

Research Article

Forecasting Crude Oil Price and Stock Price by Jump Stochastic Time Effective Neural Network Model

Jun Wang, Huopo Pan, and Fajiang Liu

Department of Mathematics, Key Laboratory of Communication and Information System, Beijing Jiaotong University, Beijing 100044, China

Correspondence should be addressed to Jun Wang, wangjun@center.njtu.edu.cn

Received 2 July 2011; Revised 19 September 2011; Accepted 10 October 2011

Academic Editor: Wolfgang Schmidt

Copyright © 2012 Jun Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The interacting impact between the crude oil prices and the stock market indices in China is investigated in the present paper, and the corresponding statistical behaviors are also analyzed. The database is based on the crude oil prices of Daqing and Shengli in the 7-year period from January 2003 to December 2009 and also on the indices of SHCI, SZCI, SZPI, and SINOPEC with the same time period. A jump stochastic time effective neural network model is introduced and applied to forecast the fluctuations of the time series for the crude oil prices and the stock indices, and we study the corresponding statistical properties by comparison. The experiment analysis shows that when the price fluctuation is small, the predictive values are close to the actual values, and when the price fluctuation is large, the predictive values deviate from the actual values to some degree. Moreover, the correlation properties are studied by the detrended fluctuation analysis, and the results illustrate that there are positive correlations both in the absolute returns of actual data and predictive data.

1. Introduction

The objective of this work is to investigate the relationships between the crude oil market and the stock market and examine whether the shocks in crude oil price transmitted to Chinese stock market will receive considerable attention from investors. In the past decade, the crude oil demand of China is growing rapidly, and China has already become the second-largest oil importer in the world, after the United States. Fourteen years ago, China from an oil-exporting country became a net oil-importing country. From then on, the movement of crude oil prices had a strong influence on the economic behavior of individuals and firms, and as a result, it affects the economic development directly. In another aspect, since July 2009, China

has taken the place of Japan to be the world's second-largest stock market, and the stock market has played an important part in its economy. China has two stock markets: Shanghai Stock Exchange and Shenzhen Stock Exchange. The indices studied in the present paper are Shanghai Composite Index (SHCI) and Shenzhen Compositional Index (SZCI). These two most influential indices play an important role in Chinese stock markets. We also consider Shenzhen Petrochemical Index (SZPI) and the stock price of China's largest oil company: China Petroleum & Chemical Corporation (SINOPEC). Daqing oil field and Shengli oil field are the first and the second largest oil fields in China respectively, the crude oil prices of Daqing and Shengli have a strong impact on Chinese energy market. The data for these crude oil prices and indices in the 7-year period is selected and analyzed by the statistical method and the neural network method.

Recently, some progress has been made in the study of fluctuations for the financial market and the energy market in China, for example see [1–7]. Artificial neural networks (ANNs) are one of the technologies that have made great progress in studying the stock markets [3, 8–11]. ANN have good self-learning ability, a strong antijamming capability, and they have been widely used in financial fields such as stock prices, profits, exchange rate, and risk analysis and prediction. Although the historical data has a great influence on the investors' positions, we think that the impacts of different historical data on the stock price are not same. In the present paper, we suppose that the degree of impact of a data depends on its occurring date (or time), we give a high level effect of a data when it is very near to the current state. Furthermore, we also introduce the Brownian motion and Poisson jump in the model [3, 6, 11–15], in order to make the model have the effect of random movement and random jump while maintaining the original trend. In a financial market, jumps in financial assets play a crucial role in volatility forecasting. And jumps have a positive and mostly significant impact on future volatility. In this work, the artificial neural network model based on jump stochastic time effective function is applied to forecast the fluctuations of SHCI, SZCI, SZPI, Daqing, Shengli, and SINOPEC. We study the statistical behaviors and the linear regression for these indices, and the simulation plots and the comparisons of the observed data are given. We introduce mean absolute error (MAE), mean relative error (MRE), Theil's inequality coefficient (Theil's IC), bias proportion (BP), variance proportion (VP) and covariance proportion (CP) to evaluate the predictive results. Detrended fluctuation analysis (DFA) is developed to study both the stock markets and the crude oil markets [16–19]. DFA is one of the statistical analysis methods, which is applied to study the extent of long-range correlations in time series, it gives a statistical approach that reduces the effects of nonstationary market trends and focuses on the intrinsic autocorrelation structure of market fluctuations over different time horizons. DFA provides a simple quantitative parameter, the scaling exponent α , to represent the correlation properties of time series. In the last part of Section 3, the empirical analysis shows the positive correlations in the absolute returns of the actual data and the predictive data by calculating the scaling exponent α .

In this paper, we introduce a new method: the jump stochastic time effective function in the neural network, to investigate the relationships between the crude oil market and the stock market. And the intelligent system, artificial neural networks with random theory are integrated in this work. The method is different from the methods used in previous papers [13, 14, 20], which also investigate the relationships between the crude oil market and the stock market. This paper also extends the method mentioned in [3] by introducing the random jump process, which can make the model have the effect of random jump while maintaining the original trend. And we do the different statistical analysis with the work in [3]. In the present paper, we improve the forecasting method in the neural network, each

historical datum is given a weight (random with jump) depending on the time it occurs in the model, and we also use the probability density functions to classify the various variables from the training samples. The empirical research exhibits that the improved neural network model takes advantage over the traditional neural network models to some degree.

2. A Brief Description of Oil Market and Stock Market in China

Chinese oil market is attracting more and more attentions from all over the world. China has been the world's second-largest oil consumer since 2003, and its oil demand reached 9% of the world's total demand in 2006. Figure 1 shows the monthly output and the monthly growth rate of the crude oil production in China from January 2003 to December 2009. The plot indicates that the crude oil output has almost reached the high limit, whereas the oil demand will grow by 4.5% in the coming three years. This displays that the stronger relationships between the international oil market and Chinese oil market become obvious.

