The Forecasts Performance of Gray theory, BP network, SVM for Stock Index

Fengqin Niu, Shuangshuang Nie, Weihong Wang
School of management, Donghua University
nfengqin@126.com, shuang689@163.com

Abstract

In order to evaluate the performance of several forecasts, the paper firstly uses three forecasting methods, namely grey model(GM (1,1)), BP neural networks and support vector machines (SVM), to forecast the Shanghai Industrial Index, the Shanghai Commercial Index, the Shanghai Real Estate Index, the Shanghai Public Utilities Index. Through evaluating the results of these forecasting methods, it is argued that, for forecasting concussing stock market, the forecast effects of BP Network method is obviously better than the forecast effects of GM (1, 1) method and SVM method; GM (1, 1) method and SVM method are more suitable for forecasting stable stock market.

1. Combining forecasts and determination of the forecasting effectiveness

With propellant by factors such as the rapid growth of China's GDP, high growth of the performance of listed companies and strong capital market, China's Shanghai index rushed to the historical high of 6124 points from 2728 points at the beginning of the year 2007. However, several substantial adjustments experienced in the period of year 2007 and year 2008, gave the investors painful lessons, which is unforgettable enough for life-long. Stock Index can reflect changes in stock market timely and comprehensively, and the trend of stock prices can be seen from it. So forecasting Stock Index is of great significance to help investors to understand and make accurate decision-making.

There are two kinds of methods in stock index forecasting. One is the modern statistical methods, which forecast the stock index through model building. The other is artificial intelligence method. Because the factors affecting the stock market is very complex and many factors themselves are fuzzy and confusing, it is very difficult to chose a statistical model to forecast the stock index. Because of its self-organizing and self-learning characteristics, Artificial intelligence thus is considered effective to forecast the nonlinear stock index fluctuation.

Previously, many domestic and foreign scholars have used various methods such as grey theory, neural networks, support vector machines etc. to forecast the stock market.

Grey system theory [1] is founded by Professor Deng Julong who is China's mathematician in 1982. It avoids the complex causal relationship, mines useful information from its own data, research and reveals internal potential law of the running of system in the future through mathematical models and processing operations. Grey model requires less information when modeling, and do not have to know the priori distribution of raw data, and have high precision. It can also maintain the characteristics of the original system and reflect the actual situation. In china, many scholars applied grey theory to the forecasting of Stock market. Panfeng Li [2] applied grey theory to the forecast of stock price. Changfeng Cai[3] selected two stocks' prices of eight day from June 1 to June 10 in the year of 1994, and used GM (1,1) model to forecast. The relative error is around 1.53 percent, it shows that GM (1, 1) model got the higher prediction accuracy and user-friendly conclusion. Guoping Li, Senfa Chen, Guangqing Yu[4] combined grey forecast, K-line theory and other traditional methods of technical analysis together to forecast the stock, and got better results. Guoping Li, Senfa Chen [5] used gold division for optimizing data of Grey forecast, and significantly improve forecasting results.

BP neural network models based on own internal relations of the data. It has a good self-organization, adaptability, a strong ability to learn and anti-jamming ability. It can automatically extract the knowledge of economic activity from the historical data. It can not only overcome many constraints and difficulties which traditional quantitative forecasting methods faced, but also can avoid many artificial factors, so BP neural network has its own unique advantages in rationality and application of construction of the stock market forecasting model. Xiaotian Zhu 、 Hong Wang[6] argued that the forecasting effect of neural networks have improved in short-term forecasts, the medium-term forecasts, long-term projections forecast three different circumstances when considering the volume. In china, Jianchen Long, Xiaoping Li [7] constructed a model of data on stock transactions, and learned network combined with parallel neural network, and forecasted the stock market trends. The simulation results showed the feasibility, and had

978-0-7695-3488-6/08 $25.00 © 2008 IEEE
DOI 10.1109/KAM.2008.161

achieved good effects. Huihui Meng, Deqian Ye, Na Liu[8] designed stock operation support system based on rough set theory and neural networks, which not only reduced the size of the neural network, but also reduce the burden of training and learning through the elimination of redundant.

