International Journal of Hybrid Information Technology
Vol.8, No.6 (2015), pp.381-390
http://dx.doi.org/10.14257/ijhit.2015.8.6.37

Predicting Inbound Tourism Demand with Optimized GM (1, 1) Model

Shasha Yu

Economic school, Anyang Normal University, Henan, China
E-mail:yushasha721@hotmail.com

Abstract

The grey model theory is widely used in many field of investigation and of course includes inbound tourism demand. The original GM (1, 1) model couldn't have accuracy prediction in some situation. Thus, an optimized GM (1, 1) model is proposed in this paper. Aiming at the deficiencies of the model, the improvements of GM (1, 1,) are include initial sequence, background value and parameter optimization. At last, the optimized GM (1, 1) model is used to predict inbound tourism demand of China and the results show that the proposed model is better than original GM (1, 1) model and time series model on prediction accuracy.

Keywords: *Prediction, inbound tourism demand, GM (1, 1), PSO, cosine transform*

1. Introduction

Prediction is a necessary part of the decision-making process and it is also an important way to make decision-making more scientific. In the process of world tourism economic development, application of appropriate prediction methods which is to obtain development of tourism economy could provide the scientific basis for decision-making of tourism economy. It is necessary to realize the sustainable development of Tourism and can also avoid economic and environmental loss caused by decision fault. Especially in tourism development planning, inbound tourism demand is very important. It is related to the investment in fixed assets, such as tourism infrastructure, service facilities, and ancillary facilities and so on. And whether civil engineering is a dequateandreception capacity reaches standard could also have critical influence on input and output of the whole process of Tourism.

Tourism demand modeling and forecasting research began in 1960s. Now, it has become one of the hot tourism researches. Many researches at home and abroad take up with this study too [1-4]. In the early 1960's to 1990's, tourism demand forecasting method is based on time series technology [5-12], mainly including the exponential smoothing method, autoregressive moving average model, the trend analysis method etc. Since the beginning of 1990s, many researchers began to study the application of econometric methods.

The grey system theory is proposed by Professor Deng Julong [13] in 1982.After 20 years of development; it has formed a theoretical system which the main content is system analysis, information processing, modeling and prediction, decision making control system. The grey model is applied in various fields and pretty of researches try to improve the model. Wu *et al.*, [14] proposed a new grey system model with the fractional order accumulation is put forward and the priority of new information can be better reflected when the accumulation order number becomes smaller in the in-sample mode. Xia and Wong [15] present a new seasonal discrete grey forecasting model based on cycle truncation accumulation with amendable items to improve sales forecasting accuracy. And the proposed model is practical for fashion retail sales forecasting with short historical data and outperforms other state-of-art forecasting techniques. He *et al.*, [16] proposed a

ISSN: 1738-9968 IJHIT

novel secondary-diagonal mean transformation Partial GM (mtPi-GM) in order to improve the simulation and prediction accuracy of GM matrix solving and the comparison results show that dtP-GM can obtain a better accuracy for periodic data sequences.

And there are also many researches using grey theory in tourism demand prediction [17-20]. But the results show that some models could reach the prediction accuracy while some couldn't. Thus, it is necessary to improve grey model.

In this paper, an optimized GM (1, 1) is proposed. Compared with original model, three aspects are optimized, include initial sequence, and background value and parameter optimization. At last, the proposed model is adopted to predict inbound tourism demand of China.

2. Original GM (1, 1) Model

Set the non-negative sequence is

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

Then the first-order accumulated generating operation sequence is

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

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

Let the background value series be

$$z^{(1)}(\mathrm{k}) = \frac{1}{2}[x^{(1)}(k-1) + x^{(1)}(k)], k = 2, 3, \cdots, \mathrm{n} \tag{3}$$

The grey differential equation is constructed with $x^{(1)}$ by Eq. (4)

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

Which a is called development coefficient. u Is called grey action variable.