In fact, China has become a net importer of crude oil since 1996; and the import dependence has exceeded 51% in 2008. Figures 2(a) and 2(b) present China's crude oil import and consumption monthly in the recent 7 years. The plots exhibit that the trends of the curves in Figures 2(a) and 2(b) are similar, which implies that the oil demand relies heavily on the international oil market. At the same time, the total values of China stock markets A shares reached 3.21 trillion US dollars on July 15, 2009, ranking as the world's second-largest stock market. The listed oil companies usually are the large cap companies, so the market capitalization value of these companies is not only a main part of the stock market value but also an important component of the stock market indices. Although some research work has been done in studying the relationship between the crude oil market and the stock market [4, 13, 14, 20–22], there has been relatively little empirical work done to analyze the relationships in China. In this paper, we select the data of SHCI, SZCI, SZPI, Daqing (Daqing crude oil price), Shengli (Shengli crude oil price), and the price of SINOPEC for each trading day in 7-year period from January 2, 2003 to December 31, 2009. And the corresponding statistical behaviors and comparisons of prices changes are studied in the following.

3. Forecasting and Statistical Analysis

In the real crude oil market, understanding the process by which oil prices evolve is fundamental to our knowledge of this market. Many empirical evidences, like the asymmetric and leptokurtic feature of return distributions and volatilities, strongly suggested an inappropriateness for the usage of Brownian motions in the Black-Scholes model. More precisely, it is often observed that the return distribution is skewed to zero and has a higher peak and fatter tails than those of the corresponding normal distribution. To explain those empirical phenomena, many researches propose innovative models such as normal jump diffusion models (see [12–15]), and continuous-time stochastic volatility models are becoming an increasingly popular way to describe moderate-and high-frequency financial data. These models introduce discontinuities, or jumps, into the volatility process, this can improve the empirical performance of these models. The distribution behavior of jumps for oil prices often represent an important piece of the temporal crude oil price dynamics. We establish the presence of jumps in the data of the financial model, where the jumps that disrupt the entire term structure represent the most significant jump events. For example, in the present paper, these jump events may include the changing of international energy

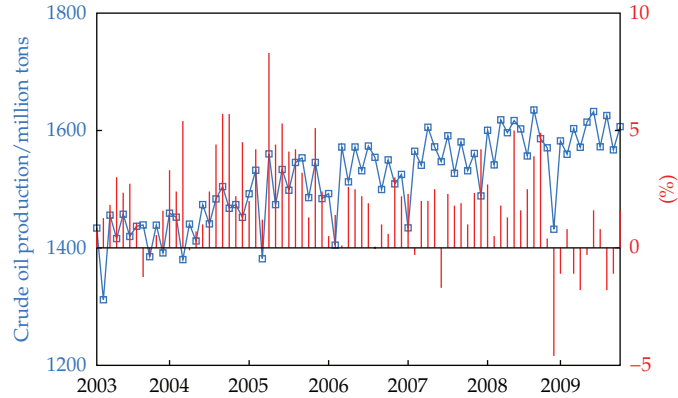


Figure 1: The output and the growth rate of crude oil in China.

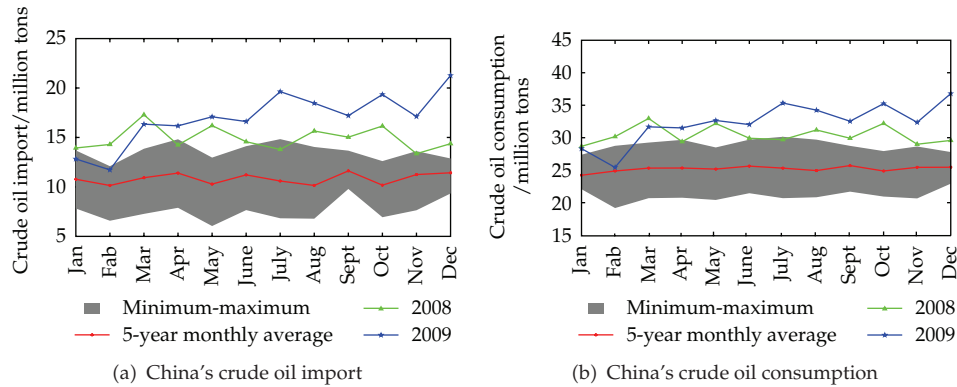


Figure 2

markets, the amount of oil production in China, the crude oil reserve in China, Chinese oil consumption, Chinese energy policy, the wars, and the political events in the world, so on. These random events may be responsible for generating jumps in crude oil price dynamics. Since the fluctuation behaviors of the crude oil prices are also nonlinear, unstable, and random, we introduce the stochastic time effective function in the neural network. The function is supposed to follow a Brownian motion plus a compound Poisson process with a random jump distribution, in order to describe the above-mentioned empirical phenomena. We assume that the historical data of the crude oil market can reflect these random events, and affect the price volatility of the current oil market. For the model, the proposed stochastic time effective function may reflect the large fluctuations of the oil prices. Further, the function is a time-dependent random variable and also shows that the recent information has a stronger effect than the old information for the investors.

3.1. Jump Stochastic Time Effective Neural Network Model for Forecasting

There are various methods to forecast the volatilities of the time series, for example, the autoregressive conditional heteroscedasticity model has been applied by many financial

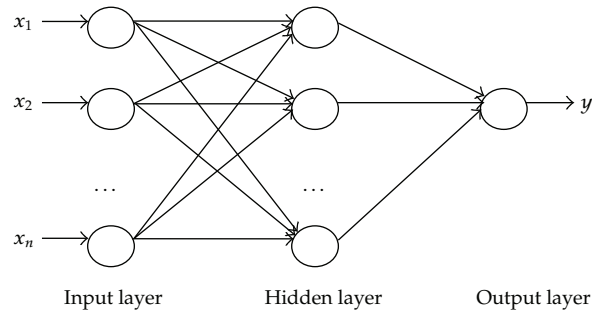


Figure 3: The plot of three-layer neural network.

analysts [23]. These financial time series models are based on the financial theories and require some strict assumptions on the distributions of the time series, so sometimes it is hard to reflect the market variables directly in the models. Usually stock prices can be seen as a random time sequence with noise, artificial neural networks, as large-scale parallel processing nonlinear systems that depend on their own intrinsic link data, providing methods and techniques that can approximate any nonlinear continuous function, without a priori assumptions about the nature of the generating process. The ANN model is a nonparametric method and can forecast future results by learning the pattern of market variables without any strict theoretical assumption [11]. Brooks demonstrated that it is applicable to forecast the volatilities of the financial time series by ANN [24].

