http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 12 (22): 7009-7013, 2013
ISSN 1812-5638 / DOI: 10.3923/itj.2013.7009.7013

Research On Business Intelligent Forecasting Method With Time Series

Sen Xin Zhou, Decheng Wu and Chao Li
Management Science And Engineering School of Anhui University
of Finance and Economics, Bengbu Anhui, 233030, China

Abstract: This study present a hybrid forecasting method with time series and intelligent error modification. It takes the historical income data of 19 months of a enterprise as the primary data, respectively using time series and grey forecasting method to train the data and to compare their results. It reveals that time series method is more accurate. However it can not meet the actual requirements. Through analysis, major income was selected from hundreds of items by using principal factor analysis method. Then, time series method was used for training major income respectively, next the large random items of income were operated with intelligent processing technology such as linear regression, neural networks and support vector machine for error modification. The result shows that the method with support vector machines is the best one.

Key words: Time series, Support vector machine, intelligent error modification, forcasting method

INTRODUCTION

Financial forecasting describes the process by which firms think about and prepare for the future. The forecasting process provides the means for a firm to express its goals and priorities and to ensure that they are internally consistent. It also assists the firm in identifying the asset requirements and needs for external financing. For example, the principal driver of the forecasting process is generally the sales forecast. Since most balance sheet and income statement accounts are related to sales, the forecasting process can help the firm assess the increase in current and fixed assets which will be needed to support the forecasted sales level. Similarly, the external financing which will be needed to pay for the forecasted increase in assets can be determined. Firms also have goals related to capital structure (the mix of debt and equity used to finance the firms assets), dividend policy and working capital management (Kumar and Jain, 2010). Therefore, the forecasting process allows the firm to determine if its forecasted sales growth rate is consistent with its desired capital structure and dividend policy.

Being able to forecast time series accurately has been quite a popular subject for researchers both in the past and at present. However, the lack of ability of conventional analysis methods to forecast time series that are not smooth leads the scientists and researchers to resort to various forecasting models that have different mathematical backgrounds, such as artificial neural networks, fuzzy predictors, evolutionary and genetic algorithms (Kayacan *et al*., 2010). A time series is a collection of data points which are generally sampled equally in time intervals. Time series prediction refers to the process by which the future values of a system is forecasted based on the information obtained from the past and current data points. Generally, a pre-defined mathematical model is used to make accurate predictions. Time series prediction models are widely used in financial area, such as predicting stock price indexes, foreign currency exchange rates (FX rates) and so on. (Boubaker and Makram, 2012).

Statistical and artificial intelligence (soft computing) based approaches are the two main techniques for time series prediction seen in the literature. While AR (Auto Regressive), MA (Moving Average), ARMA (Auto Regressive Moving Average), ARIMA (Auto Regressive Integrated Moving Average) and Box-Jenkins models (Box *et al*., 1976) can be mentioned as statistical models. Other intelligent approaches seen in the literature for the analysis of time series include Linear regression, hidden markov models (Boyacioglu *et al*., 2009). Some hybrid models are also seen in the literature: (Chen and Hsiao, 2008) a combination of genetic algorithms and neural networks has been proposed. Cui and Lin (2009) Support Vector Regression (SVR) and a Self-organizing Feature Map (SOFM) technique have been hybridized to reduce the cost of training time and to improve prediction accuracies. De Choudhury *et al*. (2008) FX rates are highly nonlinear, stochastic and highly non-stationary financial time series and as such, it is very difficult to fit a model to them by the use conventional linear statistical methods or artificial neural networks.

Corresponding Author: Sen Xin Zhou, Management Science and Engineering School of Anhui University of Finance and Economics, Bengbu Anhui, 233030, China

The firm has a large scale and rich historical data, in which there hides important decision-making information. Nowadays the competition in the market has become increasingly fierce, so it has become a priority to mine information for scientific decision. Income forecast plays a vital role in the decision-analysis system and provides important decision support for decision-makers to make business development plans, to improve management decisions. This study takes the historical income data of 19 months of a enterprise as the primary data, respectively using time series and grey forecasting method to train the data and to compare their results. It reveals that time series method is more accurate. However it can not meet the actual requirements.Through analysis, major income was selected from hundreds of items by using principal factor analysis method. Then time series method was used for training major income respectively, next the larger random items of income were operated with intelligent processing technology such as linear regression, neural networks and support vector machine for accurate correction. It concluded that the method with support vector machines is the best one. In SPSS Clementine experiment environment, through the historical test data experiment, it was found that the model has high forecast accuracy. Generating a script in SPSS clementine environment and embeding it into the server our forcasting method is run triggering by time event , the result indicated that the model completely meet the needs of the firm and its accuracy was much higher than other statistical methods.