The discrete form of Eq is.

$$x^{(0)}(\mathrm{k}) + az^{(1)}(\mathrm{k}) = u \tag{5}$$

Since $\frac{\Delta x}{\Delta t} = \frac{x^{(1)}(k+1) - x^{(1)}(k)}{k+1-k} = x^{(1)}(k+1) - x^{(1)}(k) = x^{(0)}(k+1)$ (6)

The matrix expression of GM (1, 1) is

$$Y_n = BA \tag{7}$$

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

The parameters is solved by least squares method

$$\hat{A} = \begin{pmatrix} \hat{a} \\ \hat{u} \end{pmatrix} = (B^T B)^{-1} B^T Y_n \tag{8}$$

Put parameters into grey differential equation (4), the solution is

$$x^{(1)}(t) = [x^{(1)}(1) - \frac{\hat{u}}{\hat{a}}]e^{-\hat{a}t} + \frac{\hat{u}}{\hat{a}} \tag{9}$$

The discrete form of solution is

$$x^{(1)}(k+1)=[x^{(0)}(1)-\frac{\hat{u}}{\hat{a}}]e^{-\hat{a}k}+\frac{\hat{u}}{\hat{a}}, \quad k=1,2,\cdots \tag{10}$$

Restore the first-order accumulated generating operation, we will get the prediction form.

$$\hat{x}^{(0)}(k+1)=\hat{x}^{(1)}(k+1)-\hat{x}^{(1)}(k)=(1-e^{\hat{a}})(x^{(0)}(1)-\frac{\hat{u}}{\hat{a}})e^{-\hat{a}k}, \ k=1,2,\cdots \tag{11}$$

3. Optimized GM (1, 1) Model

The GM (1, 1) model also has its shortcomings just like other models. In order to increase the prediction accuracy, the original model should be optimized by some methods. In the following, the GM (1, 1) model is optimized from three aspects: Initial sequence, background values and parameters optimization.

3.1. The Smoothness Optimization of Initial Sequence

Properties of smooth discrete function are very important to grey model. The smoothness optimization of initial sequence is the effective method to improve the accuracy of grey model [13]. In this paper, the cosine transform is used to improve the smoothness of original sequence.

Set the original sequence with standardization is

$$Y^{(0)}(\mathrm{k})=\{y^{(0)}(1),y^{(0)}(2),\cdots,y^{(0)}(\mathrm{n})\} \tag{12}$$

Which $y^{(0)}(\mathrm{i})\in(0,\frac{\pi}{2})$, $i=1,2,\cdots$

Then the sequence is processed by cosine transform, that is

$$X^{(0)}=\{x^{(0)}(1),x^{(0)}(2),\cdots,x^{(0)}(n)\} \tag{13}$$

Which $x^{(0)}(k)=\cos y^{(0)}(k)$, $i=1,2,\cdots$

Then the first-order accumulated generating operation sequence is

$$X^{(1)}=\{x^{(1)}(1),x^{(1)}(2),\cdots,x^{(1)}(n)\} \tag{14}$$

Which $x^{(1)}(k)=\sum_{i=1}^{k}x^{(0)}(i)$. $k=1,2,\cdots$

And we use the new sequence to establish grey model, the matrix expression of GM (1,1) is

$$Y_n=BA \tag{15}$$

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

The parameters is solved by least squares method

$$\hat{A}=\begin{pmatrix} \hat{a} \\ \hat{u} \end{pmatrix}=(B^TB)^{-1}B^TY_n \tag{16}$$

And the discrete form of solution is

$$x^{(1)}(k+1) = [x^{(0)}(1) - \frac{\hat{u}}{\hat{a}}]e^{-\hat{a}k} + \frac{\hat{u}}{\hat{a}}, \quad k = 1, 2, \cdots \tag{17}$$

Restore the first-order accumulated generating operation, we will get the prediction form.

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) = (1 - e^{\hat{a}})(x^{(0)}(1) - \frac{\hat{u}}{\hat{a}})e^{-\hat{a}k}, \quad k = 1, 2, \cdots \tag{18}$$

At last, let

$$y^{(0)}(k) = \arccos \hat{x}^{(0)}(\mathrm{k}) \tag{19}$$

3.2. Background Optimization

From original GM (1, 1) modeling, we can get that

$$\left.\frac{dx^{(1)}}{dt}\right|_{t=k} = x^{(1)}(k) - x^{(1)}(k-1) \tag{20}$$

And from the Lagrange mean value theorem, we can also get that

$$x^{(1)}(k) - x^{(1)}(k-1) = F'(k-\xi), \xi \in (0,1) \tag{21}$$

Actually, it is not reasonable that use derivative of $k-\xi$ instead of derivative of k. The background value is also not reasonable by being set as the mean value (Figure 1).

