International Journal of Information Technology & Decision Making

A New Approach for Stock Price Analysis and Prediction Based on SSA and SVM

--Manuscript Draft--

Manuscript Number:	IJITDM-D-16-00043
Full Title:	A New Approach for Stock Price Analysis and Prediction Based on SSA and SVM
Article Type:	Special Issue Paper
Section/Category:	Special Issue on International Conference on Information Technology and Quantitative Management
Keywords:	Stock Price Series; Singular Spectrum Analysis; Support Vector Machine; Combination Predictive Methods
Corresponding Author:	meirui zhong CHINA
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	
Corresponding Author's Secondary Institution:	
First Author:	jihong xiao, PhD
First Author Secondary Information:	
Order of Authors:	jihong xiao, PhD
	xuehong zhu, phD
	chuangxia huang, PhD
	xiaoguang yang, PhD
	fenghua wen, PhD
	meirui zhong, PhD
Order of Authors Secondary Information:	
Abstract:	The stock price, affected by many factors, may have certain features over different time horizons, and analyses of stock price features can help us to better understand the intrinsic law of the stock price and improve the predictive accuracy. This paper, using the singular spectrum analysis (SSA), decomposes the stock price into terms of the trend, the market fluctuation, and the noise to study stock price features over different time horizons, and then introduce these features into the support vector machine (SVM) to make price predictions. The empirical evidence shows that compared with the SVM without these price features, the combination predictive methods——the EEMD-SVM and the SSA-SVM, which combining the price features into the SVM perform better, with the best prediction to the SSA-SVM.
Suggested Reviewers:	zhifeng dai, doctor Hong Kong Polytechnic University zhifeng.dai@polyu.edu.hk
	zhifang he, doctor University of Windsor zhifangh@uwindsor.ca

International Journal of Information Technology & Decision Making

 World Scientific Publishing Company

A New Approach for Stock Price Analysis and Prediction Based on SSA and SVM

The stock price, affected by many factors, may have certain features over different time horizons, and analyses of stock price features can help us to better understand the intrinsic law of the stock price and improve the predictive accuracy. This paper, using the singular spectrum analysis (SSA), decomposes the stock price into terms of the trend, the market fluctuation, and the noise to study stock price features over different time horizons, and then introduce these features into the support vector machine (SVM) to make price predictions. The empirical evidence shows that compared with the SVM without these price features, the combination predictive methods——the EEMD-SVM and the SSA-SVM, which combining the price features into the SVM perform better, with the best prediction to the SSA-SVM.

Keywords: Stock Price Series; Singular Spectrum Analysis; Support Vector Machine; Combination Predictive Methods

1. Introduction

Accurate stock price predictions are conducive to investors' avoiding risk effectively, the management's making reasonable operational strategies, and the government's understanding the economic trends better. The stock price prediction has therefore always been a hot spot in financial research. However, it is universally acknowledged to be the most challenging research topic in financial research, as it is affected by a number of factors, economic and non-economic.

Traditional prediction models[1-4] are generally based on the assumption that time series are linearly stationary. However, the stock price series have the features of complexity, nonlinearity and non-stationarity[5-6], so traditional prediction models have certain limitations in price prediction. Recently, neural networks[7-8] and the support vector machine[9-11] have been introduced into the financial field to predict the nonlinear and non-stationary time series, of which the support vector machine has been widely used in price predictions[12-21], as it can solve problems of being easily trapped in a local optimum and slow convergence existing in the learning and training process of the artificial neural network prediction. In stock price predictions, Trafalis and Ince (2000,2008)[12,13] made price predictions with the support vector machine and got better predictive results than those made with multi-layer perceptron (MLP) and autoregressive integrated moving average model (ARIMA); Kim, et al (2003)[14] made stock trends predictions with the SVM and proved that the SVM has better predictive results than the case based reasoning (CBR) and the back-propagation (BP) neural networks; Yeh, et al (2011)[15] developed a multiple-kernel support vector regression approach for stock market price forecasting. However, as the stock market is affected by economic, political, financial, social factors, noises and investors' behavior[22], stock prices may have different features (such as the fractal and chaotic features[23]) in different time horizons. But few studies have introduced the price features into the SVM to make price predictions.

Zhang, et al (2008)[24] used the empirical mode decomposition (EMD) to analyze fundamental features of petroleum price series over different time horizons and pointed out that the decomposed terms can be introduced into the SVM to make predictions. YU, et al (2008)[25] and Yang Yunfei, et al (2010)[26] used the EMD to decompose petroleum price series into time series with different economic implications and combined time series with neural networks and with the SVM to make combination predictions. The above literature demonstrates that making predictions about time series with its features can be easily extended to many fields. The EMD[27] and later-proposed an ensemble empirical mode decomposition (EEMD)[28] can be used in time series analysis, but they have some limitations in the analysis of stock price series. The residual terms used to study the long-term trend of time series have the features of monotonity, which do not necessarily conform to the long-term trend of stock prices. In addition, the EMD can not effectively extract noise from the price prediction, but the impact of noise is prevalent in the stock market. Therefore, the EMD can not capture this feature well.

The SSA, a method for analyzing non-linear, non-stationary time series, was first proposed by Broomhead and King (1986)[29]. Combined the classical analysis of time series, multi-variate geometry, multi-variate statistics, dynamical systems and many other elements, the SSA is to get a series of singular values which contains the information of the original series through the singular spectral decomposition (SVD). By analyzing singular values of different information, we can derive time series with different features. The series extracted for the analysis of the long-term trend of the time series are not limited by the monotonity, and they are often used for extracting noise in time series. Hassani (2007)[30] used the SSA method to extract information from the time series of the accidental death in U.S. and obtained the term of trend, harmonic and noise; Lian Jijian, et al (2008)[31] used order determination and noise reduction based on singular entropy to reduce noise in series; Beneki, et al (2012)[32] extracted the trend term and economic fluctuation term, and put them into the analyses of tourist revenues series in Britain; Zhang Yi, et al (2012)[33] used the SSA to extract noise term and trend term and put them into the analyses of exchange rate; Beneki et al (2013)[34] adopted SSA as the nonparametric time series analysis and forecasting technique to analysis and forecasting european union energy data. Huang et al(2014)[35] Forecasted the Volatility in Chinese Stock Market by HAR-CJ-M Model. In this paper, we also use the SSA to decompose the stock price, extract its features and make analyses.

