

Predicting Electricity Consumption Based on Optimized Model of GM(1,1)

Zhang Zheng^{1*}, Yang Shanlin¹, Liu Huizhou² and Yu Bengong¹

¹*School of Management, Hefei University of Technology, Hefei, 230009, P.R. China*

²*Tongling Power Supply Company, Anhui Electric Power Corporation, Tongling, 244000, P.R. China*

**zhzheng99@sina.com*

Abstract

Optimized GM(1,1) model based on least absolute criteria is proposed in this paper. Since the initial condition of original GM(1,1) model is not very suitable, we use the modified latest data which generating from the accumulative generating operation as the new initial condition. And the least absolute criteria is applied instead of least square criteria to improve the stability and prediction accuracy of GM(1,1) model. Then the particle swarm optimization is adapted to the parameters optimization. At the end, the optimized GM(1,1) model is used to predict the whole social electricity consumption of China and the result shows its prediction accuracy is better than the original model and the GM(1,1) model with latest initial condition.

Keywords: Prediction; GM(1,1) model, Least absolute criteria, PSO, Initial condition

1. Introduction

With the rapid economic development in China, electricity demand increases with incredible speed. In the production and living of society, electricity rationing still happens and it does significant influence to enterprise's production activities. The electricity is an indispensable part of people's daily life. But if produce electricity too much, it will cause a great waste of resources and serious environmental pollution. Thus, precise prediction of electricity consumption will make production; supply and sales of power system achieve dynamic balance. It will also guarantee economic development and increase the economic benefit and social benefit.

There are many methods used to predict electricity consumption and some certain results have achieved. Zafer and Lester[1] applied the structural time series technique to estimate an industrial electricity demand function for Turkey and this technique also uncovered the electricity Underlying Energy Demand Trend (UEDT) for the Turkish industrial sector to identify the size and significance of the price and industrial value added (output) elasticities. Azadeh *et al.*, [2] used a hybrid adaptive network based fuzzy inference system (ANFIS), computer simulation and time series algorithm to estimate and predict electricity consumption estimation. Then the proposed method is applied to the monthly electricity consumption in Iran from 1995 to 2005. Kavaklioglu [3] adopted the Support Vector Regression (SVR) model to predict Turkey's electricity consumption and achieved good results. Geoffrey and Kelvin [4] compared the three methods of traditional regression analysis, decision tree and neural networks for the prediction of electricity energy consumption.

Grey system theory was first proposed by Deng [5] in 1982 and grey model is one of the most popular approaches in prediction field. A lot of experts and scholars take part in this

research and apply to various fields of economy, environment and society. In these grey models, GM(1,1) model is the most studied model. In the meanwhile, the initial condition, background value *etc.*, have great influence to prediction results from GM(1,1) modeling process. Pretty of researches concentrate on them in order to enhance the precision of prediction. Xiao *et al.*, [6] studied GM (1, 1) model by using the thought of matrix analysis and put forward extension form GGM (1, 1) model based on the fractional order accumulated generating. Zhou and He [7] pointed out that there are two shortcomings in grey model, the homogeneous-exponent simulative deviation in GM (1, 1) model, and the unequal conversion between the original and white equations in DGM (1, 1) model. They proposed a novel model based on GM (1, 1) and DGM (1, 1) models. Huang *et al.*, [8] proposed an improved error GM (1,1) model to predict cultivated land in Yiyang and the results show that the model had high prediction accuracy. Tien [9] did some research in model GM(1,n) and applied the algorithm of GMC(1,n) to explain the existing GM(1,n) model is incorrect.

On the one hand, lots of researches focus on initial condition of GM(1,1) model. Zhang *et al.*, [10] point out that $x^{(0)}(1)$ shouldn't be limited as the only know condition when forming the prediction formula. He used $x^{(0)}(k)$ which could make the model of highest accuracy as the initial condition. Wang *et al.*, [11] take the first item and the last item of a sequence generated from applying the first-order accumulative generation operator as the new initial condition. Shih *et al.*, [12] replace the initial value of grey differential equation to the latest point to enhance its accuracy.