Support vector machine [9] is a new machine learning algorithms raised by Vapnik by based on the statistical and the principle of structural risk minimization. SVM training algorithm has no problem of local minimum and dimension disaster, with automatic designing Model complexity and generalization ability, and has been successfully used for optimal control, pattern recognition and financial forecasts, and other fields. Kim [10] forecasted the daily direction of change of Korea Composite Stock Price Index (Composite Index) with SVM; the study selected 12 technical indicators, and compared the forecasts of SVM and BP neural network, and concluded that SVM is better than BP neural network in time series forecasting. Kyoung-jae Kim [11] SVM forecasted the stock index and compared forecasting results with the back-propagation neural network method, the results showed that SVM methods forecasted better than the back propagation neural network in the stock market, and SVM Method would have good prospects in the stock market forecasts. In China, Yanjie Shi [12] raised using support vector machines to forecast stock market based on comparison of forecasting methods, and established a mathematical model to forecast Xinjiang Public. The results showed that support vector machine forecasted better than neural network. Lifang Peng, Zhiqing Meng, Hua Jiang and Mi Tian [13] established stock closing price forecasting model using the stock data of Shahe shares. The model overcame limitation to the linear system of the traditional time-series forecasting model. The results showed the method had the highest forecast accuracy than the neural network method and traditional time-series forecasting model, and can be applied to forecasting certain nonlinear time series.

2. The forecasting performance of combining forecasts methods for the Shanghai Stock Index

The following is an empirical study in the Shanghai stock exchange market. Since the stock price fluctuation is strongly influenced by the industry of China, we choose the time series of Shanghai Industrial Index, Shanghai Commercial Index, Shanghai Real Estate Index and Shanghai Public Utilities Index respectively to do the research work. Firstly, we use three forecasting methods (grey system theory (GM (1, 1)), BP neural networks, Support Vector Machines (SVM)) to forecast; then, we compare and analyze the result of three forecasts methods and get a conclusion.

Simple descriptions of the process of various forecasting methods are listed below. For the sake of simplicity, the paper puts the forecasting results of three methods in the same table (Table 2). Only the forecasting results of the Shanghai Industrial Index are given in Table 2 owing to the limitation of space.

2.1. Stock index forecast based on SVM

Taking the Shanghai Industrial Index, the Shanghai Commercial Index, the Shanghai Real Estate Index and the Shanghai Public Utilities Index as study objects, the paper uses all data of 2007 as a training sample and a total of 141 data from January 4 to July 29 in 2008 as a testing sample together with 18 vectors in Table 1 as input vectors and the next day's closing price as output vector. $\varepsilon - SVR$ (ε -Support Vector Regression) model, the radial-Gaussian kernel, and LIBSVM are used during the forecasting process. Parts of the forecasting results will be provided in Table 2.

Table 1. Inputs of BP network and SVM

X1	Today's highest price
X2	Today's lowest price
X3	Yesterday's highest price
X4	Yesterday's lowest price
X5	The highest price of the day before yesterday
X6	The lowest price of the day before yesterday
X7	15-day moving average of highest price
X8	15-day moving average of highest price
X9	Today's opening price
X10	Today's closing price
X11	Yesterday's opening price
X12	Yesterday's closing price
X13	Today's turnover
X14	Yesterday's turnover
X15	The turnover of the day before yesterday
X16	Today's volume
X17	15-day moving average of volume
X18	15-day moving average of turnover

2.2. Stock index Forecast based on BP network.

We choose the same training samples, testing samples and the input and output vector as of SVM here. Input has total 18 neurons, and the output has only one neuron. Since it is difficult to determine the neuron numbers of the middle layer, which impact on the network performance to a large extent, we take 10,15,20,25,30,35,40 respectively as the neuron number of the middle layer; Take 2000 as the training times and 0.001 as the goal of training targets. When the network's predictive errors are the lowest, the neuron number of middle layer of the

neural network is the best number. We use MATLAB programming in forecasting process. Parts of forecasting results for the industrial index are listed in Table 2.