As mentioned above, the status quo shows that the SVM method is still widely used to make stock price predictions. To make predictions more accurate, we use the SSA to decompose stock price series into the terms of trend, market fluctuation and noise, and study the stock price features over different time horizons, and then introduce these features into the SVM to make price predictions. Considering the great impact of the kernel function on SVM predictive accuracy, we complete the selection of the most suitable kernel function by index evaluation rather than by experience which is adopted by many scholars in SVM predictions. This paper is divided into four parts: introduction, introduction to predictive methods, empirical research and conclusion.

2. Introduction to Predictive Methods

2.1. *SSA*

The core idea of the SSA is to obtain a series of singular values which contain the information of the original series through the singular spectral decomposition (SVD), and then select different singular values to construct series with different constituents. Specific steps are as follows:

2.1.1. *Phase Space Reconstruction of Time Series*

Given a time series $X_N = \{x_1, x_2, ... x_N\}$, N is the effective length. Reconstruct the series' phase space, and then we can derive the trajectory matrix:

$$D_m = \begin{bmatrix} x_1, x_2 \dots, x_n \\ x_2, x_3 \dots, x_{n+1} \\ \dots, \dots, \dots, \dots \\ x_m, x_{m+1} \dots, x_{n+m-1} \end{bmatrix} \tag{1}$$

Where n is the length of the window, m is the embedding dimension, and $2 \le m \le N/2$, $m \le n$,. $N = n + m - 1$.

2.1.2. *Singular Value Decomposition*

Take matrix D_m for its singular value decomposition, $D_m = USV^T$, where U and V are *m*×*m and n*×*n* matrixes, s is *m*×*n* diagonal matrix, with the diagonal components $\lambda_1, \lambda_2, ... \lambda_p$, and $\lambda_1 \ge \lambda_2 \ge ... \ge \lambda_p$. $\lambda_1, \lambda_2, ... \lambda_p$ are the singular values of matrix D_m , and p is the number of the order of the singular value. U and V are the right singular matrix and the left singular matrix. If the time series only contain effective information, then the rank of the matrix s is $k < m$; if the time series contain both effective information and noise, then $k = m$. The following example is to extract the noise series.

2.1.3. *Determination of Order of Noise Reduction Based on Singular Entropy*

The concept of singular entropy[28] is introduced in order to study the law of the amount of information in time series changing with the number of the order of the singular spectrum:

$$E_K = \sum_i^k \Delta E\left(K < P\right) \tag{2}$$

Where k is the order of singular entropy, ΔE_i represents an increment of singular entropy in order i, by the following formula we can obtain that:

$$\Delta E = -\left(\lambda_i / \sum_{j=1}^{p} \lambda_j\right) \log\left(\lambda_i / \sum_{j=1}^{p} \lambda_j\right) \tag{3}$$

When the increment of the singular entropy falls to the asymptotic value, the effective information of the time series saturates, that is, the information is essentially complete, later increments of the singular entropy are caused by noise, and at this point the number of the order *i* is therefore selected as the one for noise reduction.

2.1.4. *Time Series Reconstruction After Noise Reduction*

After noise reduction, the time series reconstructed can be divided into two processes:

(1) We can obtain the number i for the order of noise reduction from the above analysis, set all singular values beyond order i zero to get a new singular values matrix s', and then use the formula $D_m' = us'v^T$ to derive a noise-free trajectory matrix.
(2) Let the reconstructed time series be $G = \{g_0, g_1, g_2, ...g_{N-1}\}$, and $x_{c,b-c+2}$ be a component of D_m', we have the following formula (essentially, the formula represents the mean of the sum total of the diagonal components of the matrix):

$$g_b = \begin{cases} \dfrac{1}{b+1} \displaystyle\sum_{c=1}^{b} x_{c,b-c+2} & 0 \le b \le m-1 \\ \dfrac{1}{m} \displaystyle\sum_{c=1}^{m} x_{c,b-c+2} & m \le b < n \\ \dfrac{1}{N-b} \displaystyle\sum_{c=N-n+2}^{N-n+1} x_{c,b-c+2} & n \le b < N \end{cases} \tag{4}$$

According to the above formula, we can obtain a time series or an effective information series G after noise reduction.

2.2. ***Support Vector Machine***

The main purpose of the SVM is to construct a hyperplane as the decision-making surface to maximize the separation margin between 2 classes, class A and class B.

Assume that hyperplane equation for separation is:

$$\omega x + b = 0 \tag{5}$$

Where ω is a normal vector, and b is a constant term.

If we can get the optimal ω_0 and b_0, then we can get the optimal hyperplane:

$$\omega_0 x + b_0 = 0 \tag{6}$$

The definition of support vectors: Data points (x_i, y_i) that can meet the requirements are called the support vectors:

$$\omega x_i + b = -1 \;\;,\;\; y_i = -1 \;\text{ or }\; \omega x_i + b = 1,\;\; y_i = +1 \tag{7}$$

The support vector is the data points nearest to the decision-making plane.

Classify the data into two classes: class A and class B. class A is x_1, class B is x_2. Where $\omega x_1 + b = 1, \omega x_2 + b = -1$, the margin between class A and B is $dis = \omega * (x_1 - x_2) / \|\omega\| = 2 / \|\omega\|$, and the maximization problem of $2 / \|\omega\|$ can be transformed into the minimization problem of $\|\omega\|^2 / 2$.