On the other hand, lots of researches focus on background value of GM(1,1) model. Wang *et al.*, [13] find that using an integral instead of original background value is more adaptive to whitenization equation. Wang[14] constructs the background value by using Gaussian quadrature formula and improves the prediction accuracy.

In this paper, we first introduce the original GM(1,1) model. Then we take the latest point as the initial condition to improve the accuracy of prediction. In addition, least absolute criteria is applied to the GM(1,1) model instead of least square criteria. It enhance the stability and prediction accuracy of GM(1,1) model. And particle swarm optimization (PSO) is used to parameters optimization of GM(1,1) model. At last, the whole social electricity consumption of China is adopted as the example and the results show better than the original model and the GM(1,1) model with latest initial condition.

2. Original GM (1,1) Model

There are many advantages in grey prediction method. This method can provide a high degree of accuracy when the information is uncertain and the sample data are less. The main process is as follows:

Assume a non-negative original sequence is that

$$X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\} \quad (1)$$

and the first order accumulative generating operation of $X^{(0)}$ is shown as follows

$$X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)\} \quad (2)$$

Where

$$x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i), \quad k = 1, 2, \dots, n \quad (3)$$

Due to the sequence $x^{(1)}(k)$ possesses exponential growth rule, and the solution of first order differential equation is just exponential form. So the sequence $x^{(1)}(k)$ can meet the form which is the solution of one order differential equation. The differential equation is that

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = u \quad (4)$$

Eq.(4) is called whitened differential equation of GM(1,1). a is called developing coefficient. It reflects the trend of development of $x^{(1)}$ and $x^{(0)}$. u is called grey action quantity[15]. It reflects the relationship between the data.

Discrediting the Eq.(4), we have

$$\Delta(x^{(1)}(k)) + az^{(1)}(k) = u \quad (5)$$

Where

$$\Delta(x^{(1)}(k)) = x^{(1)}(k) - x^{(1)}(k-1) = x^{(0)}(k) \quad (6)$$

$$z^{(1)}(k) = \frac{1}{2}x^{(1)}(k) + \frac{1}{2}x^{(1)}(k-1) \quad (7)$$

Eq.(5) is a grey differential equation, and also called GM(1,1) model. $z^{(1)}(k)$ is called background value.

Let

$$Y = \begin{pmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{pmatrix}, B = \begin{pmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{pmatrix}, \alpha = \begin{pmatrix} a \\ u \end{pmatrix}$$

Then the Eq.(5) turn to

$$Y = B\alpha \quad (8)$$

In terms of least square method, we have

$$\min_{a,u} \sum_{k=2}^n [x^{(0)}(k) - (-az^{(1)}(k) + u)]^2 \quad (9)$$

And

$$\hat{\alpha} = \begin{pmatrix} \hat{a} \\ \hat{u} \end{pmatrix} = (B^T B)^{-1} B^T Y \quad (10)$$

Substituting the Eq.(10) into the Eq.(4), the time response function of the whitened differential equation is got. That is

$$x^{(1)}(t) = (x^{(1)}(1) - \frac{\hat{u}}{\hat{a}})e^{-\hat{a}(t-1)} + \frac{\hat{u}}{\hat{a}} \quad (11)$$

Then, the time response function of grey differential equation is

$$\hat{x}^{(1)}(k+1) = (x^{(0)}(1) - \frac{\hat{u}}{\hat{a}})e^{-\hat{a}k} + \frac{\hat{u}}{\hat{a}} \quad (12)$$

$$k = 1, 2, \dots, n$$

By inverse accumulative generating operation, the recovered value is shown below

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) = (1 - e^{-\hat{a}})(x^{(0)}(1) - \frac{\hat{u}}{\hat{a}})e^{-\hat{a}k} \quad (13)$$

