International Journal of Database Theory and Application
Vol.6, No.6 (2013), pp.29-38
http://dx.doi.org/10.14257/ijdta.2013.6.6.03

Base a EMD-Grey Model for Textile Export Time Series Prediction

Hua quanping and Yang xiaoyi

Information Engineering Institute,Zhejiang Textile & Fashion College, Ningbo 315211, China

watchping@163.com

Abstract

The grey forecasting model has been used widely. This paper proposes a hybrid method that combines the GM (1,1) grey forecasting model and the EMD method to forecast textile export time series with seasonality characteristics. Empirical mode decomposition is good at dealing with nonlinear, non-stationary signal and GM (1,1) model can be used to prediction time series. Through the empirical mode decomposition, the textile export series is decomposed into several intrinsic mode function component and a trend component, then using GM (1,1) model to predict these components, and reconstruct the final prediction results. The time series data of the total number of textile export (2003 to 2011) are used as test data sets. Results show that the method has better prediction effect compared with the direct prediction method.

***Keywords**: empirical mode decomposition; grey system; prediction; textile export*

1. Introduction

The grey system theory was initially presented by Deng (1982) [1–3]. The grey forecasting model has been successfully used in finance, physical control, engineering and economics. The grey forecasting model can be used in circumstances with relatively little data and it can use a first-order differential equation to characterize a system. Therefore, only a few discrete data are sufficient to characterize an unknown system.

Reference [4] presents an algorithm of grey model-GM (1,1) to forecast the energy consumption of enterprise. The result shows that grey model-GM (1,1) has higher prediction precision and the trend of energy consumption can be reflected accurately in actual energy consumption forecasting. In reference [5], an improved hybrid optimization model based on grey GM (1, 1) model is proposed to develop the prediction model in power systems. The improved model is defined as RGM (1, 1). Furthermore, Markov chain model is applied to RGM (1, 1) to enhance the prediction performance. Reference [6] introduce local grey SVR (LG-SVR) integrated grey relational grade with local SVR for financial time series forecasting. Pattern search method and leave-one-out errors are adopted for model selection. Experimental results of three real financial time series prediction demonstrate that LG-SVR can speed up computing speed and improve prediction accuracy. Reference [7] makes an evaluation and early warning prediction using grey prediction models and the index of process ability. Reference [8] uses the grey system theory to establish the prediction model. This paper predicted the growth of China's total social logistics costs in 2009 and 2010, and verified the credibility of the prediction model from the angle of qualitative analysis. In reference [9], in order to improve the prediction capability of grey model, the exponential smoothing (ES) method is integrated into GM (1, 1) through the preprocessing for original data set. In the paper [10], a hazard-risk assessment model and a grey hazard-year prediction model (GHYPM)

ISSN: 2005-4270 IJDTA
Copyright ⓒ 2013 SERSC

are constructed by integrating recent advances in the fuzzy mathematics, grey theory and information spread technique, and then applied to 17-year tropical cyclones hazards in Southern China. In the paper [11], the United States dollar to Euro parity is used to compare the performances of the different models. Among the grey models, the modified GM (1,1) using Fourier series in time is the best in model fitting and forecasting. The paper[12]propose a novel grey prediction based particle filter (GP-PF), which incorporates the grey prediction algorithm into the standard particle filter (SPF). The basic idea of the GP-PF is that new particles are sampled by both the state transition prior and the grey prediction algorithm. The grey prediction model presented paper [13] is called first-entry GM (1,1), abbreviated as FGM (1,1), which is based on the existing GM (1,1) but modeled with data including the first-entry's messages of the original series.

This paper combines the advantages of grey and EMD theory to design a grey prediction algorithm. A grey prediction structure is used to obtain some important information. An EMD mechanism is constructed to provide an appropriate forecasting time serial to the grey predictor. From the comparison with the other control structure, it is clear that the proposed structure and method are available.

The remainder of this paper is organized as follows: Section 2 describes the EMD method and the grey system theory. In Section 3, the proposed approach is applied to predict the quantity of textile export. Section 4 describes the evaluation methods used for comparing the performance of forecasting techniques. Section 5 presents conclusions.

2. Theoretical Foundation

The traditional quantitative analysis method need complete statistical data. However, in application, the textile export demand is influenced by so many factors, such as political and legal factors, economic factors, social and cultural factors, transport factors. The quantitative analysis models can not analysis all the related factors, and some factors are not quantifiable. But these factors can have a profound impact on the textile industry, which can not be ignored.