This study is organized into four sections including this introduction. Section 2 gives a comparison of grey forecasting with time series models and Section 3 presents hybrid forecasting method with time series and intelligent error modification. Finally, summary and conclusions are presented in Section 4.

A COMPARISON OF GREY FORECASTING WITH TIME SERIES MODELS

Grey system theory is an interdisciplinary scientific area that was first introduced in early 1980s.Since then, the theory has become quite popular with its ability to deal with the systems that have partially unknown parameters. As a superiority to conventional statistical models, grey models require only a limited amount of data to estimate the behavior of unknown systems.

Grey-forecasting model that adopts the essential part of Grey system theory. GM (1, 1) Grey forecasting can be used in circumstances with relatively little data (n>4). The GM (1, 1) uses a first order differential equation to characterize an unknown system. A GM (1, 1) modelling algorithm is described below:

- **Step 1:** Assume that the original raw data series $y^{(0)}$ with n samples is expressed as:

$$y^{(0)} = [y^{(0)}(1), y^{(0)}(2),..., y^{(0)}(n)], n \geq 4$$

Where the subscript (0) represents the original series. The original data are assumed to be positive. Negative values are prohibited in grey-modelling:

- **Step 2:** A new series y(1) is generated by accumulated generating operation (AGO):

$$y^{(1)} = [y^{(1)}(1), y^{(1)}(2),..., y^{(1)}(n)]$$

Where:

$$y^{(1)}(k) = \sum_{i=1}^{k} y^{(0)}(i),\ k = 1, 2,..., n$$

- **Step 3:** Establish a first order differential equation, where the parameters a_g and u_g are called the development coefficient and grey input, respectively. This grey model is referred as:

$$\frac{dy^{(1)}}{dt} + a_g y^{(1)} = u_g$$

GM (1, 1), in which the first number in the brackets denotes the order of differential equation and second number indicates the number of variables. We define $z^{(1)}(k)$ as a sequence obtained by applying the MEAN operation to $y^{(1)}$:

$$z^{(1)}(k) = \text{MEAN}\, y^{(1)} = \frac{1}{2}\left[y^{(1)}(k) + y^{(1)}(k-1)\right]$$

- **Step 4:** Taking the inverse AGO on sequence by $y^{(1)}(k)$, we have:

$$\hat{y}^{(0)}(n+p) = \hat{y}^{(1)}(n+p) - \hat{y}^{(1)}(n+p-1)$$

$$\hat{y}^{(0)}(n+p) = \left(\hat{y}^{(0)}(1) - \frac{u_g}{a_g}\right).\ e^{-a_g.(n+p-1)}.\ (1 - e^{a_g})$$

This give the predicted values. The results respectively using time series and grey forecasting method are as follows (Fig. 1and 2):

The result of parameter test is shown in Table 1:

The result of parameter test shows that time series model is more accurate than the grey forecasting method. So, we predicted the total income with the time series model.