3. Optimized GM(1,1) Model

3.1. Optimization of Initial Condition

Through original GM(1,1) model, the initial condition which make the time response function be established is $x^{(1)}(1) = x^{(0)}(1)$. But the fitting curve does not necessarily pass through the initial data point for prediction. The point $x^{(0)}(1)$ is an old data; it doesn't have close relationship with the future data. And the latest data have more information and the relationship with future data is close. Thus, let the initial condition be

$$\hat{x}^{(1)}(n) = cx^{(1)}(n) \quad (14)$$

The time response function of grey differential equation becomes

$$\hat{x}^{(1)}(k+1) = (cx^{(1)}(n) - \frac{\hat{u}}{\hat{a}})e^{-\hat{a}(k-n+1)} + \frac{\hat{u}}{\hat{a}} \quad (15)$$

$$k = 1, 2, \dots, n$$

and by inverse accumulative generating operation, the recovered value becomes

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) = (e^{-\hat{a}} - 1)(cx^{(1)}(n) - \frac{\hat{u}}{\hat{a}})e^{-\hat{a}(k-n)} \quad (16)$$

3.2. PSO-GM(1,1) Model based on Least Absolute Criteria

The GM(1,1) model uses least square method to obtain parameters from Eq.(9).and Eq.(10). Least square method is widely used in regression analysis. It is easy to calculate and has good analytical performance. But the error will be expanded when the sequence has singular points by least square method and the accuracy of prediction will be decrease.

The least absolute criteria use the first power of error instead of the quadratic error. Thus the singular points have less influence on the model and error will not be expanded. So the least absolute method has better stability than least square method, and also has better performance on statistics. The least absolute criteria use absolute error to describe the deviation, that is

$$\min_{a,u} \sum_{k=2}^n |x^{(0)}(k) - (-az^{(1)}(k) + u)| \quad (17)$$

Since Eq.(17) is non-smooth and non-analytical, it is difficult to solve the parameters. Thus intelligent optimization algorithms are considered. In this paper, particle swarm optimization is adopted to optimize parameters.

Kennedy and Eberhart [16] proposed particle swarm optimization(PSO) in 1995. At first some scientists studied the behavior of a flock of birds. They found that a huge group of birds could change direction, scatter and restructure group, *etc.*, when they were flying and focused on the processing of individual distance. Then Kennedy and Eberhart amended the model,

making particles fly to the solution space and land on the optimal solution. And the flow chart of PSO algorithm is shown in Figure 1.

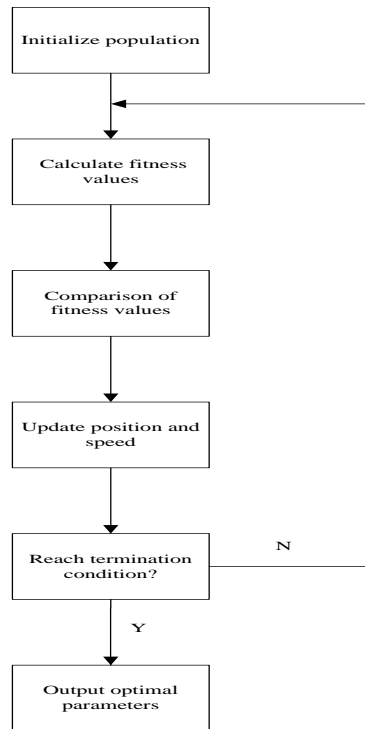


Figure 1. Flow Chart of PSO Algorithm is Shown

The principle of PSO is as follows. Set the number of particles is N . The particles fly in the d -dimensional space in a certain speed. Every particle considers the historical best point of itself and the historical best point of group when they search in the space. And update the position of particles on the basis of historical best point of itself and group.