Empirical mode decomposition algorithm decomposes the physical measured sequence into a small number of IMF component and a trend component. The trend component is the remaining component after IMF separated from the original sequence .The trend component is flat and smooth. Empirical mode decomposition is suitable for nonlinear and non-stationary sequences.

Grey system theory is a kind of tool to study and solve the grey system analysis, modeling, prediction, decision making. Grey prediction is the prediction of grey system. At present, some commonly used prediction method (such as regression analysis) need for larger samples. If the sample is smaller, these methods often result in larger error and make the forecast target failure. Grey theory [5], originally developed by Prof .Deng in 1982, has become a very effective method of solving uncertainty problems under discrete data and incomplete information. Here, we give some basic definitions of grey system.

2.1. GM (1,1) Model

Model procedure of GM (1,1) is given as follows:

Grey derivative and grey differential equation is defined based on the relational space and smooth discrete function in grey system theory. And then we use the discrete data sequence to establish dynamic model in the form of differential equations. This model is the basic model of grey system, known as the grey model (GM).

The formula $x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n))$ is set as the original series, and the formula $x^{(1)} = (x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n))$ is its one order accumulated generating sequences, in which $x^{(1)}(k)=\sum_{i=1}^{k}x^{(0)}(i), k=1,2,\cdots,n,$ and the grey derivative of $x^{(1)}$ is defined as $d(k)=x^{(0)}(k)=x^{(1)}(k)-x^{(1)}(k-1)$.

Let $z^{(1)}=(x^{(1)}(2),x^{(1)}(3),\cdots,x^{(1)}(n))$ denote adjacent mean sequences of the sequence $x^{(1)}$, i.e. $z^{(1)}(k)=ax^{(1)}(k)+(1-a)x^{(1)}(k-1)$.

Then GM (1, 1) gray differential equation model can be defined as $d(k)+\alpha z^{(1)}(k)=b$.

Namely $x^{(0)}(k)+\alpha z^{(1)}(k)=b$ (1)

In (1), $x^{(0)}(k)$ is known as the grey derivative; the variable a is known as the coefficient of development; $z^{(1)}(k)$ is called the white background value; b is called the ash dosage.

When $k=2,3,\cdots,n$ is substituted into the formula (1), we get

$$\begin{cases} x^{(0)}(2)+az^{(1)}(2)=b, \\ x^{(0)}(3)+az^{(1)}(3)=b, \\ \cdots\cdots\cdots\cdots \\ x^{(0)}(n)+az^{(1)}(n)=b, \end{cases} \tag{2}$$

After introducing vector matrixes：

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

GM (1, 1) module can be expressed as $Y=\mathbf{B}u$.

Their value is estimated as $\hat{u}=\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix}=(B^T B)^{-1}B^T Y$ by least square method.

As to grey differential equation of GM (1, 1) module, if the time variable $k=2,3,\cdots,n$ is regarded as continuous variable t, then $x^{(1)}(t)$ is the time function of $x^{(1)}$. Grey differential equation of GM(1,1) module of GM(1,1) module can be expressed as $\frac{dx^{(1)}(t)}{dt}+\alpha x^{(1)}(t)=b$.

Taken $x^{(1)}(t=1)=x^{(0)}(1)$ as initial value, we get solution.

$$x^{(1)}(t)=(x^{(0)}(1)-\frac{b}{a})e^{-a(t-1)}+\frac{b}{a} \tag{3}$$

2.2. Empirical Mode Decomposition

The EMD method is under such assumption: any complex signal is composed of simple intrinsic mode function composition. These mode are linear and smooth, or be nonlinear and non-stationary. The essence of EMD method is using the signal’s characteristic scale to obtain intrinsic mode function. Characteristic scale is defined by

Copyright ⓒ 2013 SERSC

adjacent extreme point in the time span. According to the concept of characteristics scale, EMD uses a shift process flow to separating intrinsic mode function from original signal.

The detail steps of EMD are given as follows: for a real signal $x(t)$, first obtain all the maximum points on $x(t)$, and then use all maximum points with three spline curve interpolation and approximate signal $x(t)$'s envelope curve $e_{\max}(t)$. Similarly, we can obtain the signal $x(t)$'s lower envelope curve $e_{\min}(t)$.

