.222	0.35366	875.741	0.29457	893.044	1.675
6	886.911	918.450	3.55600	918.164	3.52380	915.438	3.21645	897.041	1.142
7	978.3029	958.402	2.03419	963.213	1.54248	956.330	2.24605	976.927	0.141
Average relative error/%		1.6347		1.6116		1.7878		0.6313	

Through the comparison of the results of the above average relative error and the fitting of the actual value and the predicted value, it can be seen that although the four models can be used for power load forecasting, the APSO-GLM model proposed in this paper is more accuracy than the traditional GLM model, linear regression model and GM(1,1). Based on the effective synthesis of the

GM(1,1) model and the linear regression model, the prediction results are optimized. The local large deviations that may occur in the prediction process of the single prediction model are avoided, and the prediction accuracy is improved.

6. Conclusion

This article has the following three main tasks:

a. Aiming at the limitations of a single theoretical method, a GLM model that combines linear and exponential information and does not require typical distribution rules is proposed. It is superior to single gray model and linear regression model in prediction accuracy, improves the accuracy of the model, reduces the prediction error, and makes the prediction effect better.

b. In view of some theoretical problems of the traditional GLM model in solving the parameters, this paper uses the APSO algorithm to replace the traditional least squares method to solve the parameters of the GLM model, and proposes an APSO-GLM prediction model, which overcomes the limitations of the GLM model, has certain theoretical significance and application value.

c. The APSO-GLM model proposed is applied to power system load forecasting, and an experiment verifies that the model has certain feasibility and high prediction accuracy.

References

[1] Zhang, L.Y., Xiao, H.R., Wu, J.H., et al. (2018) Review of power system load forecasting. Power Systems and Big Data, 21: 52-56.

[2] Dou, Z.H., Yang, R.G., Jiao, J. (2013) Method of short-term load forecasting based on mean generating function-optimal subset regression. Transactions of the Chinese Society of Agricultural Engineering, 29: 178-184.

[3] Dou, Z.H., Yang, R.G., Zou, L., et al. (2014) Prediction method of wind speed on wind farm using optimal sub-set regression model. Renewable Energy Resources, 32: 1161-1167.

[4] Liu, W., Zhang, B., et al. (2018) Short-term load forecasting based on elastic net improved GMDH and difference degree weighting optimization. Applied Sciences, 8: 1603-1624.

[5] Wang, W.G., Shen, J., et al. (2017) Short-term load forecasting based on fuzzy mean generating function improved by back-stepping theory. Water Resources and Power, 35: 208-211.

[6] Wang, W.G., et al. (2018) Prediction method of wind speed on wind farm using optimal sub-set regression model. Renewable Energy Resources, 55: 19-24.

[7] Lee, W., Hong, J. (2015) A hybrid dynamic and fuzzy time series model for mid-term power load forecasting. International Journal of Electrical Power and Energy Systems, 64: 1057-1062

[8] Zhao, H.R., Guo, S. (2016) An optimized gray model for annual power load forecasting. Energy, 107: 272-286.

[9] Jiao, R.H., Su, C.J., Lin, B.Y., et al. (2013) Short-term load forecasting by grey model with weather factor-based correction. Power System Technology, 37: 720-725.

[10] Zhang, B., Liu, W., Li, S.T., et al. (2018) Short-term load forecasting based on wavelet neural network with adaptive mutation bat optimization algorithm. IEEJ Transactions on Electrical and Electronic Engineering, 14.

[11] Wu, G.N., Ni, S.X., Song, Z.J., et al. (2017) Prediction for substation equipment failure rate based on improved grey combination model. High Voltage Engineering, 43: 2249-2255.

[12] Cheng, S.F., Cheng, X.H., Yang, L. (2014) Application of wavelet neural network with improved particle swarm optimization algorithm in power transformer fault diagnosis. Power System Protection and Control, 42: 37-42.

[13] Li, S., Li, L.J., Zhai, M. (2012) Prediction for short-term traffic flow based on modified PSO optimized BP neutral network. Systems Engineering-Theory & Practice, 32: 2045-2049.