Let the position of particle i is

$$x_i = (x_{i1}, x_{i2}, \dots, x_{id})$$

the speed of particle i is

$$v_i = (v_{i1}, v_{i2}, \dots, v_{id})$$

the historical best point of particle i is

$$p_i = (p_{i1}, p_{i2}, \dots, p_{id})$$

the historical best point of group is

$$p_g = (p_{g1}, p_{g2}, \dots, p_{gd})$$

The position and speed of particle change rule is as follows

$$v_{id}^{k+1} = \omega v_{id}^k + c_1 \eta_1 (p_{id}^k - x_{id}^k) + c_2 \eta_2 (p_{gd}^k - x_{id}^k) \quad (18)$$

$$x_{id}^{k+1} = x_{id}^k + v_{id}^{k+1} \quad (19)$$

Where c_1 and c_2 are called learning factor. Generally, $c_1 = c_2 = 2 \cdot \eta_1$ and η_2 are random number which are subjected to uniform distribution $U[0,1]$. The maximum speed of particles is v_{max} . In order to make the particle have better exploration ability at the early stage and have better development ability at the later stage, the weight coefficient is set to

$$\omega = \omega_i = \omega_{max} - \frac{\omega_{max} - \omega_{min}}{Iter_max} * i \quad (20)$$

Where $Iter_max$ is maximum iterations. $\omega_{max} = 0.9$, $\omega_{min} = 0.4$.

The optimized initial condition of PSO-GM(1,1) algorithm based on least absolute criteria is as follows:

Step 1: The population size of particles is set to $N = 50$ and The particle is set to two-dimensional vector $x = (x_1, x_2)$. (x_1, x_2) is stand for (a, u) . Initialize position and speed of particles randomly.

Step 2: Calculate the fitness values of every particles. The fitness function is

$$F = \sum_{k=2}^n |x^{(0)}(k) - (-az^{(1)}(k) + u)| \quad (21)$$

Step 3: For each particle, compared its fitness value with historical best fitness value of individual. If the current fitness value of individual is lower, it is set to the historical best fitness value of individual particle and the individual historical best position is updated by the current position.

Step 4: For each particle, compared its historical best fitness value with the group historical best fitness value. If the historical fitness value of individual is lower, it is set to the historical fitness value of group and the group historical best position is updated by the individual best position.

Step 5: Update the speed and position of particles according to Eq.(18) and Eq.(19).

Step 6: Determine whether the algorithm reaches termination condition. If doesn't, turn to Step 2. The termination condition is $Iter_max = 1000$.

After the process above, the optimal parameters a and u will be found. Then we need to solve parameter c of initial condition. Hence, the same method is adopted while the fitness function is

$$F^* = \sum_{k=2}^n |x^{(1)}(k) - \hat{x}^{(1)}(k)| \quad (22)$$

Then PSO algorithm is also used to find the optimal parameter c .

4. Model Construction and Prediction

We construct the optimized GM(1,1) model based on least absolute criteria by Matlab and the data is the whole social electricity consumption of China from 1980 to 2010. Here the data from 1980 to 2006 are used to construct the model. The others which are from 2007 to 2010 are used to predict. Then the original GM(1,1) model, GM(1,1) model with latest initial condition and optimized GM(1,1) model proposed in this paper are established for comparative analysis. The results are shown as follows.