Averaging the upper envelope curve $e_{\max}(t)$ and lower envelope curve $e_{\min}(t)$, we get envelope mean curve $m_1(t)$.

$$m_1(t) = \frac{e_{\max}(t) + e_{\min}(t)}{2} \tag{4}$$

Subtracted $m_1(t)$ from $x(t)$, we get the first component:

$$x(t) - m_1(t) = h_1(t) \tag{5}$$

Under ideal conditions, $x(t)$ should be an intrinsic mode. Because the construction process for $m_1(t)$ is to make it meet the intrinsic mode function condition. But the overshoot and undershoot effects on the fitting process could lead to a new extreme points, or translate and enlarge existing extreme point.

Because upper and lower envelope mean curve take part in the shift processing flow so this impact is indirect. But there are still some problems. Even if the fitting process is agree very well, a micro convex or concave point on the slopes of signal, will be enlarged into a new local extreme point. Thus we need to repeat shift process to restore all the low amplitude superposition waves.

The purpose of the shift process is twofold: on the one hand it eliminates the riding waves on signal; on the other hand it makes the waveform profile more symmetrical. such that every IMF has the following two features: (1) the number of the local extreme (maximum or minimum) and the number of zero-crossing point is equal or differ at most one; (2) the average value of the upper envelope which formed by the local maxima and the lower envelope which formed by the local minimum values constitute is zero. The above two characteristics of IMF, is also the EMD decomposition convergence criterion. $h_1(t)$ is regard as the original signal, and repeat the above steps can get the following result:

$$h_1(t) - m_{11}(t) = h_{11}(t) \tag{6}$$

If $h_{11}(t)$ does not meet the IMF conditions, we should continue repeating shift process for k times, until the $h_{11}(t)$ is an IMF. That is: $h_{1(k-1)}(t) - m_{1k}(t) = h_{1k}(t)$.

Define $h_{11}(t)$ as:

$$h_{1k}(t) = c_1(t) \tag{7}$$

$c_1(t)$ is the first IMF component of the original data. As the short period component of the original signal $x(t)$n, $c_1(t)$ is the highest frequency component. When the $c_1(t)$ is isolated from the original signal $x(t)$, we get the residual component $r_1(t)$:

$$x(t) - c_1(t) = r_1(t) \tag{8}$$

 Copyright ⓒ 2013 SERSC

Regarded the $r_1(t)$ as the new signal and repeat the above process, we get the results as follows: $r_{n-1}(t) - c_n(t) = r_n(t)$

In order to extract second intrinsic mode function $c_2(t)$, repeat n times, until the first n intrinsic mode function $c_n(t)$. Finally, because the $r_n(t)$ into a monotone signal, no intrinsic mode function can be extracted. If the decomposed components combined, can reconstruct the original signal $x(t)$:

$$x(t) = \sum_{i=1}^{n} c_i(t) + r_n(t) \tag{9}$$

Then, through the EMD can convert the original data into n intrinsic mode components, and a residual component of $r_n(t)$, the component or a representation of the signal representative of an average trend or is a constant. EMD decomposition process also has the following rules: because the EMD method from the characteristic time scales, the signal characteristic time scale of the smallest modal separation, and then separating the characteristic time scale of larger modal function, the final separation characteristic time scales for the largest component of the EMD method as a group of high pass filter.

3. Empirical Analysis

3.1. Experimental Data

The textile industry is the basis industry of the national economy. To the current national economic growth, production and export of textile product plays an important role. At the same time, the balance of supply and demand of the world textile also plays a vital role in the national economic growth. The textile industry competition is increasingly fierce. China is facing the challenge of textiles with opportunities and risk in the world trade market.

From a lot of literatures we can see that there is no one kind of model and method is the best for prediction. Thus, some simple prediction method can also obtain ideal prediction. We should select a suitable method according to the actual problems in the study.

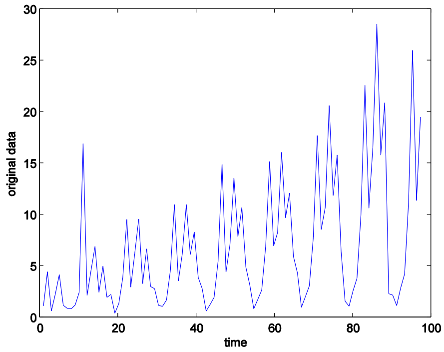
The prediction should base on the analysis of large number of historical data. Most prediction method requires support of at least 5 years to 10 years data. With more data, the result will be more perfect. This paper selects 2003-2011 china monthly total number of export of fabric production as the experimental data (see Table 1, unit of measure is billion dollar).