Extend to (x_i, y_i), then the problem of maximizing the separation margin between class A and class B is transformed into a quadratic optimization programming problem:

$$\min \frac{\|\omega\|^2}{2} \tag{8}$$

$$s.t: y_i (\omega x_i + b) \geq 1, i = 1, 2 \ldots n$$

Let the training set of samples be $T = \{(x_i, y_i)\}_{i=1}^{n}$ (where x_i is an input variable, y_i is an output value, and i = 1,2 ... n, n is the number of samples), if the training samples are not separable linearly, the input variable x should be mapped onto a high dimensional feature space through a non-linear mapping ϕ; then perform linear regression in a high dimensional space. Construct the optimal learning machine:

$$f(x) = \omega^T \phi(x) + \tag{9}$$

In 1998, Vapnik[10] proposed the ε - support vector machine to solve problems in predictions. Here we mainly introduce the ε - SVM. As we can see, the linear quadratic programming problem of equation (9) can be transformed into the following expression:

$$\min_{\omega, b\ \xi\ \xi^*} \frac{\omega^T \omega}{2} + c \sum_{i=1}^{n} \xi_i + c \sum_{i=1}^{n} \xi^*$$

$$s.t: \omega^T \phi(x_i) + b - y_i \leq \varepsilon + \xi_i \tag{10}$$

$$\omega^T \phi(x_i) + b - y_i \geq \varepsilon + \xi_i^*$$

$$\xi_i, \xi_i^* \geq 0, i = 1, 2 \ldots n$$

To solve the above problem, we need to introduce a Lagrange function, and according to the Duality and the saddle point condition, we can get the dual form:

$$\min_{\partial, \partial^*} \frac{1}{2} (\partial - \partial^*)^T Q (\partial - \partial^*) + \varepsilon \sum_{i=1}^{n} (\partial_i + \partial_i^*) + \sum_{i=1}^{n} y_i (\partial_i - \partial_i^*) \tag{11}$$

$$s.t:\sum_{i=1}^{n}\left(\partial_i-\partial_i^*\right)=0, 0\le\partial_i,\partial_i^*\le c, i=1,2...n$$

$$\omega=\sum_{i=1}^{n}\left(\partial_i^*-\partial_i\right) \tag{12}$$

$Q_{ij}=k\left(x_i,x_j\right)\equiv\phi\left(x_i\right)^T\phi\left(x_j\right)$, $k\left(x_i,x_j\right)$ are kernel functions. As long as a function meets the Mercer requirement, it can be used as a kernel function. The main kernel functions are: (1) Linear kernel $k\left(x,x_i\right)=x^T x_i$; (2) Polynomial kernel $k\left(x,x_i\right)=\left(\gamma x^T x_i+r\right)^p$; (3) RBF kernel $k\left(x,x_i\right)=\exp\left(-\gamma\left\|x-x_i\right\|^2\right)$; (4) Sigmoid kernel $k\left(x,x_i\right)=\tanh\left(\gamma x^T x_i+r\right)$。From the above analysis, we can obtain the optimal classifier:

$$f\left(x\right)=\sum_{i=1}^{n}\left(\partial_i^*-\partial_i\right)k\left(x_i,x\right)+b \tag{13}$$

3. Empirical Research

3.1. *Research Steps*

We study the effect of various factors on the stock price over three different time horizons of the long term, the mid-term, and the short term with reference to the study of Zhang, et al [24] on the analysis of the petroleum price. Over the long-term horizon, we focus on the long-term trend of the stock price; over the mid-term horizon, we pay our attention to the impacts of various events with economic and non-economic implications on the stock price; and over the short-term horizon, we turn our eyes to the noise in the market. The SSA method has its advantages in decomposing the stock price: (1) SSA can be employed to analyze non-linear and non-stationary time series; (2) the extracted series for the long-term trend of the stock price with the SSA is not limited by monotonity; and (3) the SSA can extract the noise series, which is conducive to the analysis of the noise—the factor universally existing in the stock market. Therefore this paper utilizes the SSA to derive the singular values which contain different information of the original series. The first singular value contains the most fundamental information of the original series, and can be used to construct the trend term to study the long-term trend of the stock price. The singular entropy noise reduction method is used to derive the singular value that contains the noise of the original series, which is exploited to construct the noise term to study the effect of the noise in the stock market. And then the remaining singular value is used to construct the market fluctuation term to study the effects of various events. With analyses of the trend term, the market fluctuation term and the noise term, the intrinsic law for the original series can be understood. And finally predictions can be made with the help of an SVM. The specific steps are as follows：

(1) SSA decomposition: first select a suitable number for the embedding dimension to reconstruct the phase space for the original series to derive the trajectory matrix, and then use the SVM to decompose the trajectory matrix to obtain singular values; determine the number of the order according to the property of the increment for the singular entropy; and finally extract the singular values to reconstruct the trend term, the market fluctuation term, and the noise term.

(2) Selection of suitable kernel functions: select with the help of evaluation indexes a kernel function for each decomposed trend term, market fluctuation term, and noise term, and perform SVM predictions for each term using a suitable kernel function to get predictive values for the terms.

(3) SVM combination prediction: use the predictive value of each term as the input, and the real price at that time as the output to construct a training model; use the training model to make a prediction; when the input variable is the predictive value for each component, the output is the final predictive value.