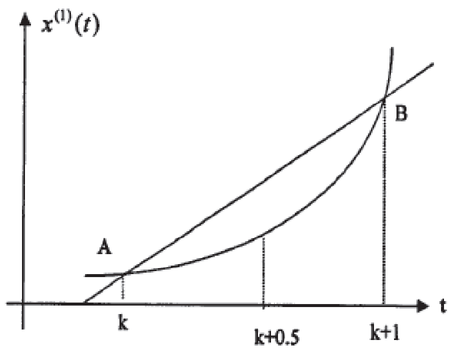

Figure 1. Background Value

So the background value should be optimized. Assume there is a system like

$$x(t) = Ae^{at} + B \tag{22}$$

The discrete form is

$$x(k) = Ae^{a(k-1)} + B \tag{23}$$

Set $\left.\frac{dx^{(1)}(t)}{dt}\right|_{t=x} = x^{(1)}(k+1) - x^{(1)}(k)$, so

$$A\alpha e^{ax} = Ae^{ak} - Ae^{a(k-1)} \tag{24}$$

The solution of Eq. (24) is

$$x = k - 1 + \frac{\ln(\frac{e^{\alpha} - 1}{\alpha})}{\alpha} \tag{25}$$

Which $e^{\alpha} = \frac{x^{(1)}(k+1) - x^{(1)}(k)}{x^{(1)}(k) - x^{(1)}(k-1)} = \frac{x^{(0)}(k+1)}{x^{(0)}(k)}$, $\alpha = \ln \frac{x^{(0)}(k+1)}{x^{(0)}(k)}$

Then we will have

$$x = k - 1 + \frac{\ln(\frac{x^{(0)}(k+1)/x^{(0)}(k) - 1}{\ln x^{(0)}(k+1) - \ln x^{(0)}(k)})}{\ln x^{(0)}(k+1) - \ln x^{(0)}(k)} \tag{26}$$

Set the point of Eq. (26) be point C. Thus, it is better that derivative of $x^{(1)}(t)$ at time C is set as the slope of line AB.

When the sequence changes greatly, the background value which is solved by mean value will cause large error. Here, the weighted method is used to solve the background value.

Assume

$$Ae^{a(k-1+\frac{\ln(\frac{e^{\alpha}-1}{\alpha})}{\alpha})} + B = \mathrm{m}(Ae^{a(k-1)} + B) + n(Ae^{ak} + B) \tag{27}$$

Which $m + n = 1$, $m = \frac{e^{\alpha}}{e^{\alpha} - 1} - \frac{1}{a}$, $n = \frac{1}{a} - \frac{1}{e^{\alpha} - 1}$, $\alpha = \ln \frac{x^{(0)}(k+1)}{x^{(0)}(k)}$

Thus, the background value is

$$\begin{aligned} z^{(1)}(\mathrm{k}) = {} & \left(\frac{\frac{x^{(0)}(k+1)}{x^{(0)}(k)}}{\frac{x^{(0)}(k+1)}{x^{(0)}(k)} - 1} - \frac{1}{\ln x^{(0)}(k+1) - \ln x^{(0)}(k)}\right) x^{(1)}(\mathrm{k}-1) \\ & + \left(\frac{1}{\ln x^{(0)}(k+1) - \ln x^{(0)}(k)} - \frac{1}{\frac{x^{(0)}(k+1)}{x^{(0)}(k)} - 1}\right) x^{(1)}(\mathrm{k}) \end{aligned} \tag{28}$$

3.3. Parameters Optimization with PSO

From the original GM (1, 1) model, we find that the parameters a , u will have a great influence on model performance. In this paper, the Particle Swarm Optimization (PSO) algorithm will used to solve the parameters.

Particle Swarm Optimization (PSO) is proposed by Eberhard and Kennedy [21] in 1995. The PSO algorithm also has been applied in a wide variety of applications [22-24]. The algorithm is shown as follows:

The position of lth particle is

$$x_l = (x_{l1}, x_{l2}, \cdots, x_{ls})$$

The speed of lth particle is

$$v_l = (v_{l1}, v_{l2}, \cdots, v_{ls})$$

The history best point of lth particle is

$$p_l = (p_{l1}, p_{l2}, \cdots, p_{ls})$$

The history best point of the group is

$p_g = (p_{g1}, p_{g2}, \cdots, p_{gs})$

The updating formula of particle about speed and position is

$$v_{ij}^{(t+1)} = w v_{ij}^{(t)} + c_1\delta_1(p_{ij}^{(t)} - x_{ij}^{(t)}) + c_2\delta_2(p_{gj}^{(t)} - x_{ij}^{(t)}) \tag{29}$$

$$x_{ij}^{(t+1)} = x_{ij}^{(t)} + v_{ij}^{(t+1)} \tag{30}$$

Where w inertia is weigh c_1 and c_2 is learning factors. δ_1 And δ_2 are random numbers subjects to the uniform distribution in the interval of $[0,1]$.