First we introduce the three-layer BP neural network model in Figure 3, (for the details see [8–10]), and for any fixed neuron n ($n = 1, 2, \dots, N$), the model has the following structure: let $\{x_i(n) : i = 1, 2, \dots, p\}$ denote the set of input of neurons, $\{y_j(n) : j = 1, 2, \dots, m\}$ denote the set of output of hidden layer neurons, V_i is weight that connects the node in i the input layer neurons to the node j in the hidden layer, W_j is weight that connects the node j in the hidden layer neurons to the node k in the output layer, and $\{o_k(n) : k = 1, 2, \dots, q\}$ denote the set of output of neurons. Then the output value for a unit is given by the following function

$$y_j(n) = f\left(\sum_{i=1}^p V_i x_i(n) - \theta_j\right), \quad o_k(n) = f\left(\sum_{j=1}^m W_j y_j(n) - \theta_k\right), \quad (3.1)$$

where θ_j, θ_k are the neural thresholds, and $f(x) = 1/(1 + e^{-x})$ is Sigmoid activation function. Let $T_k(n)$ be the actual value of data sets, then the error of the corresponding neuron k to the output is defined as $\varepsilon_k = T_k - o_k$.

Obviously, the real data follow normal distribution ingeneral. However, the tail of the real distribution is fatter than the normal, which is called fat-tail phenomena. It is caused by drastic fluctuation of stock price. Moreover, we can find that the log return of stock price will fluctuate rapidly at intervals. In view of the above reality problem, the error of the output is defined as $\varepsilon = \varepsilon_k^2/2$, then the error of the sample n ($n = 1, 2, \dots, N$) is defined as

$$e(n, t) = \frac{1}{2} \phi(t) \sum_{k=1}^q (T_k(n) - o_k(n))^2, \quad (3.2)$$

where $\phi(t)$ is the jump stochastic time effective function. Now we defined $\phi(t)$ as follows

$$\phi(t_1 - t_n) = \frac{1}{\tau} \exp \left\{ - \int_{t_n}^{t_1} \mu(t) dt - \int_{t_n}^{t_1} \sigma(t) dB(t) + \sum_{l=1}^{N(t_1-t_n)} J_l \right\}, \quad (3.3)$$

where $\tau(> 0)$ is the time strength coefficient, t_1 is the current time or the time of newest data in data set, and t_n is an arbitrary time point in data set. J_l ($l = 1, 2, \dots, N(t)$) are independent and identically distributed jump processes and J_l obey the normal distribution with mean μ_J and variance σ_J . $N(t)$ ($t \geq 0$) is a Poisson process with intensity λ . $\mu(t)$ is the drift function (or the trend term), $\sigma(t)$ is the volatility function, and $B(t)$ is the standard Brownian motion [5]. The stochastic time effective function implies that the recent information has a stronger effect for the investors than the old information. In detail, the nearer the events happened, the greater the investors and market are affected. Then the total error of all data training set in the set output layer with the jump stochastic time effective function is defined as

$$\begin{aligned} E &= \frac{1}{N} \sum_{n=1}^N e(n, t) \\ &= \frac{1}{N} \sum_{n=1}^N \frac{1}{\tau} e^{-\int_{t_n}^{t_1} \mu(t) dt - \int_{t_n}^{t_1} \sigma(t) dB(t) + \sum_{l=1}^{N(t_1-t_n)} J_l} \sum_{k=1}^q \frac{1}{2} (T_k(n) - o_k(n))^2. \end{aligned} \quad (3.4)$$

Data is divided into two sections: the data from 2003 to 2007 is used for training and the rest is used for testing. For the stock indices, we input five kinds of stock prices: daily open price, daily closed price, daily highest price, daily lowest price, and daily trade volume, and one price of stock prices in the output layer: the closed price of the next trade day. And for the crude oil prices, we input five kinds of prices: the crude oil price of Brent, WTI, Dubai, Daqing, and Shengli, and the crude oil price of Daqing (or Shengli) of the next trade day is in the output layer. The number of neural nodes in input layer is 5, the number of neural nodes in the hidden layer is 13, and the number of neural nodes in output layer is 1. In this section, we take μ_J and σ_J to be the mean and the variance of reality historical data of SHCI, and let the intensity λ be $1/30$. That is to say, jump will happen 10 times a year in average. Moreover, we suppose that the values of vector $(\mu(t), \sigma(t))$ are $(1, 1)$. The training algorithms procedures of the neural network is described as follows.

Step 1. Normalize the data as follows: $S(t)' = (S(t) - \min S(t)) / (\max S(t) - \min S(t))$.

Step 2. At the beginning of data processing, connective weights V_i and W_j follow the uniform distribution on $(-1, 1)$, and let the neural threshold θ_k, θ_j be 0.

Step 3. Introducing the jump stochastic time effective function $\phi(t)$ in the error function $e(n, t)$. Choosing different volatility parameter. Giving the transfer function from input layer to hidden layer and the transfer function from hidden layer to output layer.

Step 4. Establishing an error-acceptable model and setting preset minimum error. If output error is below preset minimum error, go to Step 6, otherwise go to Step 5.

Table 1: Linear regression parameters.