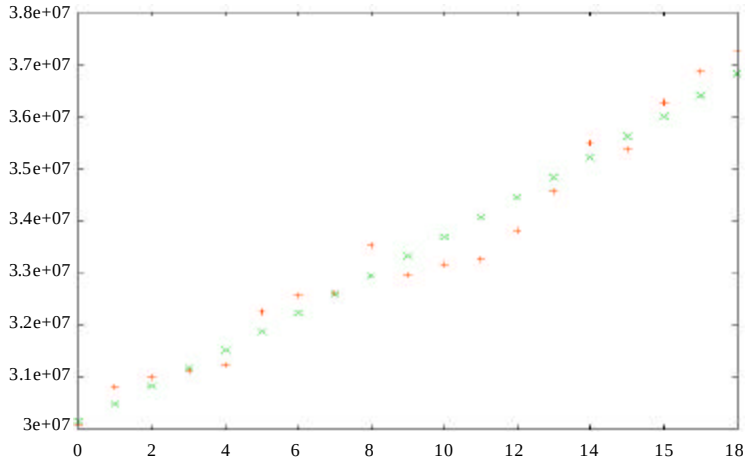

Fig. 1: Result for time serie

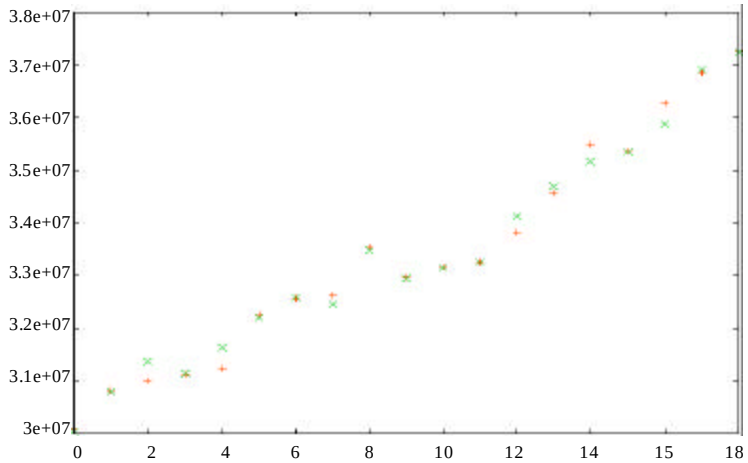

Fig. 2: Result for grey forecasting

Table 1: Result of parameter test

Predicted method	Sequence correlation	The original series variance	Residual variance	Variance ratio
Time series model	0.812122	2063596.238	192307.3834	0.09319
Grey forecasting	0.732618	2063596.238	392925.8012	0.19040

HYBRID FORECASTING METHOD WITH TIME SERIES AND INTELLIGENT ERROR MODIFICATION

Though time series method is more accurate than the grey forecasting method, it can not meet the actual requirements of the firm.Through analysis, the total income was made up of hundreds of items. In order to improve forecasting accurate we selected the major items using principal factor analysis method. Then time series method was used for training major income respectively, next the large random items of income were operated with intelligent processing technology such as linear regression, neural networks and support vector machine for error modification. The forecasting result of the major items with time series model in Table 2 is as follow.

Time is a phenomenon which is both very complex and very important in many real-world problems. Its importance comes from the fact that almost every kind of data contains time-dependent information, either explicitly in the form of time stamps or implicitly in the way that the data is collected from a process that varies with time. A

Table 2: Forecasting result of the major items

		Item (%)					
Accuracy	Month	1	2	3	4	5	6
Fitting accuracy	201001	61.09	97.59	98.93	100.00	100.00	99.61
	201002	52.82	99.37	99.97	100.00	99.49	99.45
	201003	47.39	98.42	99.30	100.00	99.89	98.90
	201004	0.62	98.22	99.93	99.96	96.85	99.66
	201005	9.97	99.89	99.70	99.94	99.72	98.76
	201006	2.56	94.10	99.88	99.92	99.42	90.06
	201007	92.16	99.98	99.95	99.88	99.10	99.84
	201008	93.04	99.99	99.96	99.84	100.00	99.87
	201009	97.36	99.99	99.96	99.79	100.00	99.89
	201010	97.88	100.00	99.97	99.73	100.00	99.90
	201011	98.30	100.00	99.97	99.64	100.00	99.92
	201012	98.50	100.00	99.97	99.53	100.00	99.93
	201101	93.05	99.40	99.01	98.49	98.87	99.59
	201102	91.16	99.33	99.98	75.01	99.47	98.48
	201103	89.94	98.96	99.28	84.67	99.89	97.84
	201104	-37.79	98.90	99.89	95.29	96.54	98.98
	201105	-28.78	99.86	99.67	96.67	99.69	97.80
	201106	79.73	96.81	99.85	99.38	99.38	89.02
Forecasting accuracy	201107	56.65	95.72	98.99	93.77	99.02	69.77