Table 2. The forecasting results of the various methods for industrial index

Data	Actual value	GM	BP neural	SVM
08-7-15	2220.8	2133.8	2282.1	2372.3
08-7-16	2201	2123.6	2214.5	2341.1
08-7-17	2264.5	2114.9	2245.3	2331
08-7-18	2327.8	2107.5	2242.9	2329.1
08-7-19	2314.2	2100.1	2255.3	2351.7
08-7-20	2300.4	2092.6	2247.1	2370.7
08-7-21	2344.5	2086	2268.2	2377.6
08-7-22	2307.1	2079	2305.4	2388.6
08-7-23	2347.9	2072.9	2253.2	2382.6
08-7-24	2318.8	2066.4	2445.1	2395.8
08-7-25	2306.4	2059.8	2455.5	2388.5

2.3. Stock index forecast based on GM (1, 1)

As for the grey theory, since the inheritance of the stock price and the stock index and the latest data contains the most abundant information, the original data for the grey theory to predict the stock index are set to be 10 and the next day's closing price is chosen as target vector. Meanwhile, in order to compare the forecasts with those of BP network, SVM, we select the data of the last 10 days in year 2007 and all the data in year 2008 as a sample, take a total of 141 samples from January 4 to July 29 as testing samples. We also use MATLAB programming in forecasting process, and parts of forecasting results for the industrial index are also listed in Table 2.

2.4. Forecasting results comparison of the three single methods.

According to the forecasting results of each single model in Table 2 and the above-mentioned definitions, we can get the forecasting error percentage comparison figures (Figure 1) respectively for the Shanghai Industrial Index, the Shanghai Commercial Index, the Shanghai Real Estate Index and the Shanghai Public Utilities Index. Through further calculation, we get the average forecast effectiveness of three forecasts methods and standard deviation of each model's effectiveness which are Mean Relative Error (MRE) and Mean Squared Error (MSE). It can be seen in Table 3.

It can be seen from Table 3:

(1) The average MRE of BP network method< the average MRE of GM (1, 1) method < the average MRE of SVM method, the average MSE of BP network method< the average MSE of GM (1, 1) method < the average MSE of SVM method. so forecast effects of BP Network method is the best, forecast effects of GM (1, 1) method is middle, and forecast effects of SVM method is the worst.

(2) The Standard Deviation of MRE of BP network method < the Standard Deviation of MRE of GM (1, 1) method < the Standard Deviation of MRE of SVM method, The Standard Deviation of MSE of BP network method < the Standard Deviation of MSE of GM (1, 1) method < the Standard Deviation of MSE of SVM method, so the forecast effects stability of BP Network method > the forecast effects stability of GM (1, 1) method > the forecast effects stability of SVM method.

Figure 2 is the curve of the industrial index, the Shanghai Commercial index, the Shanghai Real Estate Index, the Shanghai public utilities index on January 4, 2007 to July 29, 2008.

It can be seen from Figure 3: In the year 2007 and year 2008, each index fluctuates much. But relatively speaking, the Shanghai industrial index, the Shanghai Commercial index is more stable, while the Shanghai Real Estate Index and the Shanghai public utilities Index is relatively more concussing. It can be seen from Table 3 and Figure 3:

(1)On the whole, for the concussing stock index, the forecast effects of BP Network method are better than the forecast effects of GM (1, 1) method and SVM method.

(2)For forecasting relatively steady industrial index and Commercial index, the forecast effects of GM (1, 1) method is obviously worse than that of BP Network method and SVM method; the forecast effects of BP Network method and SVM method are very good, the MSE of SVM method in forecasting Shanghai Industrial Index is less than that of BP Network method.

(3)For forecasting relatively concussing Real Estate Index and public utilities Index, the forecast effects of BP Network method is obviously better than that of GM (1, 1) method and SVM method; Specially the forecast effects of SVM method in forecasting Real Estate Index and public utilities Index, is much worse than the forecast effects of SVM method in forecasting industrial index and Commercial index, even worse than the forecast effects of GM (1, 1) method.