Parameter	SHCI	SZCI	SZPI	Daqing	Shengli	SINOPEC
a	0.9940	0.9715	0.9107	0.9217	0.8776	0.9934
b	8.5849	510.7603	91.7800	4.9774	5.6977	0.1957
r	0.9806	0.9824	0.9749	0.9942	0.9941	0.9792

Step 5. Modify connective weights by calculating backward for the node in output layer:

$$\delta_o(n) = \frac{1}{\tau} e^{-\int_{t_n}^{t_1} \mu(t) dt - \int_{t_n}^{t_1} \sigma(t) dB(t) + \sum_{l=1}^{N(t_1-t_n)} J_l} o(n)[o(n) - T(n)][1 - o(n)]. \quad (3.5)$$

Calculate δ backward for the node in hidden layer:

$$\delta_h(n) = \frac{1}{\tau} e^{-\int_{t_n}^{t_1} \mu(t) dt - \int_{t_n}^{t_1} \sigma(t) dB(t) + \sum_{l=1}^{N(t_1-t_n)} J_l} o(n)[1 - o(n)] \sum_{h'} W_j \delta_{h'}(n), \quad (3.6)$$

where $o(n)$ is the output of the neuron n , $T(n)$ is the actual value of the neuron n in data sets, $o(n)[1 - o(n)]$ is the derivative of the sigmoid activation function and h' is each of the node which connect with the node h and in the next hidden layer after node h . Modifying the weights from this layer to the previous layer:

$$W_j(n+1) = W_j(n) + \eta \delta_o(n) y(n) \quad \text{or} \quad V_j(n+1) = V_j(n) + \eta \delta_k(n) x(n), \quad (3.7)$$

where η is learning step, which usually take constants between 0 and 1.

Step 6. Output the predictive value.

Next, according to the computer simulations of the given neural network model, we do the comparisons between the predictive data of the model and the actual data of SHCI, SZCI, SZPI, Daqing, Shengli, and SINOPEC. And these comparison results are plotted in Figure 4.

In Figure 5, by using the linear regression method, we compare the predictive data of the neural network model with the actual data of SHCI, SZCI, SZPI, Daqing, Shengli, and SINOPEC. It is known that the linear regression attempts to model the relationship between two variables by fitting a linear equation to observed data. And it is usually used to fit a predictive model to an observed data set of two variables. Through the regression analysis, there are different linear equations in SHCI, SZCI, SZPI, Daqing, Shengli, and SINOPEC respectively, in Figure 5. We set the predictive data as x -axis and set the actual data as y -axis, and the linear equation is $y = ax + b$. A valuable numerical measure of association between two variables is the correlation coefficient r . Table 1 shows the values of a , b , and r for the indices.

3.2. Experiment Analysis

In Section 3.1, the financial price model is modeled by the neural network system. In order to evaluate the prediction of the model, we introduce some statistics in this section: mean

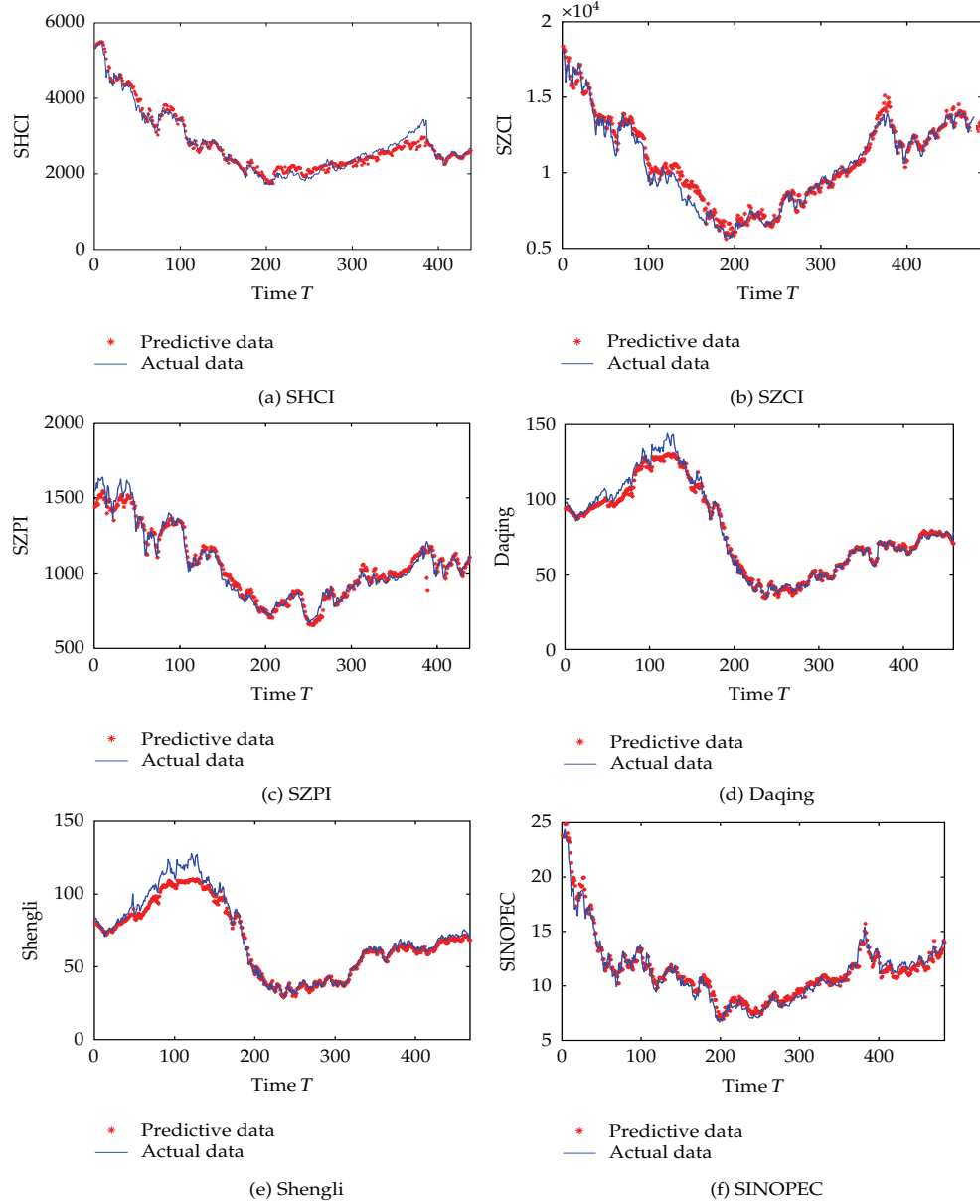


Figure 4: Comparisons of the predictive data and the actual data.