The above steps are presented in the following flow chart:

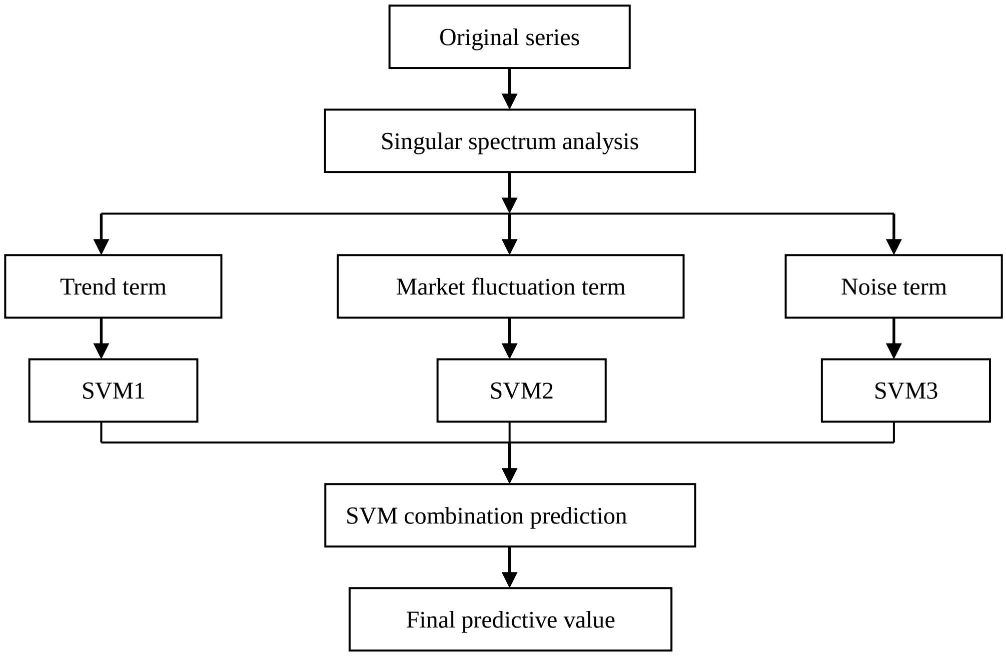

Figure.1. Empirical Research Process Flow Chart

3.2. *Experiment Data and Evaluation Index*

We select the closing price of the Shanghai Stock Exchange (SSE) Composite Index from Jan. 5, 2009 to Sept. 30, 2013 as the empirical research subject, which offers 1150 data points (the data supplied by the Guotai Jun'an CSMAR research databank). We use two-thirds of the research data points (the first 767 closing price) as the training data, and one-third of the research data points (the final 383 closing price) as the test data. In the empirical research, the predictive value is derived one step ahead by moving the window. The evaluation indexes are the mean squared error (MSE), the mean absolute prediction error (MAPE), the directional symmetry (DS). The less the values for the MSE and the MAPE are, the better the predictions are; for the DS, the greater the value is, the more accurate the prediction is. The formulae of these evaluation indexes are listed as follows:

$$\text{MSE}=\frac{1}{n}\sum_{t=1}^{n}\left[x(t)-\bar{x}(t)\right]^{2} \tag{14}$$

$$MAPE=\frac{1}{n}\sum_{t=1}^{n}\left|\frac{x(t)-\bar{x}(t)}{x(t)}\right| \tag{15}$$

$$DS=\frac{1}{n}\sum_{i=1}^{n}d_i$$

$$d_i = \begin{cases} 1, \left[x(t) - x(t-1)\right] * \left[\bar{x}(t) - x(t-1)\right] \geq 0 \\ 0, \left[x(t) - x(t-1)\right] * \left[\bar{x}(t) - x(t-1)\right] < 0 \end{cases} \tag{16}$$

Where $x(t)$ is the actual closing price for the stock at time t, $\bar{x}(t)$ is the predictive value at time t, and n is the number of the actual values or the predictive values.

3.3. *Empirical Study Process*

Let the original series be $x(t)$, and then take $x(t)$ for phase space reconstruction. The selection of the embedding dimension is very important in phase space reconstruction, as it should be large enough but without exceeding half the length of the original series so that we can get all the information from the original series. Specifically, if we know the periodic dimension of the original series, the number of the embedding dimension should be an integer multiple of the periodic dimension; if there exist different periodic dimensions, and then the embedding dimension should cover all periodic dimensions. Practically, one-third or one-fourth of the length of the original series is selected to obtain all the information from the original series. This paper selects one-third the length of the original series (approximately 400) as the embedding dimension, decomposes the derived trajectory matrix into the singular value matrix through the SVD, and then constructs the graph of the increment of the singular value, as shown in Figure 2.

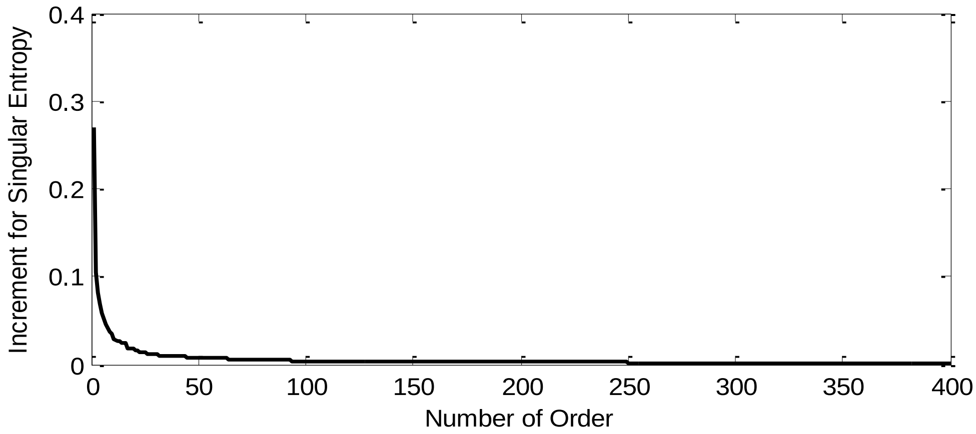

Figure.2. the Increment of the Singular Entropy

As seen in Figure 2, when the number of the order reaches 50, the increment changes little and the effective information saturates. Reconstruct the singular values beyond the number of the order 50 as the noise term $x_3(t)$, and thus the first singular value contains the most fundamental information of the series, use it to reconstruct the trend term $x_1(t)$ and use singular value 2-50 to reconstruct the market fluctuation term $x_2(t)$. Then make a graphical comparison of $x(t)$, $x_1(t)$, $x_2(t)$ and $x_3(t)$, see Figure 3.