Generally, the particle should have good exploration ability at the beginning of the flight and have good development ability later. Thus, a time varying weight can be achieved to the target. Let

$$w_i = w_{\max} - \frac{w_{\max} - w_{\min}}{Iter\_\max} * i \tag{31}$$

Where $w_i \in [w_{\min}.w_{\max}]$, $Iter\_\max$ is the maximum number of iterations.

In order to let the particle has better self-learning ability and smaller social learning ability at the beginning and has smaller self-learning ability and better social learning ability later, the learning factors can be described like

$$c_1 = (c_{1b} - c_{1a})\frac{Iter}{Iter\_\max} + c_{1a} \tag{32}$$

$$c_2 = (c_{2b} - c_{2a})\frac{Iter}{Iter\_\max} + c_{2a} \tag{33}$$

Where $c_{1a}, c_{1b}, c_{2a}, c_{2b}$ are initial value and final value of c_1, c_2, $Iter$ is current number of iterations.

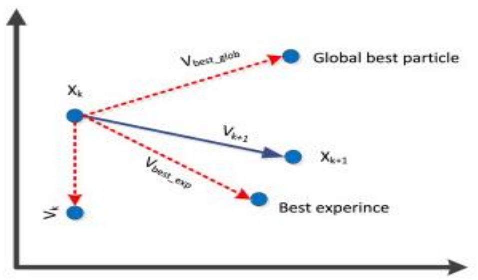

Figure 2. Dynamic of PSO Algorithm

In order to reflect the fact that the recent observations have great influence on prediction value while the later observations have small influence on prediction value, the weighted least squares method is used as the fitness function in PSO algorithm.

The steps of particle swarm optimization algorithm are shown in bellow.

Step 1: Initialize population and the parameters.

Step 2: Update the speed and position of particles.

Step 3: Calculate the fitness values. The fitness function is

$$F = \sum_{k=2}^{n} \alpha^{n-k}\left(x^{(0)}(k) - az^{(1)}(k) - u\right)^2 \tag{34}$$

Which $0 < \alpha < 1$.

Step 4: Update the speed, best position of particle and best position of group.

Step 5: Determine whether the termination condition is reached. If it do not reach the termination condition, turn to Step 2. Otherwise, turn to Step 6.

Step 6: Output the optimal parameters.

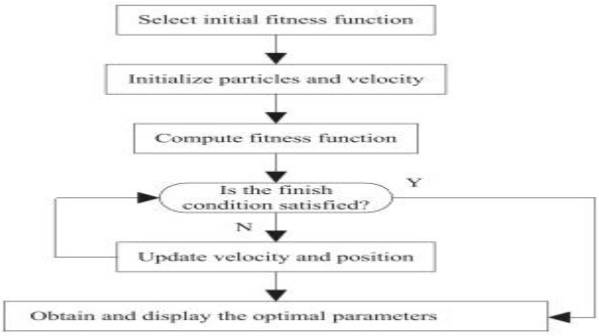

Figure 3. Parameters Optimization of PSO Algorithm

4. Model Construction and Prediction

In this section, the optimized GM (1, 1) model is constructed and it is used to predict inbound tourism demand of China from 1978 to 2011. The inbound tourism demand of China is shown in Fig 4. The unit of inbound tourism demand of China is ten thousand person-times. The data from 1978 to 2006 is used to establish models. Then the models are used to predict inbound tourism demand of China from 2007 to 2011.

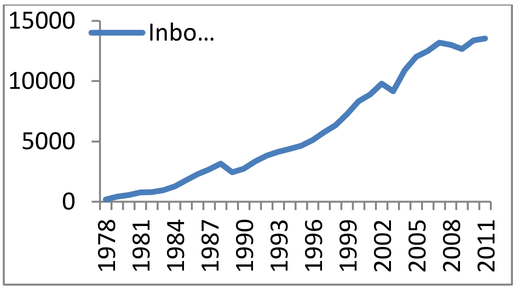

Figure 4. Inbound Tourism Demand of China

From Table 1, we can find that the optimized GM (1, 1) model is better than original GM (1,1) model and ARIMA model on RMSE and MAE. It means that the optimization method of grey model is effective and it is also have better performance than ARIMA model in certain situation.