absolute error (MAE), mean relative error (MRE), Theil inequality coefficient (Theil's IC), bias proportion (BP), variance proportion (VP) and covariance proportion (CP). We set x_i , y_i , \bar{x} , \bar{y} , σ_x , σ_y , and r as the predictive value, the actual value, the mean of the predictive value, the mean of the actual value, and the variance of the predictive value, the variance of the actual value and the correlation, respectively. These statistics are defined as follows:

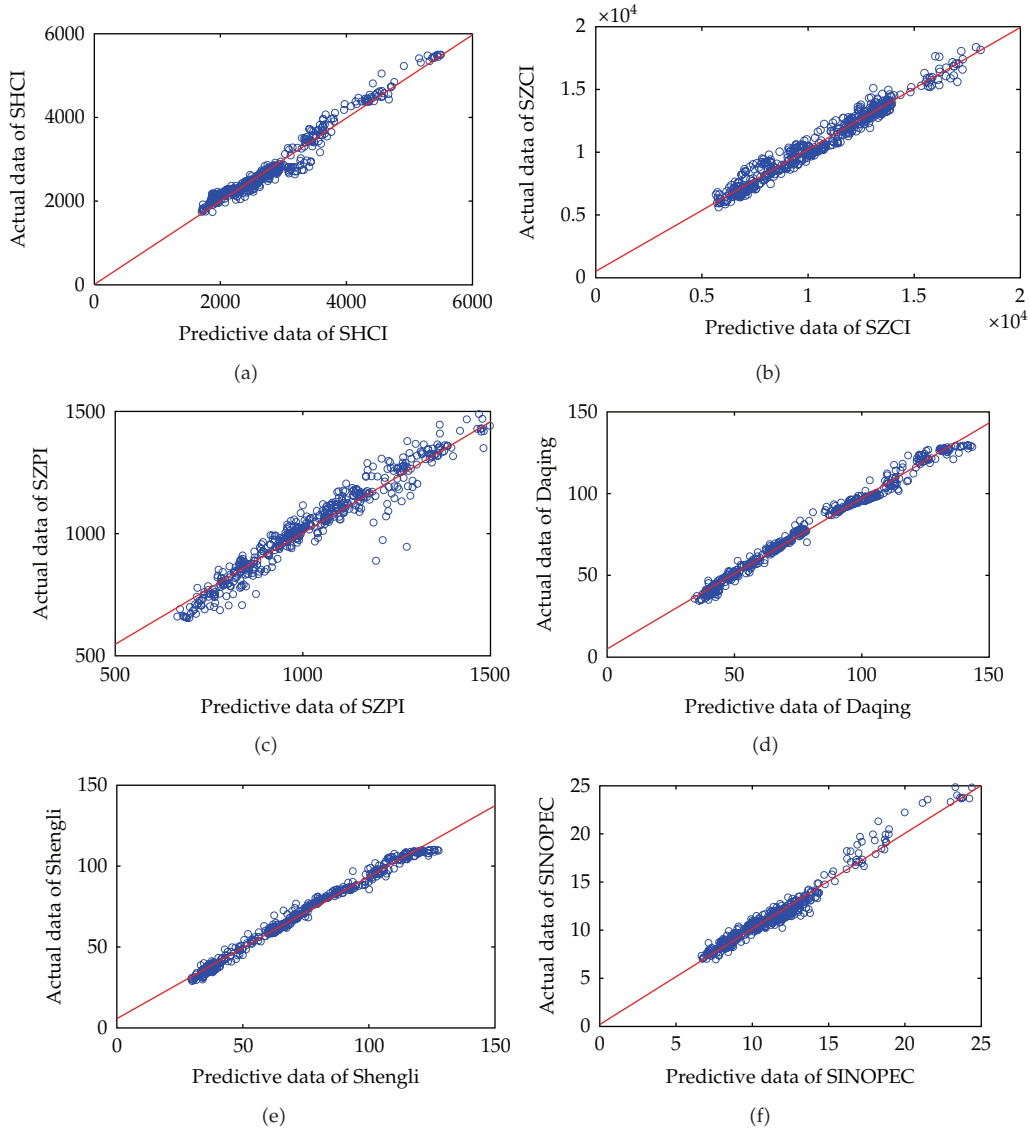


Figure 5: Regressions of the predictive data and the real data.

$$\begin{aligned}
 \text{MAE} &= \frac{1}{n} \sum_{i=1}^n |x_i - y_i|, & \text{MRE} &= \frac{1}{n} \sum_{i=1}^n \left| \frac{x_i - y_i}{y_i} \right|, \\
 \text{Theil's IC} &= \frac{\sqrt{(1/n) \sum_{i=1}^n (x_i - y_i)^2}}{\sqrt{(1/n) \sum_{i=1}^n x_i^2 + (1/n) \sum_{i=1}^n y_i^2}},
 \end{aligned} \tag{3.8}$$

where the value of Theil IC is in $[0, 1]$, and the smaller value means the better prediction of the model.