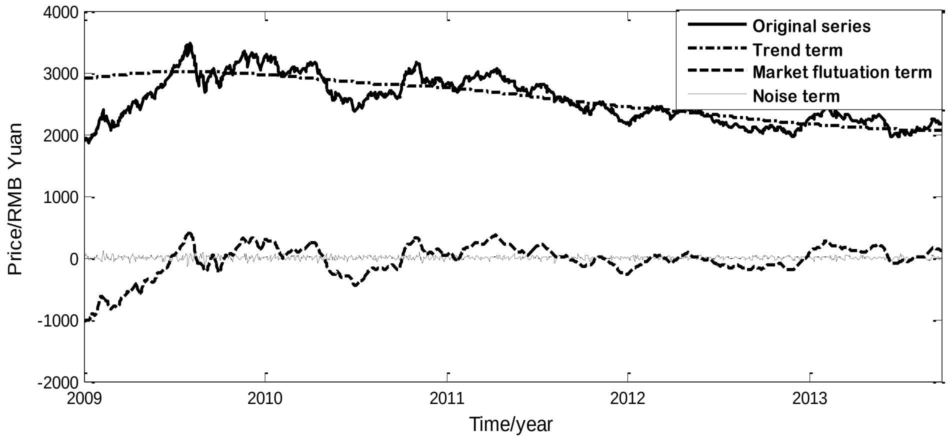

Figure.3. Original Series and Decomposed Terms

3.3.1. *An Analysis of Economic Implications of Decomposed Terms*

If the correlation between the trend term and the original series is up to 74%, it indicates that the trend term plays a decisive role in the long-term trend of stock prices. Changes in the trend term are inseparable from the economic condition of a country; in turn, the economic development determines the changes in the trend term to some extent, while changes in the trend term can also reflect some problems that exist in the economic development. Since the financial crisis in 2008, the global stock market suffered a setback, so did the domestic stock market and the Shanghai Stock Exchange (SSE) Composite Index. After the crisis, the government introduced massive stimulus plans. Then the economy of the country began to upturn and the stock index began to go up. The trend term in the graph shows that the Shanghai Stock Exchange (SSE) Composite Index began to rise from 2009, but declined since 2010, which seems to be inconsistent with the 7 percent economic growth the country achieved in recent years. However, there exist various problems in China's economic growth—too hot economy, excessive dependence on investment, and thus too many bubbles in the economy with a serious lack of confidence in the economy on the part of the general public, and the necessity of returning to the real growth and pricking the bubbles emerges, as the stock index rise up to a certain level.

The market fluctuation term is used primarily to reflect the impacts of mid-term economic or non-economic events on the stock price, as events will bring considerable shocks to the stock price. The curve of the market fluctuation term shows that the fluctuation is volatile from 2009 to 2010, but it is relatively mild from 2011 to 2013, indicating that the effect of events on price fluctuations lasts from a couple of weeks to a few months, while the effect of the events with considerable implications lasts usually longer. In addition, some data points fluctuate violently, indicating that the effect of events on stock prices is very significant. The violent fluctuations over the mid-term horizon are mostly brought about by the shocks of various events, as the trend term changes slowly, and the noise term fluctuates mildly. That's why the curve of the original stock index is so similar to the shape of the graph of the market fluctuation term. Extract the events from the original index series is conducive to carry out stock index predictions. Thus, when there are similar events taking place, such as frequent economic events or specific non-economic events, we can predict their gains and durations with reference to past studies.

The noise term mainly reflects the effect of the universal noise in the stock market on the stock price. China's stock market has a very short history with many market rules and regulations yet to be improved. Many market participants lack financial expertise, and are very liable to herding behavior. What is worse, China's stock market lacks transparency, more often than not with incomplete information to investors. It can be seen from the Graph for the noise term, the effect of the noise presents in the short term with mild but frequent fluctuations of some dozen points or even less, the significance of which, compared with those of the trend term and of the market fluctuation term is not so meaningful. But the fluctuations of the noise term occur at a high frequency with local impacts on the stock index, and it is therefore very important in short-term predictions, albeit not much significance for the long-term trend of the stock price.

3.3.2. *SVM Prediction*

Since the kernel function has a direct impact on the SVM prediction, we'll select a kernel function every time when we make a SVM prediction. Firstly we make a single prediction based on decomposed terms.

Select the kernel function from the trend term $x_1(t)$, predictive results on different kernel functions are shown in Table 1.

Table 1 Prediction Errors of Different Kernel Functions Based on Trend Terms

	Linear Kernel Function	Polynomial Kernel Function	RBF Kernel Function	Sigmoid Kernel Function
MSE	0.000041	0.005953	0.000046	0.167835
MAPE	0.000918	0.009672	0.000970	0.057335
DS	45.6693%	43.0446%	59.0551%	24.6719%

From Table 1, we can get a linear kernel function more suitable for the SVM trend term predictions, and then get the predictive value $x_1'(t)$ of trend terms by linear kernel function. The comparison of predictive and real values is shown in Figure 4.