Table 1. Experimental Data

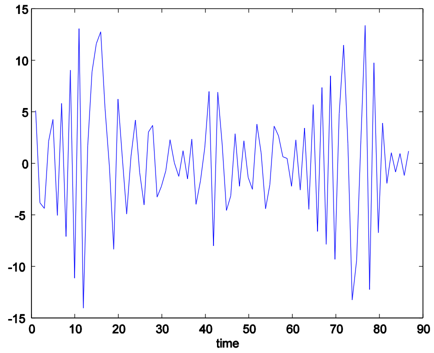
1	2	3	4	5	6	7	8	9	10	11	12
38.1	38.4	35.3	35.6	35.9	35.6	59.9	41.4	43.2	41.6	84.4	43.6
42.3	41.5	42.1	44.4	40	38.2	33	48.5	46.6	45.9	52.1	54.3
48.6	45.5	52	49.5	38.9	41.9	50.5	59.8	58.5	56.5	61.5	60.5
59.4	49.9	50.1	49.6	47.2	25.7	47.3	51.2	48.9	49.2	53.2	51.6

Copyright ⓒ 2013 SERSC

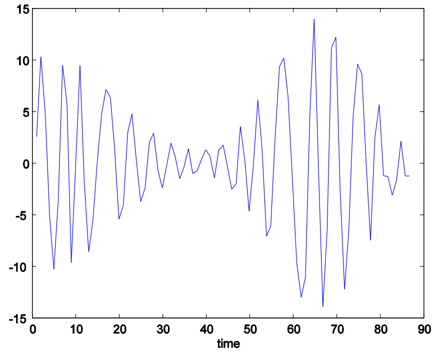
56.7	52.9	53.7	62.1	55.8	45.8	50.5	63.9	69.6	70.9	71.9	66.7
68.2	62.8	70.7	73.7	82.4	40.3	32.7	133.9	86.2	83.5	87	83.6
79.4	74.5	80.2	83	76.8	42.6	85	84.1	87	92.2	84.1	81.2