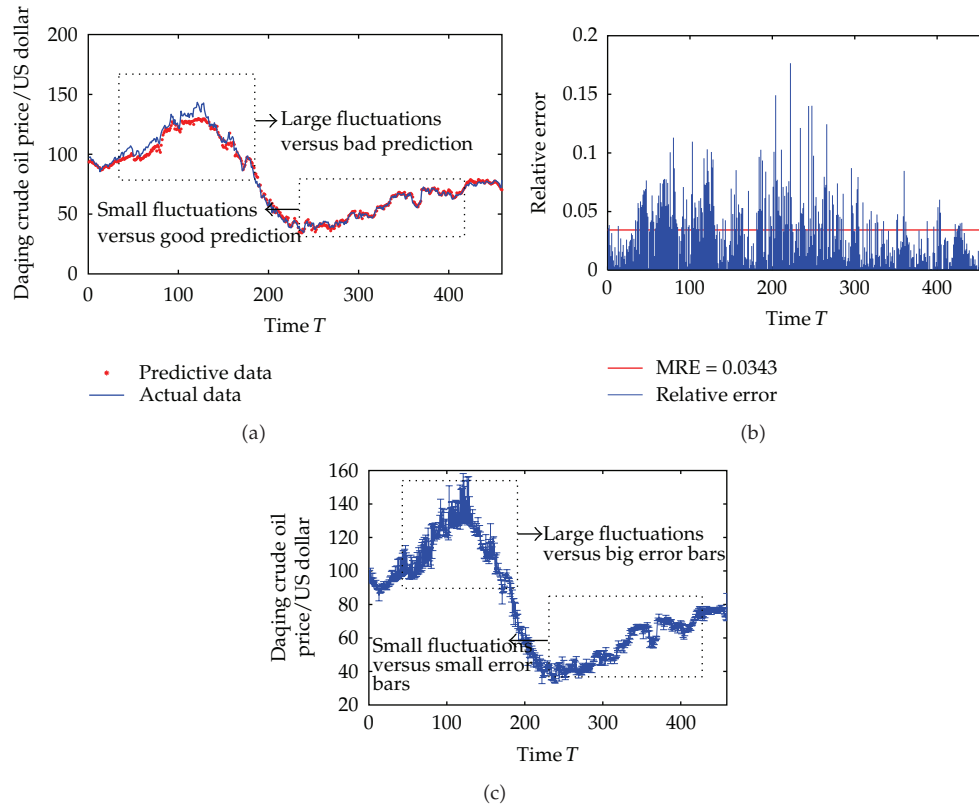


Figure 6: Comparisons of the fluctuation and the prediction of Daqing.

$$\begin{aligned}
 BP &= \frac{(\bar{x} - \bar{y})^2}{\sum_{i=1}^n (x_i - y_i)^2 / n}, & VP &= \frac{(\sigma_x - \sigma_y)^2}{\sum_{i=1}^n (x_i - y_i)^2 / n}, \\
 CP &= \frac{2(1-r)\sigma_x\sigma_y}{\sum_{i=1}^n (x_i - y_i)^2 / n} = 1 - BP - VP,
 \end{aligned} \tag{3.9}$$

where BP denotes the normalized difference between the mean of the predictive value and the mean of the actual value, and VP denotes the normalized difference between the variance of the predictive value and the variance of the actual value. Their values range from 0 to 1. The prediction of the model is effective when the value of CP is close to 1. From the computer computation, Table 2 presents the values of the above statistics. Table 2 also gives a description of the deviating degrees between the predictive data and the actual data.

In the next part, we will discuss the relationship between the crude oil price fluctuation of Daqing and the predictive values of the model. It is apparent in Figure 6(a) when the fluctuation is small, the predictive values are close to the actual values. In another aspect, when the fluctuation is large, the predictive values deviate from the actual values in some extent. We also can see in Figures 6(b) and 6(c) that the small fluctuation leads to the small relative errors and the small errorbars and the large fluctuation leads to the big relative errors and the big errorbars. So there is a relationship between the fluctuation and the prediction. To

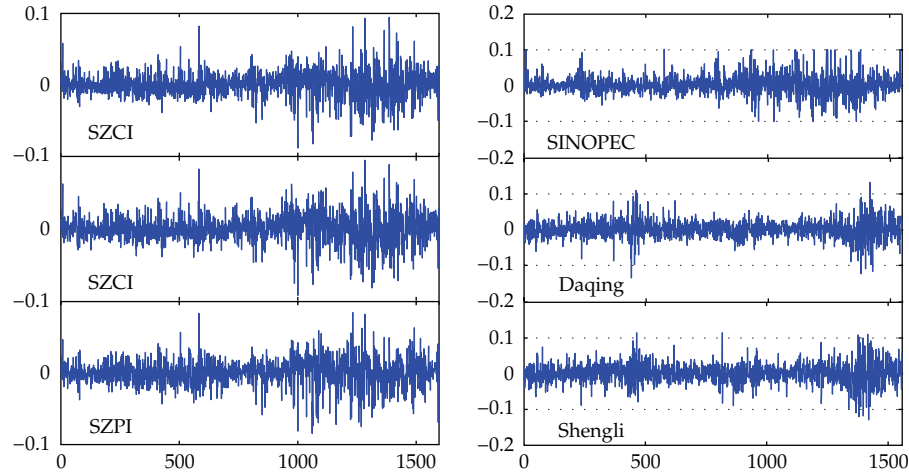


Figure 7: Returns of the indices in the 7-year period from January 2003 to December 2009.

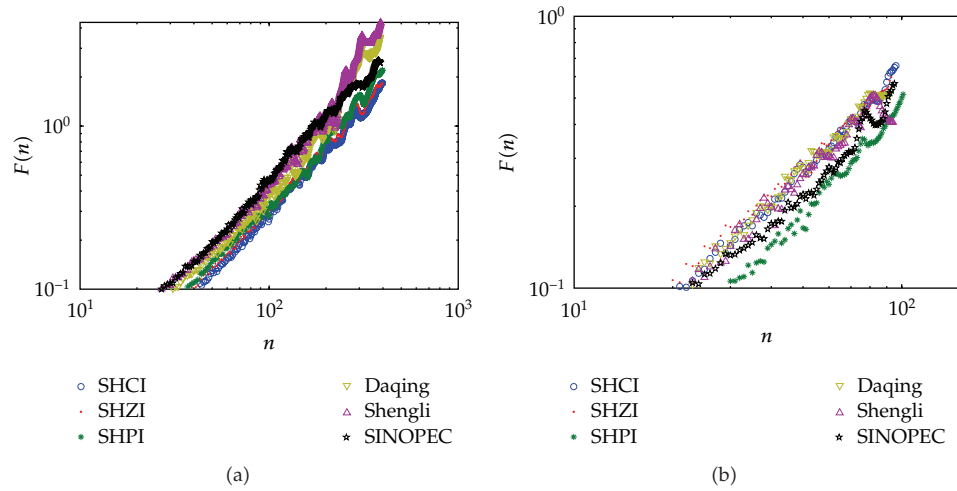


Figure 8: Detrended fluctuation analysis for the absolute returns of the actual data and the predictive data. (a) The plot of the absolute returns for the actual data from January 2003 to December 2009. (b) The plot of the absolute returns of the predictive data from January 2008 to December 2009.