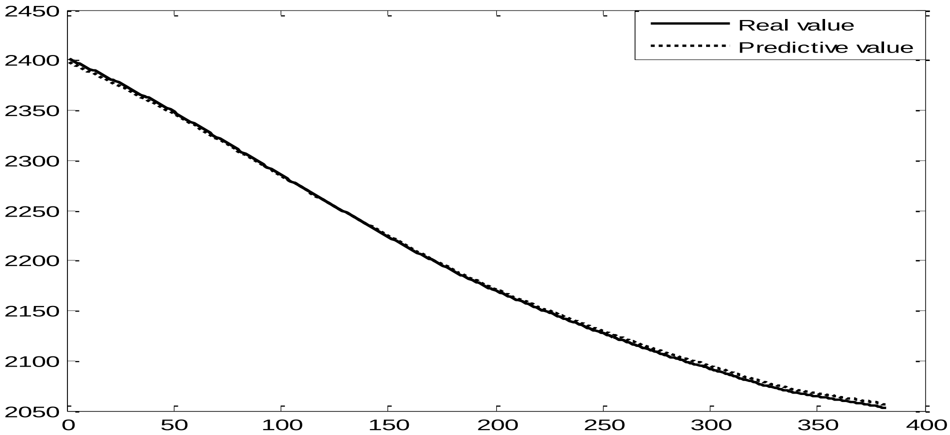

Figure.4. A Comparison of Real and Predictive Values Based on Trend Terms

Figure 4 shows a very satisfactory predictive result, as the real values and the predictive values essentially form the same curve. The trend term reflects a long-term trend, which changes slowly with considerable regularity and is thus fairly easy to be forecasted.

Select kernel functions from the market fluctuation term $x_2(t)$, predictive results on different kernel functions are shown in Table 2.

Table 2 Prediction Errors of Different Kernel Functions Based on Market Fluctuation Terms

	Linear Kernel Function	Polynomial Kernel Function	RBF Kernel Function	Sigmoid Kernel Function
MSE	0.000316	0.007628	0.000322	0.155809
MAPE	0.183587	0.394703	0.186350	3.954080
DS	47.5066%	48.0315%	47.5066%	50.9186%

From Table 2, we can get a linear kernel function suitable for the SVM predictions based on market fluctuation terms, and then derive the predictive value $x_2'(t)$ of market fluctuation terms from the linear kernel function. The comparison of predictive and real values is shown in Figure 5.

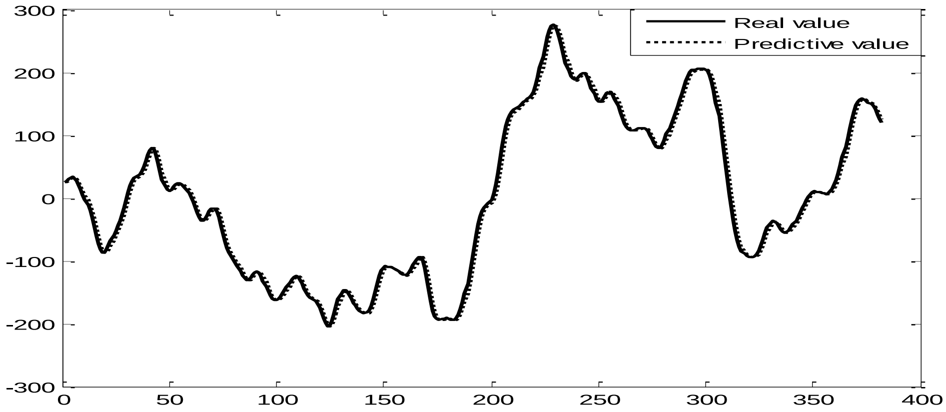

Figure.5. Real and Predictive Values Based on Market Fluctuation Terms

Figure 5 shows that the real values and the predictive values essentially form the same curve, i.e., the predictive values characterize the actual fluctuation pretty well. The market fluctuation term primarily reflects the impacts of events on the stock price. Although the stock price fluctuates frequently with a considerable magnitude, the effects of events last a certain period of time from a couple of weeks to a few months, which indicates fairly obvious regularity and is thus capable of predicting with pretty good accuracy.

Select kernel functions from the noise term $x_3(t)$, predictive results on different kernel functions are shown in Table 3.

Table 3 Prediction Errors of Different Kernel Functions Based on Noise Terms

	Linear Kernel Function	Polynomial Kernel Function	RBF Kernel Function	Sigmoid Kernel Function
MSE	0.024478	0.024012	0.024367	0.031451
MAPE	2.581213	1.938445	2.529581	7.718691
DS	66.1417%	66.6667%	66.4042%	65.8793%

From Table 3, we can get a Polynomial kernel function suitable for the SVM predictions based on market fluctuation terms, and then derive the predictive value $x_3'(t)$ of market fluctuation terms from the Polynomial kernel function. The comparison of predictive and real values is shown in Figure 6.

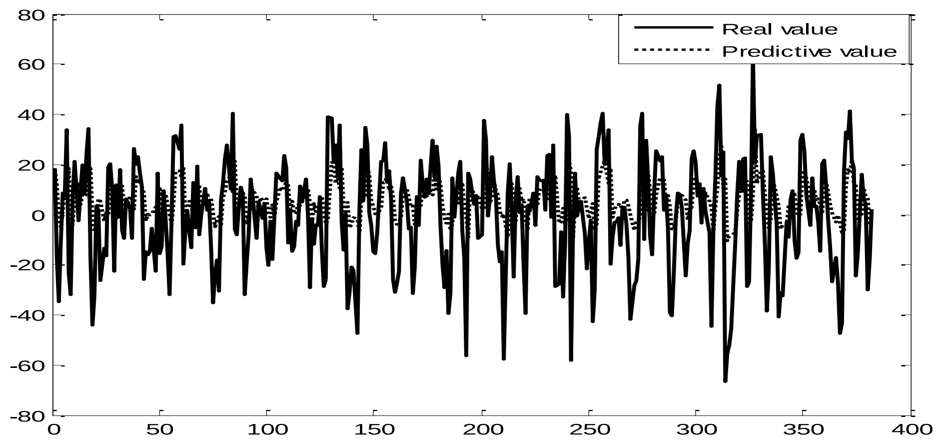

Figure.6. Real and Predictive Values Based on Noise Terms

It can be seen from Figure 6 that there are serious deviations of the predictive values from the real values, the predictive results are thus unsatisfactory. This is because noise traders often trade for short-term profits, and their investment decisions change pretty fast, thus making it difficult to present any regularity for predictions with some accuracy.