Table 1. The Performance of Different Models

	Optimized GM(1,1)	Original GM(1,1)	ARIMA
RMSE	0.01458965	0.017413058	0.017297 424
MAE	0.030937182	0.053575704	0.047646 87

Which

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{35}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \tag{36}$$

Table 2. The Prediction Values of Different Models

Year	True value	Optimized GM(1,1)	Original GM(1,1)	ARIMA
2007	13187.33	12441.72	12062.37	12231.09
2008	13002.74	12582.23	12389.76	13622.48
2009	12647.59	12482.13	12221.23	13409.36
2010	13376.22	13087.86	12623.57	13029.11
2011	13542.35	13120.29	12926.83	13974.07

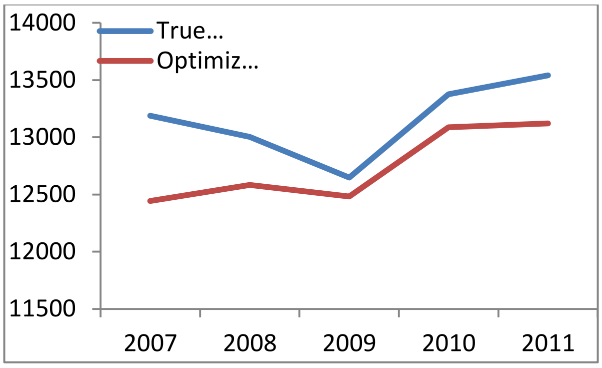

Figure 5. The Prediction Curve of Optimized GM (1, 1) Model

The prediction values of different models in this paper are shown in Table 2. We can see that all these models could give reasonable results. And from Fig 5, we can find that the optimize GM (1, 1) model have the good performance on predicting inbound tourism demand of China from 2007 to 2011.

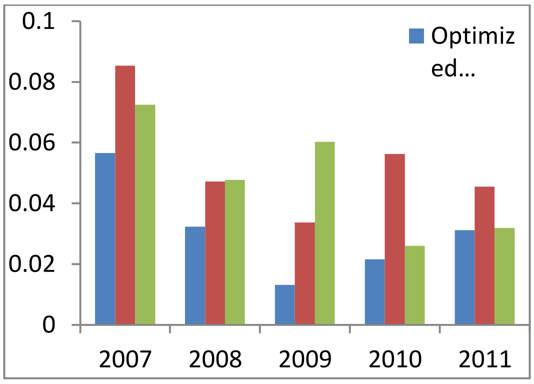

Figure 6. The Absolute Error of Different Models

Table 3. The Absolute Error of Different Models

Year	Optimized GM(1,1)	Original GM(1,1)	ARIMA
2007	0.056539876	0.085306123	0.072511385
2008	0.032340107	0.047142372	0.047662906
2009	0.013082334	0.03371077	0.060231237
2010	0.02155766	0.056267765	0.025949306
2011	0.031165935	0.045451491	0.031879517

Table 3 gives the absolute error of different models on predicting inbound tourism demand of China from 2007 to 2011.From Table 3, the optimized GM (1, 1) shows better than the other models on absolute error in each year. The absolute errors are almost in the range of 5%, except in 2007. And Fig 6 is more intuitively to show the absolute error of different models, we can easily find that the optimized model proposed in this paper is the best model compared with other models on absolute error.

5. Conclusion

In order to increase the prediction accuracy, the optimized GM(1,1) model is proposed. There are three optimized aspects compared with original GM(1,1) model. Firstly, cosine transform is used to improve the smoothness of original sequence. Secondly, the background value is optimized by index method. Thus, the restructured background value is more reasonable. Thirdly, the optimized PSO algorithm is used to search the parameters of GM (1, 1) in order to get better solution. At last, predicting results of inbound tourism demand of China from 2007 to 2011show that the proposed optimized GM(1,1) model is superior to other alternatives.

Acknowledgment

This project is supported by the Joint Funds of the National Natural Science Foundation of China (No.U1404706).