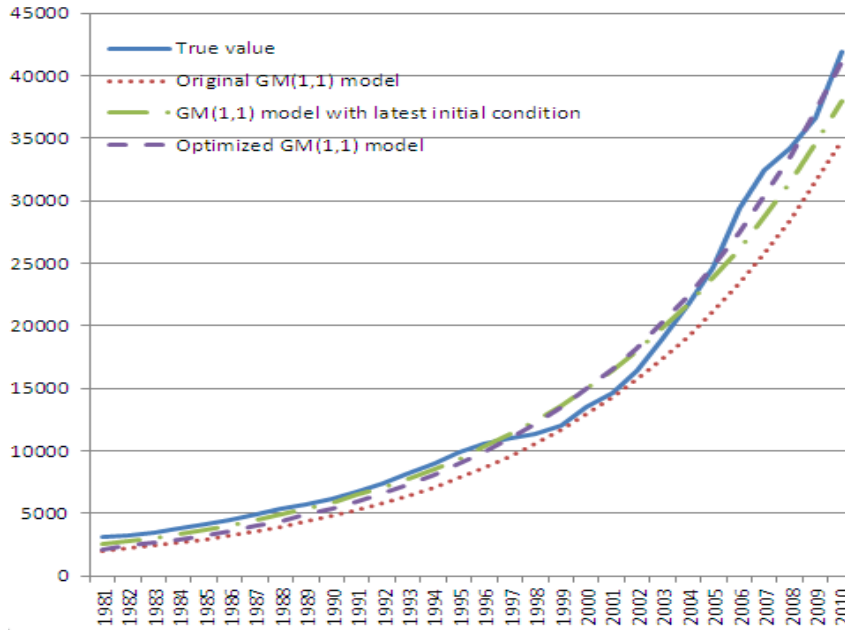


Figure 2. The True Value and the Prediction Curves of Different Methods (Hundred Million KWH)

Table 1. The Prediction Results of Different Methods from 1981 to 2006

Time	True value	Original GM(1,1) model	GM(1,1) with latest initial condition	Optimized GM(1,1) model
1981	3095.7	1966.47	2544.301	2166.575
1982	3280.1	2171.259	2792.93	2398.161
1983	3518.7	2397.373	3065.856	2654.501
1984	3777.6	2647.036	3365.451	2938.242
1985	4117.6	2922.698	3694.323	3252.312
1986	4429.04	3227.068	4055.333	3599.953
1987	4902.69	3563.135	4451.62	3984.753
1988	5358.65	3934.2	4886.633	4410.685
1989	5761.98	4343.907	5364.155	4882.145
1990	6125.96	4796.282	5888.341	5404
1991	6696.79	5295.767	6463.75	5981.635
1992	7455.39	5847.268	7095.388	6621.015
1993	8201.08	6456.202	7788.75	7328.738
1994	9046.49	7128.551	8549.867	8112.11
1995	9886.36	7870.919	9385.36	8979.217
1996	10570.29	8690.596	10302.5	9939.009
1997	11039.11	9595.635	11309.26	11001.39
1998	11347.3	10594.92	12414.4	12177.34
1999	12092.28	11698.28	13627.54	13478.98

2000	13466.22	12916.54	14959.22	14919.75
2001	14682.51	14261.67	16421.04	16514.53
2002	16386.28	15746.88	18025.7	18279.78
2003	18891.21	17386.76	19787.18	20233.71
2004	21761.3	19197.41	21720.78	22396.5
2005	24688.54	21196.63	23843.34	24790.47
2006	29368	23404.05	26173.31	27440.34

Table 2. The Total Prediction Accuracy of Different Methods

	Original GM(1,1) model	GM(1,1) with latest initial condition	Optimized GM(1,1) model
RMSE	2876.4973	1490.3345	1078.7505
MAPE	0.1892229	0.0804278	0.1158562
a	-0.093235	-0.093235	-0.101555
u	1963.4871	1963.4871	1014.188
c	--	--	1.0044508

Where $RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}$, $MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right|$, \hat{y}_i is prediction value, y_i is true value.

The initial condition of GM(1,1) model with latest initial condition is $\hat{x}^{(1)}(n) = x^{(1)}(n)$.