Secondly we make a combination prediction based on the predictive values $x_1'(t)$, $x_2'(t)$ and $x_3'(t)$. predictive results on different kernel functions are shown in Table 4.

Table 4 Predictive Errors Based on combination prediction

	Linear Kernel Function	Polynomial Kernel Function	RBF Kernel Function	Sigmoid Kernel Function
MSE	0.001852	0.002324	0.001785	2.40823
MAPE	0.007454	0.008417	0.007245	0.27686
DS	64.3035%	61.9423%	67.979%	56.168%

Table 4 show that the predictive results of RBF kernel functions are more satisfactory than those of other kernel functions, So we make SVM combination predictions based on RBF kernel functions and get a final prediction value $x'(t)$. The final comparison of the real and predictive values is shown in Figure 7.

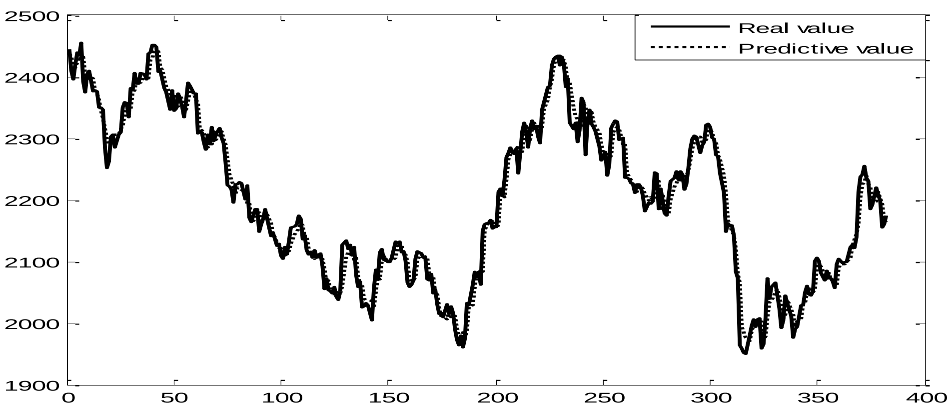

Figure.7. Final Comparison of the Real and Predictive Values

Figure 7 shows that the final predictive value provides a good fitting to the trend, and the change of the wave shapes of the actual price, and thus making a good predictive result.

With the same sample data, the prediction made only with the SVM, the EEMD-SVM combination prediction (the EMD decomposition has the effect of modal overlapping, but the EEMD decomposition gets rid of the above effect, and is more suitable for the decomposition of the original series) and the prediction made with our method (SSA-SVM) are compared with their results tabulated in Table 5.

Table 5 Prediction Error

	SVM Prediction	EEMD-SVM	SSA-SVM
MSE	0.002315	0.001972	0.001785
MAPE	0.008138	0.007696	0.007245
DS	52.493%	62.729%	67.979%

From Table 7, we can see that making combination predictions by decomposing the original index series into series with economic implications is more desirable than making mere SVM predictions, and SSA-SVM combination predictions are better than EEMD-SVM combination predictions, which indicates that SSA decomposition of the original index series can better grasp the features of the original index series, and thus get better predictive results.

4. Conclusion

Based on the SSA and the research subject selected from the closing price of the Shanghai Stock Exchange (SSE) Composite Index from Jan. 5, 2009 to Sept. 30, 2013, this paper decomposes the original stock index into the trend terms, the market fluctuation terms, the noise terms and time series with different economic implications by using the SVD and studies the stock index features over different time horizons. Research result indicates that the trend term plays a decisive role in the long-term trend of stock index, as changes in the trend term are closely related to the economic situation of China, and it can also reflect the country's economic situation to some extent; the obvious fluctuations are mostly brought about by the shocks of various events, and the effects of events on the stock price usually last a certain period of time, from a couple of weeks to a few months; moreover, the analysis of past events can help to improve the predictive accuracy; localized changes in the stock market index are resulted from the impact of noise traders, though they can not affect the long-term trend of the stock, they play a huge role in the short-term prediction.

Then we introduce the index features into the SVM method for prediction. Considering the importance of the kernel function in SVM prediction, every time we make a SVM prediction, we select the most suitable kernel function by making empirical studies on the four kernel functions and evaluating the indexes rather than by experience. When making predictions on the decomposed terms, we find that the predictive results of the trend term rank the highest in accuracy, of the market fluctuation term ranks second because of the uncertainty of future events, of the noise term ranks third, because it is difficult to obtain a satisfactory predictive result due to the frequent fluctuations and randomness of the noise term. Finally, we compare the predictive effect of the SSA-SVM combination prediction with the SVM prediction and the EEMD-SVM combination prediction, and find that the combination prediction which combines the decomposition of original index into series with certain economic implications to the SVM is more effective than the SVM prediction; and the SSA can better grasp the features of the original index series than the EEMD, while the SSA-SVM combination prediction have better predictive effect than that of the EEMD-SVM combination prediction.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (No.71171024, No.71371195, No.71431008, No.71471020), and the major program of the National Social Science Foundation（13&ZD024）.