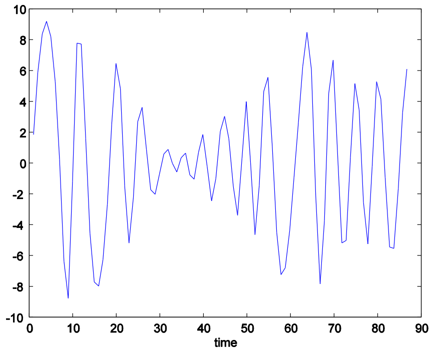
(a)Original Signal



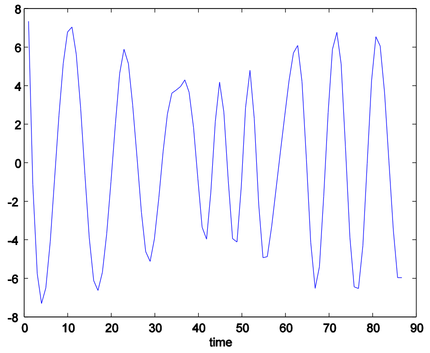
(b)IMF1



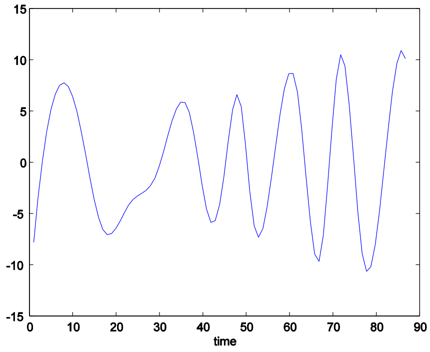
(c) IMF2



(d) IMF3

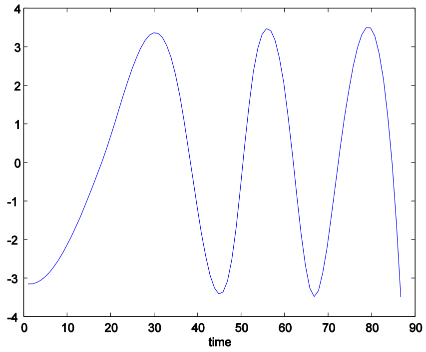


(e) IMF4

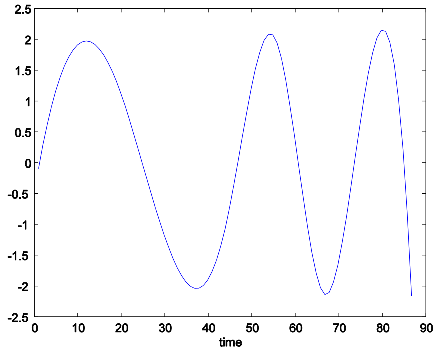


(f) IMF5

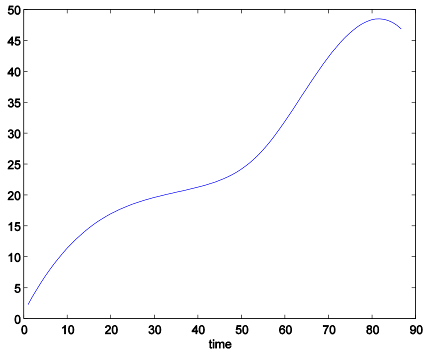
Copyright ⓒ 2013 SERSC



(g) IMF6



(h) IMF7



(i) IMF8

Figure 1. Original Signal and each IMF Component

Figure 1 is the eight intrinsic component decomposed from the original signal. From the amplitude, it can see first intrinsic component has the biggest amplitude, the highest frequency and has the shortest wavelength fluctuations. The intrinsic mode function’s amplitude becomes smaller gradually and wave length becomes longer .The trend component has the lowest frequency.

The reason for intrinsic component is distributed in the above mode is determined by the intrinsic component essence. It is always extracted the most important signals firstly. The first few decomposed intrinsic mode function have the most significant information of data.

3.2. The Accuracy of the Algorithm

In order to facilitate the comparison of the proposed algorithm and the GM (1, 1) algorithm, we establish improved GM (1, 1) model and verify the forecasting precision. Test results are shown in Table 2. The average relative error of the original GM (1, 1) model is 4.0894%. The mean relative deviation of GM (1, 1) combining EMD model is 0.64327%. This reflects that the prediction precision of the proposed model is much higher than that of the original data of GM (1, 1) model.

Table 2 shows that the relative error of the combined model is reduced to a lower magnitude. So for the textile export, the prediction accuracy of model based on the EMD and GM (1, 1) meets the requirements of practical application.

Table 2. Performance Comparison of Algorithms

	original	component 1	component 2	component 3	component 4	component 5	final
error	4.0894%	2.5125%	2.0709%	3.4815%	0.50408%	0.61444%	0.64327%

3.3. Algorithm Effectiveness

The uncertainty system which has less data and poor information is abundant. So GM(1,1) model has very wide application areas. However, this does not mean that GM(1,1) model can be used arbitrarily. Just as any other mathematical model, GM(1,l) model has limitation application range. Beyond its range of application, the GM(1,1) model is difficult to obtain the ideal result. For this reason, we have the follow conclusions. In the formula (3)

1) When -a<0.3, GM(1,l) model can be used for middle and long term forecasting;

2) When 0.3<-a<0.5, GM(1,l) model can be used for short term prediction, prediction of medium and long term;

3) When 0.5<-a<0.8, GM(1,1) short-term predictions should be very cautious;

4) When 0.8<-a<1, we should use the modified GM(1, l) model;

5) When -a>1, it can not predict by GM(1,l) model.

Table 3. Algorithm Effectiveness

	original	component 1	component 2	component 3	component 4	component 5	combination
a	-0.016068	-0.00010688	2.0404e-005	4.9993e-005	-0.0019884	0.001459	-0.002185

Table 3 shows that the combined model is can be used for long term forecasting, just as the original GM(1,1) model.

4. The Further Research Direction and Conclusions

It can be seen from the textile export demand forecasting, the combination forecast method with fixed weights optimization is better than other each single forecast method. Results have improved greatly. However, the method based on the fixed weight coefficient is often unstable. It can not fully reflect the actual situation. To further improvement the method, we should research the variable weight combination forecasting method. So the prediction result will be more close to the actual value.

The traditional GM(1,1) model has demerits in itself, so that it can only be used to deal with the original data sequence which changes exponentially and not too quickly, but for the time series data which fluctuate fiercely, data fitting effect is not ideal. Algorithm using EMD decomposition data to generate grey model can weaken the original data‘s randomness and make the time sequence regularity. Thus the proposed algorithm improves GM(1,1) model prediction accuracy effectively.