Table 3: Result of intelligent error modification

Month	Accuracy before modification	Modification method	Modification value	Accuracy after modification (%)
201107	99.45%			
		SVM	250717.65	99.88
		ANN	224108.54	99.95
		linear regression	-18628.77	99.40

reason for its complexity is that time can be represented in a multitude of different representations. As always, an effect that is obvious in one representation may be very much hidden in another representation. Time series analysis is an important and complex problem in machine learning and statistics. Real-world applications can consist of very large and high dimensional time series data. Support Vector Machines (SVMs) are a popular tool for the analysis of such data sets. The fundamental concept of an artificial neural network (ANN) is inspired by the neural architecture of the human brain. As a non-parametric model, ANN does not rely on assumptions such as normality and stationarity that are often adopted to make traditional statistical methods tractable. ANN systems learn by example and dynamically modify themselves to fit the data presented. They, also have the ability to learn from very noisy, distorted or incomplete sample data. ANN models have outperformed the traditional statistical models in forecasting stock prices, stock returns, inflation, imports and exchange rate.

Intelligent error modification is made up of two steps. The first step is to compute major items forecasting value of 06/2011, respectively by linear regression model, neural networks model and support vector machine model. Their train data are major items data before 06/2011.The step two is to modify forecasting value of 07/2011 with error value of 06/2011 respectively by linear regression model, neural networks model and support vector machine model. The result of intelligent error modification in Table 3.

CONCLUSION

Computational intelligence is a newly emerging discipline. Obviously, this is a very young topic and the

definition and content of which are still undergoing development and change. Application of computational intelligence in financial computation is very impotant. We present a hybrid forecasting method with time series and intelligent error modification. It takes the historical income data of 19 months of a enterprise as the primary data, respectively using time series and grey forecasting method to train the data and to compare their results. It reveals that time series method is more accurate. However it can not meet the actual requirements.Through analysis, major income was selected from hundreds of items by using principal factor analysis method. Then time series method was used for training major income respectively, next the large random items of income were operated with intelligent processing technology such as linear regression, neural networks and support vector machine for error modification. The result shows that the method with support vector machines is the best one.In SPSS clementine experiment environment the high forecast accuracy model was made and it is applied in practice. Our method can be widely applied to other financial forecasting.

ACKNOWLEDGEMENT

This study is supported by National Nature Science Foundation of china under Project Number: 60473042 and 2009 Major Science Foundation Subject for the Education Department of Anhui Province under Project Number ZD200905.

REFERENCES

Boubaker, A. and B. Makram, 2012. Modelling heavy tails and double long memory in North African stock market returns. J. North Afr. Stud., 17: 195-214.

Box, G.E.P., G. Jenkins and G.C. Reinsel, 1976. Time Series Analysis: Forecasting and Control. 4nd Edn., Wiley, USA.

Boyacioglu, M.A., Y. Kara and O.K. Baykan, 2009. Predicting bank financial failures using neural networks, support vector machines and multivariate statistical methods: A comparative analysis in the sample of Savings Deposit Insurance Fund (SDIF) transferred banks in Turkey. J. Expert Syst. Appl., 36: 3355-3366.

Chen, L.H. and H.D. Hsiao, 2008. Feature selection to diagnose a business crisis by using a real GA based support vector machine: An empirical study. Expert Syst. Appl., 35: 1145-1155.

Cui, M. and Y. Lin, 2009. Nonlinear Numerical Analysis in Reproducing Kernel Space. Nova Science Publishers, New York, ISBN: 9781604564686, Pages: 226.

De Choudhury, M., H. Sundaram, A. John and D.D. Seligmann, 2008. Can blog communication dynamics be correlated with stock market activity. Proceedings of the 9th Conference on Hypertext and Hypermedia, November 5, 2008, Pittsburgh, pp: 55-60.

Government of India, 2007. Ministry of Finance, Economic Division, India 2007. Economic Survey. Oxford University Press, India.

Kayacan, E., B. Ulutas and O. Kaynak, 2010. Grey system theory-based models in time series prediction. Expert Syst. Appl., 37: 1784-1789.

Kumar, U. and V.K. Jain, 2010. Time series models (Grey-Markov, Grey Model with rolling mechanism and singular spectrum analysis) to forecast energy consumption in India. Energy, 35: 1709-1716.