From Figure 2, these three methods all have a good fitting precision in early years. However, the fitting and prediction values of original GM(1,1) model become more and more far away from true values with the passage of time. From Table 2, the root mean square error (RMSE) and mean absolute percentage error (MAPE) of original GM(1,1) model are the biggest in three methods. In addition, RMSE of optimized GM(1,1) model proposed in this paper are lower than GM(1,1) model with latest initial condition, but MAPE of optimized GM(1,1) model proposed in this paper are higher than GM(1,1) model with latest initial condition. It only shows that the GM(1,1) model with latest initial condition has better performance in fitting precision to some extent. It doesn't mean that this method is better in prediction.

Table 3. Prediction Value of Different Methods

Time	True value	Original GM(1,1) model	GM(1,1) with latest initial condition	Optimized GM(1,1) model
2007	32458	25841.3463	28730.96543	30373.44458
2008	34268	28532.4641	31538.55745	33620.07504
2009	36595	31503.8349	34620.50756	37213.73921
2010	41923	34784.64426	38003.62606	41191.53168

Table 4. Prediction Accuracy of Different Methods

	Original GM(1,1) model	GM(1,1) with latest initial condition	Optimized GM(1,1) model
RMSE	6195.81639	3185.936117	1191.9587
MAPE	0.170155155	0.085480337	0.029371599

The prediction values and prediction accuracy of three methods are shown in Table 3 and Table 4. According to RMSE and MAPE, optimized GM(1,1) model based on least absolute criteria is the best of three methods in prediction accuracy on the whole.

Table 5. Relative Error of Different Methods

	Original GM(1,1) model	GM(1,1) with latest initial condition	Optimized GM(1,1) model
2007	-20.3853%	-11.4826%	-6.4223%
2008	-16.7373%	-7.9650%	-1.8908%
2009	-13.9122%	-5.3955%	1.6908%
2010	-17.0273%	-9.3490%	-1.7448%

Where relative error is $rel = \frac{\hat{y}_i - y_i}{y_i}$, \hat{y}_i is prediction value, y_i is true value.

From Table 5, we can clearly find that the optimized GM(1,1) model has the best performance in the relative error of three methods. The prediction accuracy is improved in each year. So maybe the fitting accuracy of optimized GM (1,1) model is not very ideal in the beginning. But it can catch the change tendency of the data better in the future. So the optimized GM(1,1) model based on least absolute criteria can provide a higher precision of prediction.

5. Conclusion

In this paper, we propose an optimized GM(1,1) model based on least absolute criteria. In the original GM(1,1) model, the initial condition is not very suitable because the fitting curve doesn't have to pass through the initial point. And we use the modified latest data which generating from the accumulative generating operation as the new initial condition. Though least square criteria are easy to calculate and widely used, it is not very reliable and will lead the accuracy of prediction to expand. So the least absolute criteria are used to overcome the problem. In addition, PSO algorithm is applied to the parameters optimization. At the end, the prediction results show that the optimized GM(1,1) model based on least absolute criteria improves the prediction accuracy by comparing with the original model and the GM(1,1) model with latest initial condition.

There also exist some improvements, such as improvement of background value and a better way to improve initial condition *etc.* These problems will be researched in the future in order to enhance the prediction accuracy.

Acknowledgements

This research is conducted with the support of National Natural Science Foundation of China(71071045) and the national key project of “863” plan(2011AA05A116).