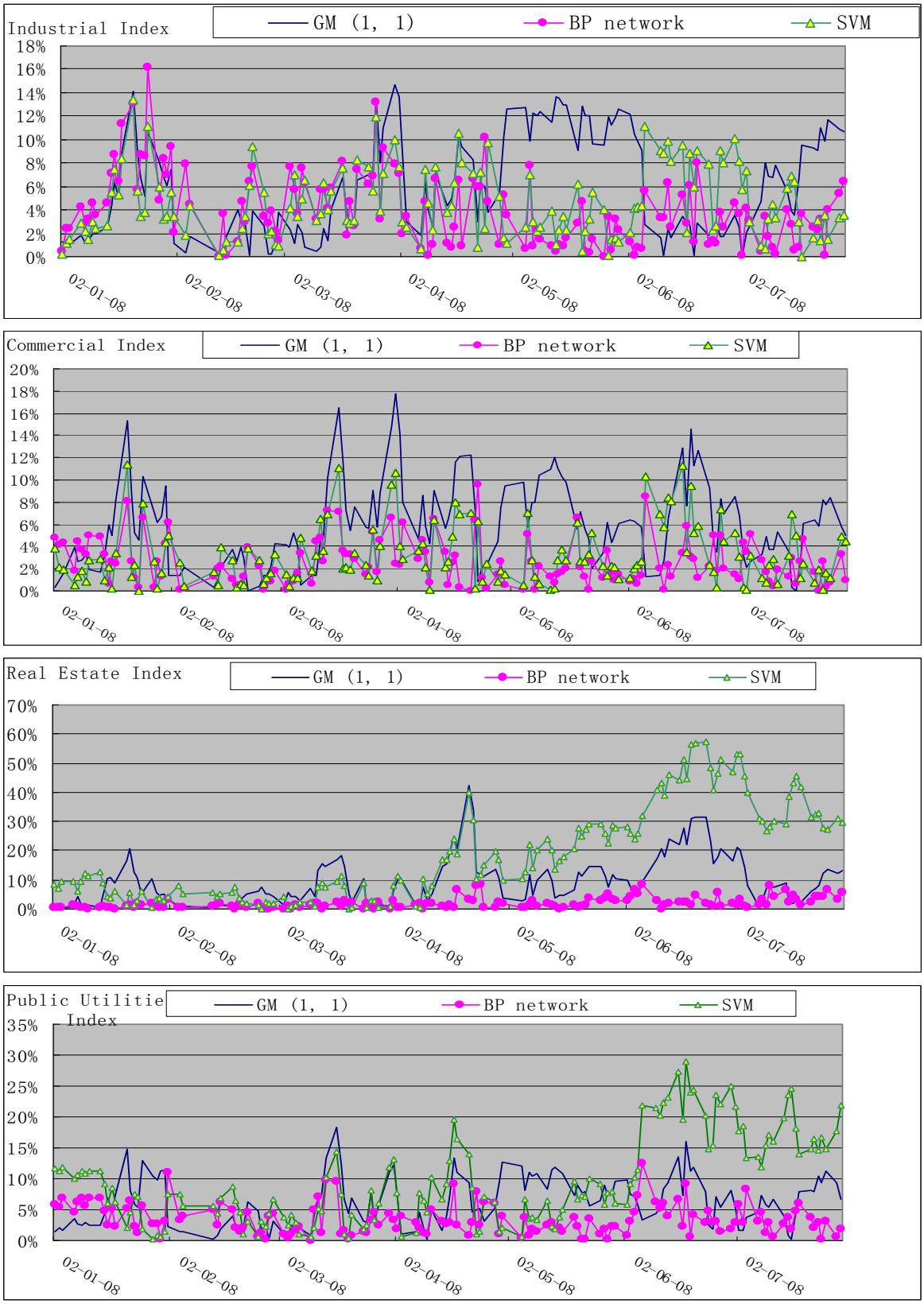

Figure 1. Forecast error percentage of three methods

Table 3. The forecast effectiveness of three methods

	MRE			MSE		
	GM(1, 1)	BP network	SVM	GM(1, 1)	BP network	SVM
Industrial Index	5.92%	3.98%	4.53%	212.20	167.41	160.06
Commercial Index	5.90%	2.57%	3.17%	231.50	115.01	134.39
Real Estate Index	9.63%	1.93%	18.56%	490.40	107.29	832.03
Public Utilities Index	6.25%	3.44%	9.65%	526.59	333.26	759.34
Average	6.93%	2.98%	8.98%	365.17	180.74	471.45
Standard Deviation	1.51%	0.24%	4.25%	616364.5	163799.4	1312498