Reference

1. Engle R E. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kindom inflation. *J. Econometrica*, 50(4) (1982) 987-1007.
2. Bollerslev T. Generalised autoregressive conditional heteroskedasticity. *J. Journal of Econometrics,* 31(3) (1986) 307-327.
3. Taylor S J. Modeling stochastic volatility. *J. Mathematical Finance*, 04 (1994) 183-204.
4. Huang C, Peng C, Chen X, Wen F. Dynamics Analysis of a Class of Delayed Economic Model. *J/OL*.Abstract and Applied Analysis, 2013 (2013) 1-12.
5. Dai Z, Wen F. Robust CVaR-based portfolio optimization under a genal affine data perturbation uncertainty set, *J. Journal of Computational Analysis & Applications,* 16(1) (2014) 93-103.
6. Liu J, Tao M, Ma C, Wen F. Utility indifference pricing of convertible bonds, *J. International Journal of Information Technology & Decision Making,* 13(2) (2014) 439-444.
7. Hunt JE, Cooke DE: Learning Using an Aritificial Immune System. *J. Journal of Network and Computer Applications,* 1996(19): 189-212.
8. Huang C, Kuang H, Chen X, Wen F. An LMI Approach for Dynamics of Switched Cellular Neural Networks with Mixed Delays. *J/OL. Abstract and Applied Analysis*, 2013 (2013) 1-8.
9. Cortes C, Vapnik V. Support-Vector network, *J. Machine Learning*. 20 (1995) 273-297.
10. Vapnik V. *Statistical Learning Theory* (Wiley, New York, NY, 1998).
11. Zhang Xuegong. About the statistical learning theory and support vector machine, *J. ACTA AUTOMATICA SINICA,* 01 (2000) 36-46.
12. Trafalis T, Ince H, *Support vector machine for regression and applications to financial forecasting*. ed. Como (Proceedings of the IEEE-INNS-ENNS International joint Conference on Neural Networks, 2000) July 24-27.

13. Ince H, Trafalis T, Short term forecasting with support vector machines and application to stock price prediction, *J. International Journal of General Systems.* 37(6) (2008) 677-687.
14. Kim K. Financial time series forecasting using support vector machines, *J. Neurocomputing.* 55(1-2) (2003) 307-320.
15. Yeh Chi-Yuan, Huang Chi-Wei, Shie-Jue Lee. A multiple-kernel support vector regression approach for stock market price forecasting, *J.Expert systems with Applications.* 38 (2011) 2177-2186.
16. Tay F E H, Cao Lijuan. Improved financial time series forecasting by combining support vector machines with self-organizing feature map, *J. Intelligent Data Analysis.* 5(4) (2001) 339-354.
17. Cao Lijuan, Tay F E H. Financial forecasting using support vector machines, *J. Neural Computing & Applications.* 10(2) (2001) 184-192.
18. Tay F E H, Cao Lijuan. Application of support vector machines in financial time series forecasting, *J. Omega.* 29 (2001) 309-317.
19. Tay F E H, Cao Lijuan. e-Descending support vector machines for financial time series forecasting, *J. Neural Processing Letters.* 15(2) (2002) 179-195.
20. Tay F E H, Cao Lijuan. Modified support vector machines in financial time series forecasting, *J. Neurocomputing.* 48 (2003) 847-862.
21. Sai Ying, Zhang Fengting, Zhangtao. China's stock index futures regression prediction research based on SVM, *J. China Journal of Management Science*, 03 (2013) 35-39.
22. WEN F, HE Z, DAI Z, YANG X. Characteristics of investor's risk preference for stock markets, *J. Economic Computation and Economic Cybernetics Studies and Research,* 48(3) (2014) 235-254.
23. Wen F, Li Z, Xie C, Shaw D. Study on the fractal and chaotic features of the Shanghai composite index, *J. Fractals-complex geometry patterns and scaling in nature and society*, 20(2) 2012 133-140.
24. Zhang X, Lai k k, Wang S Y.A New Approach for Crude Oil Price Analysis Based on Empirical Mode Decomposition, *J. Energy Economics,* 30(3) (2008) 905-918.
25. Yu L, Wang S Y, Lai K K. Forecasting Crude Oil Price with an Emd-Based Network Ensemble Learning Paradigm, *J. Energy Economics,* 30(5) (2008) 2623-2635.
26. Yang Yunfei, Bao Yukun, Hu Zhongyi, Zhang Rui. Rrude oil price forecasting method based on EMD and SVMS, *J. Journal of Management*, 12 (2010) 1884-1889.
27. Huang N E, Shen Z, Long S R. The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis. *J. Process of the Royal Society of London A*, 454 (1971) (1998) 903-995.
28. Wu Zhao-hua, Huang N E. Ensemble empirical mode decomposition: a noise assisted data analysis method, *J. Advances in Adaptive Data Analysis,* 1(1) (2009) 1-41.
29. Broomhesd, D. S. King, G. P. Extracting Qualitative Dynamics from Experimental Data, *J. Physica D:Nonlinear Phenomena,* 20(2-3) (1986) 217-236.
30. Hossein Hassani, Singular Spectrum Analysis: Methodology and Comparison, *J. Data Sci,* 5(1) (2007) 239-257.
31. Lian Jijian, Li Huokun, Zhang Jianwei. Hydranlic structure vibration modal ERA identification method based on the singular entropy order noise reduction, *J. SCI CHINA SER,* 09 (2008) 1398-1413.
32. Beneki,C.,Eeckels,B.,and Leon,L.,Signal Extraction and Forecasting of the UK Tourism Income Time Series: A Singular Spectrum Analysis Approach, *J. Journal of Forecasting,* 31(5) (2012) 351-400.
33. Zhang Yi, Hui Xiaofeng. The exchange rate prediction research based on singular spectrum analysis, *J. Statistics and Decision,* 06 (2012) 29-31.
34. Beneki C, Silva E S. Analysing and forecasting European Union energy data, *J. International journal of Energy and Statistics*, 02 (2013) 127-141.
35. Huang C, Gong X, Chen X, Wen F. Measuring and Forecasting Volatility in Chinese Stock Market Using HAR-CJ-M Model, *J/OL. Abstract and AppliedAnalysis,* 2013 (2013) 1-13.