Table 2: Evaluation of the prediction.

	SHCI	SZCI	SZPI	Daqing	Shengli	SINOPEC
MAE	116.5679	433.4098	37.8812	2.8028	3.6815	0.4957
MRE	0.0431	0.0448	0.0343	0.0349	0.0466	0.0448
Theil's IC	0.0269	0.0260	0.0236	0.0348	0.0367	0.0273
BP	$1.1760e-8$	$5.6445e-7$	$1.8592e-8$	$4.4732e-7$	$1.5301e-6$	$1.5276e-7$
VP	0.0496	0.4514	0.15795	0.0047	0.0045	$7.8972e-7$
CP	0.9504	0.5486	0.84205	0.9942	0.9955	1.0000

Table 3: The relationship between the fluctuation and the prediction by the absolute return intervals.

$ R(t) $ intervals	MAE	MRE
$[0, 0.01)$	0.0312	17.3123
$[0.01, 0.02)$	0.0363	16.4007
$[0.02, 0.03)$	0.0401	16.3452
$[0.03, 0.04)$	0.0387	14.8879
$[0.04, M]$	0.0376	26.2455
$[0, M]$	0.0382	18.2984

Table 4: Returns statistics of the real data.

	SHCI	SZCI	SZPI	Daqing	Shengli	SINOPEC
Mean	$7.6353e-4$	$1.2217e-3$	$8.7236e-4$	$6.8830e-4$	$8.4244e-4$	$1.5475e-3$
Variance	$3.3633e-4$	$3.9532e-4$	$3.6080e-4$	$6.3243e-4$	$7.9565e-4$	$7.4543e-4$
Skewness	-0.0779	-0.1274	-0.3832	-0.1276	-0.0643	0.3245
Kurtosis	3.1675	2.4920	2.2832	3.1553	1.8875	2.4495
Minimum	-0.0884	-0.0932	-0.0844	-0.1352	-0.1263	-0.1030
Maximum	0.0954	0.0963	0.0843	0.1323	0.1146	0.1015

investigate this relationship, we choose the predictive values and the actual values of Daqing as the research object. First, we measure the fluctuation in absolute returns, which is denoted by $|R(t)|$. Then we divide the data into five groups by the absolute return intervals. The intervals are $[0, 0.01)$, $[0.01, 0.02)$, $[0.02, 0.03)$, $[0.03, 0.04)$, and $[0.04, M]$, where M denotes the maximum of absolute returns. Table 3 shows the relationship between the actual fluctuation and the prediction by the absolute return intervals.

3.3. Return Analysis

In this section, we discuss the statistical properties of SHCI, SHZI, SZPI, Daqing, Shengli, and SINOPEC in the 7-year period from January 2003 to December 2009. Figure 7 presents the figures of the returns time sequence for these indices. We denote the daily price at time t by $S(t)$ ($t = 0, 1, 2, \dots$), then the return of the stock price (or index) is given by

$$R(t) = \frac{S(t+1) - S(t)}{S(t)} = \frac{S(t+1)}{S(t)} - 1. \quad (3.10)$$

Table 4 presents the statistical analysis of the returns for the actual data. Note that the daily price fluctuation is limited in China, that is, the changing limits of the daily returns for stock prices and stock indices are between 10% and -10%, whereas the returns of the crude oil price can change in a larger value range. Table 5 presents the statistical analysis of the returns for the predictive data. In these two tables, they show the values of mean, variance, kurtosis and skewness of the returns, and we also can compare these values between the actual data and the predictive data.

Table 5: Returns statistics of the predictive data.

	SHCI	SZCI	SZPI	Daqing	Shengli	SINOPEC
Mean	$-9.9940e-4$	$-4.3613e-5$	$2.2083e-4$	$-7.4974e-5$	$-2.0295e-4$	$-4.9588e-4$
Variance	$8.3044e-4$	$1.1969e-3$	$9.5615e-4$	$1.0991e-3$	$1.1038e-3$	$1.2144e-3$
Skewness	0.6257	0.0838	1.0189	0.2120	0.1417	0.5700
Kurtosis	2.6270	2.3022	1.9964	3.1030	2.0811	1.7350
Minimum	-0.08211	-0.1327	-0.2057	-0.1227	-0.1101	-0.1073
Maximum	0.1525	0.1340	0.2815	0.1589	0.1250	0.1597

Table 6: Scaling exponent of the absolute returns.

Scaling exponent	SHCI	SZCI	SZPI	Daqing	Shengli	SINOPEC
α_A	3.5852	3.6523	3.7007	3.8456	3.9760	3.6342
α_P	5.6438	5.7443	5.6002	5.9843	5.7712	5.8732

3.4. Detrended Fluctuation Analysis

Detrended fluctuation analysis (DFA) is a scaling analysis method providing the scaling exponent α to represent the correlation properties [7, 16–18]. There are two advantages in DFA method. One is that it permits the detection of long-range correlations embedded in seemingly nonstationary time series. The other is that it avoids the spurious detection of apparent long-range correlations that are artifact of nonstationarity. Briefly, for a given stochastic time series $S(i)$, $i = 1, 2, \dots, N$, with the sampling period Δt , the DFA method can be implemented as follows.

Step 1. Compute the mean $\bar{S} = (1/N) \sum_{i=1}^N S(i)$ and obtain an integrated time series $y(j) = (1/N) \sum_{i=1}^j (S(i) - \bar{S})$. Then divide the integrated time series into boxes of equal size, n .