References

- [1] D. Zafer and C. H. Lester, “Industrial electricity demand for Turkey: A structural time series analysis”, *Energy Economics*, vol. 33, no. 3, (2011), pp. 426-436.
- [2] A. Azadeh, M. Saberi, A. Gitiforouz and Z. Saberi, “A hybrid simulation-adaptive network based fuzzy inference system for improvement of electricity consumption estimation”, *Expert Systems with application*, vol. 36, no. 8, (2009), pp. 11108-11117.
- [3] K. Kavaklioglu, “Modeling and prediction of Turkey’s electricity consumption using Support Vector Regression”, *Applied Energy*, vol. 88, no. 1, (2011), pp. 368-375.
- [4] K. F. TSO Geoffrey and K. W. Yau Kelvin, “Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks”, *Energy*, vol. 32, no. 9, (2007), pp. 1761-1768.
- [5] D. Julong, “Control problems of grey systems”, *Systems & Control Letters*, vol. 1, no. 5, (1982), pp. 288-294.
- [6] X. Xinping, G. Huan and M. Shuhua, “The modeling mechanism, extension and optimization of grey GM (1, 1) model”, *Applied Mathematical Modelling*, vol. 38, no. 5–6, pp. 1896-1910.
- [7] Z. Wei and H. JianMin, “Generalized GM (1, 1) model and its application in forecasting of fuel production”, *Applied Mathematical Modelling*, vol. 37, no. 9, pp. 6234-6243.
- [8] H. Changjun, T. Zhengqi, Z. Qingshan and C. Yuanzhi, “Application of Improved Error GM (1, 1) Model on Predicting of Cultivated Land in Yiyang”, *Energy Procedia*, vol. 5, pp. 1172-1176.
- [9] T. Tzuli, “A research on the grey prediction model GM(1,n)”, *Applied Mathematics and Computation*, vol. 218, no. 9, pp. 4903-4916.
- [10] Z. Dahai, J. Shifang and S. Kaiquan, “Theoretical Defect of Grey Prediction Formula and Its Improvement”, *Systems Engineering-theory & Practice*, vol. 22, no. 8, (2002), pp. 140-142.
- [11] W. Yuhong, D. Yaoguo and L. Yueqing, “An approach to increase prediction precision of GM(1,1) model based on optimization of the initial condition”, *Expert Systems with Applications*, vol. 37, no. 8, (2010), pp. 5640-5644.
- [12] S. Chisheng, H. Yentseng and J. Yeh, “Grey number prediction using the grey modification model with progression technique”, *Applied Mathematical Modelling*, vol. 35, no. 3, (2011), pp. 1314-1321.
- [13] W. Zhengxin, D. Yaoguo and L. Sifeng, “Optimization of Background Value in GM(1,1) Model”, *Systems Engineering - Theory & Practice*, vol. 28, no. 2, (2008), pp. 61-67.
- [14] W. Xiaojia, “Application Research on Electricity Demand Forecasting Based on Gaussian Quadrature Formula”, *Procedia Engineering*, vol. 15, (2011), pp. 5574-5578.
- [15] Y. Tianxiang, L. Sifeng and X. Naiming, “On the properties of small sample of GM(1,1) model”, *Applied Mathematical Modelling*, vol. 33, no. 4, (2009), pp. 1894-1903.
- [16] J. Kennedy and R. C. Eberhart, “Particle swarm optimization”, *Proc. IEEE International Conference on Neural Networks*, Perth, Australia, vol. 4, (1995), pp. 1942-1948.
- [17] H. Chechiang and C. Chiayon, “Applications of improved grey prediction model for power demand forecasting”, *Energy Conversion and Management*, vol. 44, no. 14, (2003), pp. 2241-2249.
- [18] L. Chintsai and Y. Shihyu, “Forecast of the output value of Taiwan's opto-electronics industry using the Grey forecasting model”, *Technological Forecasting and Social Change*, vol. 70, no. 2, (2003), pp. 177-186.

Authors

Zhang Zheng, he was born in 1987 and pursues a doctorate of management science and engineering at school of management of Hefei University of Technology in China. His research interests are prediction and data mining.

Yang Shanlin, he was born in 1948. He is a professor and Ph.D. supervisor in school of management of Hefei University of Technology in China. His research work is mainly engaged in the field of decision science, artificial intelligence, information management and information system.

Liu Huizhou, he was born in 1985 and received his master degree from school of management of Hefei University of Technology in China in 2012. He is an engineer of Tongling Power Supply Company and his research interests are Power information system and computer network.

Yu Bengong, he was born in 1971. He is an associate professor in school of management of Hefei University of Technology in China. His research interests are computer network, information system, data structure and database.