Combining the gray prediction and empirical mode decomposition algorithm, this paper put forward a kind of improvement GM(1,1) prediction model. First several intrinsic mode function component and a trend is obtained by the empirical mode decomposition, then we predict each component by GM(1,1) model, and the final prediction results of the method is reconstructed, compared with the algorithm without EMD treatment ,our forecast method has higher forecast accuracy.

 Copyright ⓒ 2013 SERSC

Acknowledgments

This work was supported by Scientific Research Project of Education Department of Zhejiang in China: Y201119385.The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

References

[1] C. -O. Oh, "Evaluating time-series models to forecast the demand for tourism in Singapore comparing within-sample and post sample results", Journal of Travel Research, vol. 43, no. 4, **(2005)**, pp. 404-413.
[2] C. Goh and R. Law, "Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention", Tourism Management, vol. 23, no. 5, **(2002)**, pp. 499-510.
[3] W. Chiang Hong, "The application of support vector machines to forecast tourist arrivals in barbados: an empirical study", International Journal of Management, vol. 23, no. 2, **(2006)**, pp. 375-385.
[4] X. Yu and Z. Lu, "Prediction of energy consumption based on grey model-GM (1,1)", Artificial Intelligence and Computational Intelligence ,Lecture Notes in Computer Science, vol. 7530, **(2012)**, pp. 192-199
[5] G. Dong Lia, S. Masudaa and M. Nagaib, "The prediction model for electrical power system using an improved hybrid optimization model", International Journal of Electrical Power & Energy Systems, vol. 44, no. 1, **(2013)**, pp. 981-987.
[6] H. Jiang, W. He, "Grey relational grade in local support vector regression for financial time series prediction Expert Systems with Applications", vol. 39, no. 3, **(2012)**, pp. 2256-2262.
[7] Y. Wang, J. Tang and W. Cao, "Grey prediction model-based food security early warning prediction", Grey Systems: Theory and Application, vol. 2, no. 1, pp. 13-23.
[8] W. Wei and B. Tang, "Prediction of China's total social logistics costs based on grey model", Advances in Electric and Electronics Lecture Notes in Electrical Engineering, vol. 155, **(2012)**, pp. 505-511.
[9] G.-D. Lia, S. Masudaa and M. Nagaib, "An optimal hybrid model for atomic power generation prediction in Japan", Energy, vol. 45, no. 1, **(2012)**, pp. 655-661.
[10] H. Liua and D.-L. Zhan, "Analysis and prediction of hazard risks caused by tropical cyclones in Southern China with fuzzy mathematical and grey models", Applied Mathematical Modelling, vol. 36, no. 2, **(2012)**, pp. 626-637.
[11] E. Kayacana, B. Ulutasb and O. Kaynaka, "Grey system theory-based models in time series prediction", Expert Systems with Applications, vol. 37, no. 2, **(2010)**, pp. 1784-1789.
[12] J.F. Chen, Z.G.Shi, S.H.Hong, and K.S.Chen, "Grey prediction based particle filter for maneuvering target tracking", Progress in Electromagnetic Research, vol. 93, **(2009)**, pp. 237-254.
[13] T. L. Tien, "A new grey prediction model FGM (1, 1)", Mathematical and Computer Modeling, vol. 49, no. 7–8, **(2009)**, pp. 1416-1426.
[14] H. Jin Hwang and J. Seruga, "An Intelligent Supply Chain Management System to Enhance Collaboration in Textile Industry", IJUNESST, vol. 4, no. 4, **(2011)** December, pp. 47-62.
[15] G. Sundari and P. E. Sankaranarayanan, "Energy-Latency Improved Sensor Networks Using Mobile Agents in Textile Industry", IJFGCN, vol. 4, no. 4, **(2011)** December, pp. 21-30.
[16] V. Golmah and G. Mirhashemi, "Implementing A Data Mining Solution To Customer Segmentation For Decayable Products: A Case Study For A Textile Firm", IJDTA, vol. 5, no. 3, **(2012)** September, pp. 73-90.

Author

Hua Quanping, received his B.Eng degree in Mathematics from Zhejiang Normal University and M.Eng degree in Computer Science and Technology from Southeast University,China in 1992 and 2005 respectively.He is a dean of Department of Information Engineering in Zhejiang Txtile & Fashion College.His current research interests on Intelligence Computation and Database Technology.

Copyright ⓒ 2013 SERSC