Step 2. In each box, fit the integrated time series by using a polynomial function, $y_{\text{fit}}(i)$. For order- l DFA, l order polynomial function should be applied for the fitting and in this paper, $l = 2$. Then calculate the detrended fluctuation function as follows:

$$Y(i) = y(i) - y_{\text{fit}}(i). \quad (3.11)$$

Step 3. For a given box size n , calculate the root mean square fluctuation:

$$F(n) = \left(\frac{1}{N} \sum_{i=1}^N [Y(i)]^2 \right)^{1/2}. \quad (3.12)$$

A power-law relation between $F(n)$ and the box size n indicates the presence of scaling: $F(n) \sim n^\alpha$. The parameter α , called the scaling exponent or correlation exponent, represents the correlation properties of the time series: if $\alpha = 0.5$, there is no correlation and the time series is uncorrelated; if $\alpha < 0.5$, the signal is anticorrelated; if $\alpha > 0.5$, there are positive correlations in the time series.

In this paper, we use DFA to analyze the absolute returns of the actual data and the predictive data, see Figure 8. α_A and α_P denote the scaling exponents of the absolute returns

for the actual data and the predictive data respectively. Table 6 shows that α_A and α_P are all larger than 0.5, which means that there are positive correlations in the absolute returns of the actual data and the predictive data.

4. Conclusion

In this paper, we introduce the jump stochastic time effective neural network model to forecast the fluctuations of SHCI, SZCI, SZPI, Daqing, Shengli, and SINOPEC. The corresponding statistical behaviors of these indices are investigated; and several kinds of comparisons between the actual data and the predictive data are given. Further, the absolute returns of the actual data and the predictive data are studied by the statistical method and the detrended fluctuation analysis.

Acknowledgments

The authors were supported in part by National Natural Science Foundation of China Grant nos. 70771006 and 10971010, and BJTU Foundation grant no. S11M00010.

References

- [1] R. G. Cong, Y. M. Wei, J. L. Jiao, and Y. Fan, "Relationships between oil price shocks and stock market: an empirical analysis from China," *Energy Policy*, vol. 36, no. 9, pp. 3544–3553, 2008.
- [2] M. F. Ji and J. Wang, "Data analysis and statistical properties of Shenzhen and Shanghai land indices," *WSEAS Transactions on Business and Economics*, vol. 4, pp. 33–39, 2007.
- [3] Z. Liao and J. Wang, "Forecasting model of global stock index by stochastic time effective neural network," *Expert Systems with Applications*, vol. 37, no. 1, pp. 834–841, 2010.
- [4] T. C. Mills, *The Econometric Modelling of Financial Time Series*, Cambridge University Press, Cambridge, UK, Second edition, 1999.
- [5] J. Wang, *Stochastic Process and Its Application in Finance*, Tsinghua University Press and Beijing Jiaotong University Press, Beijing, China, 2007.
- [6] J. Wang, Q. Wang, and J. Shao, "Fluctuations of stock price model by statistical physics systems," *Mathematical and Computer Modelling*, vol. 51, no. 5-6, pp. 431–440, 2010.
- [7] T. Wang, J. Wang, and B. Fan, "Statistical analysis by statistical physics model for the stock markets," *International Journal of Modern Physics C*, vol. 20, no. 10, pp. 1547–1562, 2009.
- [8] E. M. Azoff, *Neural Network Time Series Forecasting of Financial Market*, Wiley, New York, NY, USA, 1994.
- [9] M. Demuth and M. Beale, *Neural Network Toolbox: For Use with MATLAB*, The Math Works, Inc., Natick, Mass, USA, 5th edition, 1998.
- [10] V. S. Desai and R. Bharati, "The efficacy of neural networks in predicting returns on stock and bond indices," *Decision Sciences*, vol. 29, no. 2, pp. 405–423, 1998.
- [11] T. Hyup Roh, "Forecasting the volatility of stock price index," *Expert Systems with Applications*, vol. 33, no. 4, pp. 916–922, 2007.
- [12] O. B. Nielsen and N. Shephard, "Power and bipower variation with stochastic volatility and jumps," *Journal of Financial Econometrics*, vol. 2004, no. 2, pp. 1–48, 2004.
- [13] U. Oberndorfer, "Energy prices, volatility, and the stock market: evidence from the Eurozone," *Energy Policy*, vol. 37, no. 12, pp. 5787–5795, 2009.
- [14] E. Papapetrou, "Oil price shocks, stock market, economic activity and employment in Greece," *Energy Economics*, vol. 23, no. 5, pp. 511–532, 2001.
- [15] D. Pirino, "Jump detection and long range dependence," *Physica A*, vol. 388, no. 7, pp. 1150–1156, 2009.
- [16] J. Alvarez-Ramirez, J. Alvarez, and E. Rodriguez, "Short-term predictability of crude oil markets: a detrended fluctuation analysis approach," *Energy Economics*, vol. 30, no. 5, pp. 2645–2656, 2008.

- [17] O. F. Ayadi, J. Williams, and L. M. Hyman, "Fractional dynamic behavior in Forcados Oil Price Series: an application of detrended fluctuation analysis," *Energy for Sustainable Development*, vol. 13, no. 1, pp. 11–17, 2009.
- [18] K. Hu, P. C. Ivanov, Z. Chen, P. Carpena, and H. E. Stanley, "Effect of trends on detrended fluctuation analysis," *Physical Review E*, vol. 64, no. 1, pp. 0111141–01111419, 2001.
- [19] Y. Wang, L. Liu, and R. Gu, "Analysis of efficiency for Shenzhen stock market based on multifractal detrended fluctuation analysis," *International Review of Financial Analysis*, vol. 18, no. 5, pp. 271–276, 2009.
- [20] S. Saif Ghouri, "Assessment of the relationship between oil prices and US oil stocks," *Energy Policy*, vol. 34, no. 17, pp. 3327–3333, 2006.
- [21] R. Gaylord and P. Wellin, *Computer Simulations with Mathematica: Explorations in the Physical, Biological and Social Science*, Springer, New York, NY, USA, 1995.
- [22] K. Ilinski, *Physics Of Finance: Gauge Modeling in Non-Equilibrium Pricing*, John Wiley, New York, NY, USA, 2001.
- [23] R. F. Engle, "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation," *Econometrica*, vol. 50, no. 4, pp. 987–1007, 1982.
- [24] C. Brooks, "Predicting stock index volatility: can market volume help?" *Journal of Forecasting*, vol. 17, pp. 59–80, 1998.