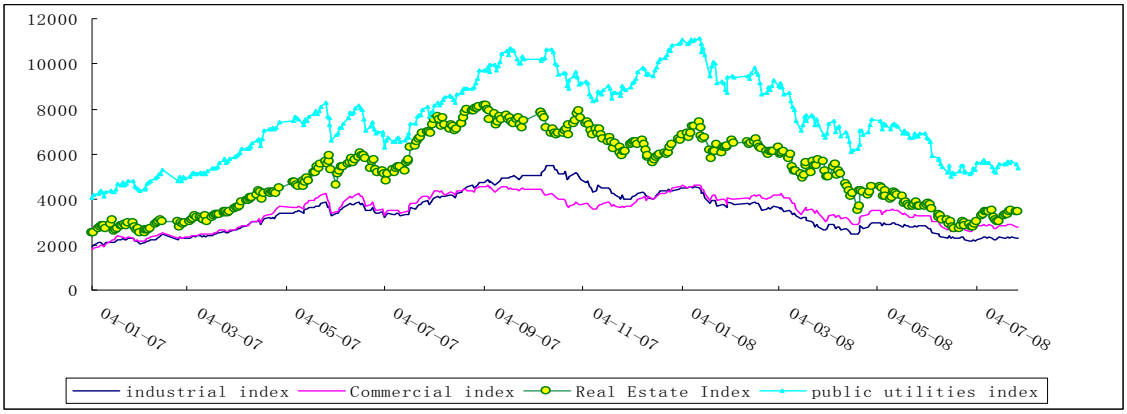

Figure 2 the curves of four indexes

We can see that, GM (1, 1) method and SVM method are more suitable for forecasting stable stock market. For forecasting concussing stock market, the forecast effects are not satisfactory.

3. Comparison and conclusions

All in all，we can get the following conclusions: For forecasting concussing stock market, the forecast effects of BP Network method is obviously better than the forecast effects of GM (1, 1) method and SVM method; GM (1, 1) method and SVM method are more suitable for forecasting stable stock market.

References

[1] Deng J. Basis Method of Grey System [M]. Shanghai, Central China Polytechnic University Press, 1998.

[2] Panfeng Li. Grey Forecasting of Stock Price [J]. EAST CHINA ECONOMIC MANAGEMENT, 1997, 4:60-61.

[3] Changfeng Cai. The Gray Prediction of Stock Market [J]. Mathematical Applicata, 2000, 13(4):78-81.

[4]Guoping Li, Senfa Chen, Guangqing Yu. Securities investment strategy based on forecast[A]. In Proceedings of the first behavioral finance and capital markets[C]. Nanjing University, 2003, 612-616.

[5] Guoping Li, Jungen Lei, Senfa Chen, Guangqing Yu, Boyu Chen, Xinping Li. The Problem of Optimizing The Data Quantity in Stock Market Gray Forecasting[A].In Proceedings of 2003 International Conference on WTO and Financial Engineering[C]. Hangzhou: Zhejiang University Press, 2004.

[6] Xiaotian Zhu, Hong Wang. Predicting stock index increments by neural networks: The role of trading volume under different horizons [J], Expert Systems with Applications, 2007(xxx) xxx－xxx.

[7] Jiancheng Long, Xiaoping Li. Study of the stock market tendency based on the neural network [J], Journal of Xidian University(Natural Science), 2005 32(3): 460-463.

[8] Huihui Meng, Deqian Ye, Na Liu. The Research of Stock Forecasting System which is Based on Neural Networks [J], Control & Automation, 2007, 23(1):240-242.

[9] Vapnik V. The nature of statistical learning theory[M]. New York: Springer- Verlag， 2000。

[10] Kim，K. J.."Financial time series forecasting using support vector machines", Neurocomputing, 2003, 55:307－319

[11] Kyoung-jae Kim. Financial time series forecasting using support vector machines. Neurocomputing. 2003，55: 307－319.

[12] Yanjie Shi. Prediction for Stock Market Based on Support Vector Machine [J]. Statistics and Decision. 2005, 2(下): 123-125.

[13] Lifang Peng, Zhiqing Meng, HuaJiang, Mi Tian. Application of Support Vector Machine Based on Time Sequence in Stock Forecasting [J]. Computing Technology and Automation. 2006, 25(3):88-91.