Reference

[1] C. Lim, “Review of international tourism demand models”, Annals of Tourism Research, vol. 24, **(1997)**, pp. 835-849.
[2] G. Li, H. Song and S. Witt, “Time varying parameter and fixed parameter linear AIDS, An application to tourism demand forecasting”, International Journal of Forecasting, vol. 22, **(2006)**, pp. 57-71.
[3] S. F. Witt, C. A. Witt, “Forecasting tourism demand, a review of empirical research”, International Journal of Forecasting, vol. 11, no. 3, **(1995)**, pp. 447-475.
[4] G. I. Crouch, “The study of international tourism demand, review of practice”, Journal of Travel Research, vol. 33, **(1994)**, pp. 41-54.
[5] M. Geurts and I. Ibrahim, “Comparing the Box–Jenkins approach with the exponentially smoothed forecasting model application to Hawaii tourists”, Journal of Marketing Research, vol. 12, **(1975)**, pp. 182–188.
[6] J. H. Kim, “Forecasting monthly tourist departures from Australia”, Tourism Economics, vol. 5, **(1999)**, pp. 277–291.
[7] N. Kulendran and M. L. King, “Forecasting international quarterly tourist flows using error-correction and time-series models”, International Journal of Forecasting, vol. 13, **(1997)**, pp. 319–327.
[8] C. A. Martin and S. F. Witt, “Forecasting tourism demand, a comparison of the accuracy of several quantitative methods”, International Journal of Forecasting, vol. 5, **(1989)**, pp. 7–19.
[9] C. Lim and M. McAleer, “Time series forecasts of international travel demand for Australia”, Tourism

Management, vol. 23, **(2002)**, pp. 389-396.
[10] C. Goh and R. Law, "Modeling and forecasting tourism demand for arrivals with stochastic no stationary seasonality and intervention", Tourism Management, vol. 23, **(2002)**, pp. 499-510.
[11] Y. M. Chan, "Forecasting tourism, a sine wave time series regression approach", Journal of Travel Research, vol. 32, **(1993)**, pp. 58–60.
[12] L. Turner, N. Kulendran and V. Pergat, "Forecasting New Zealand Tourism demand with disaggregated data", Tourism Economics, vol. 1, **(1995)**, pp. 51–69.
[13] J. L. Deng, "Control problem of grey system", Systems and Control Letters, no. 1, **(1982)**, pp. 288-294.
[14] L. F. Wu, S. F. Liu, L. G. Yao, S. L. Yan and D. L. Liu, "Grey system model with the fractional order accumulation", Communications in Nonlinear Science and Numerical Simulation, vol. 18, no. 7, **(2013)**, pp. 1775-1785.
[15] M. Xia and W. K. Wong, "A seasonal discrete grey forecasting model for fashion retailing", Knowledge-Based Systems, vol. 57, **(2014)**, pp. 119-126.
[16] Z. S. He, X. Liu and Y. N, "Chen Secondary-diagonal mean transformation Partial Grey Model based on matrix series", Simulation Modeling Practice and Theory, vol. 26, **(2012)**, pp. 168-184.
[17] C. H. Wang, "Predicting tourism demand using fuzzy time series and hybrid grey theory", Tourism Management, vol. 25, no. 3, **(2004)**, pp. 367-374.
[18] X. Y. Liu, H. Q. Peng, Y. Bai, Y. J. Zhu and L. Liao, "Tourism Flows Prediction based on an Improved Grey GM(1,1) Model", Procedia - Social and Behavioral Sciences, vol. 138, no. 14, **(2014)**, pp. 767-775.
[19] Q. L. Liu, and B. Li, "The Grey Forecasting Model of the International Tourist of Henan", Henan Sciences, vol. 28, **(2010)**, pp. 113-115.
[20] R. Hu, Q. J. Zhuang, L. Zhu and Y. Fu, "Application of improved discrete grey model in medium-long term power load forecasting", Journal of Electric Power Science and Technology, vol. 24, **(2009)**, pp. 50-53.
[21] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory, Micro Machine and Human Science", Proceedings of the Sixth International Symposium, **(1995)**, pp. 39-43.
[22] I. Ziari and A. Jalilian, "Optimal placement and sizing of multiple APLCs using a modified discrete PSO", International Journal of Electrical Power & Energy Systems, vol. 43, no. 1, **(2012)**, pp. 630-639.
[23] M. I. Menhas, L. Wang, M. R. Fei and H. Pan, "Comparative performance analysis of various binary coded PSO algorithms in multivariable PID controller design", Expert Systems with Applications, vol. 39, no. 4, **(2012)**, pp. 4390-4401.
[24] M. A. Mohandes, "Modeling global solar radiation using Particle Swarm Optimization (PSO)", Solar Energy, vol. 86, no. 11, **(2012)**, pp. 3137-3145.

Author

Shasha Yu, born in 1984, she received her PhD in economics (2011) from Kangwon National University (Republic of Korea). Now she is an Assistant Professor in the department of economic at Anyang Normal University (China). Her primary research interests are business management and industrial economy.

