

Research Article

Integrating Independent Component Analysis and Principal Component Analysis with Neural Network to Predict Chinese Stock Market

Haifan Liu and Jun Wang

*Institute of Financial Mathematics and Financial Engineering, College of Science,
Beijing Jiaotong University, Beijing 100044, China*

Correspondence should be addressed to Haifan Liu, haifanliu@yeah.net

Received 24 March 2011; Revised 16 May 2011; Accepted 17 May 2011

Academic Editor: Kuppapalle Vajravelu

Copyright © 2011 H. Liu and J. Wang. 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.

We investigate the statistical behaviors of Chinese stock market fluctuations by independent component analysis. The independent component analysis (ICA) method is integrated into the neural network model. The proposed approach uses ICA method to analyze the input data of neural network and can obtain the latent independent components (ICs). After analyzing and removing the IC that represents noise, the rest of ICs are used as the input of neural network. In order to forecast the fluctuations of Chinese stock market, the data of Shanghai Composite Index is selected and analyzed, and we compare the forecasting performance of the proposed model with those of common BP model integrating principal component analysis (PCA) and single BP model. Experimental results show that the proposed model outperforms the other two models no matter in relatively small or relatively large sample, and the performance of BP model integrating PCA is closer to that of the proposed model in relatively large sample. Further, the prediction results on the points where the prices fluctuate violently by the above three models relatively deviate from the corresponding real market data.

1. Introduction

Recently, some progress has been made to investigate the statistical behaviors of the financial market fluctuation, see [1–6] and some forecasting methods of price changes have been developed and studied by using the theory of artificial neural networks [7–9]. Financial market is a complex system which has many influence factors and many kinds of uncertainties, and its fluctuation often represents strong nonlinear characteristics, so that the forecasting of financial time series has long been a focus of financial research. Unlike the traditional time-series analysis methods, such as exponential smoothing, GARCH [10],

and ARIMA [11], artificial neural network which can handle disorderly comprehensive information does not require strong model assumptions and also has good nonlinear approximation, strong self-learning, and self-adaptive abilities. Therefore, the neural network is usually applied to forecast the financial time series; see [12–15]. The most popular neural network training algorithm for financial forecasting is the backpropagation neural network (BP), which has powerful problem-solving ability. When the BP neural network is applied to forecast financial time series, we need to consider the related factors which are correlated and contain a large amount of noise included into the system. The noise in the data could lead to overfitting or underfitting problems. The key of using BP neural network is to eliminate the correlation and noise among the input data as far as possible, so that it can improve the performance of the BP model and the prediction accuracy.

In the present paper, the methods of independent component analysis (ICA) and principal component analysis (PCA) are integrated into BP neural network for forecasting financial time series, which are called ICA-BP model and common PCA-BP model, respectively. Independent component analysis is a recently developed and new method of feature extraction in blind source separation which is the process of separating the source signal from the mixed signal in the case of unknown mixing signal model; see [16]. The explication of ICA is that multichannel observed signals, according to the statistical independence, are decomposed into several independent components (ICs) through the optimization algorithm. ICA can also be used for the financial model to extract latent factors of influencing financial market; for example, see [17]. In this model, the input data is firstly analyzed by using ICA and obtains several latent ICs. After identifying and eliminating the IC that represents the noise, the remaining ICs are conducted as the input of the BP model, so that the noise component of original data is removed and the input data is made independent of each other. The approach of PCA-BP is to extract the principal components (PCs) from the input data according to the PCA method and conduct PCs as the input of BP model which can eliminate redundancies of original information and remove the correlation between the inputs. In order to forecast the fluctuations of Chinese stock market, we compare the forecasting performance of ICA-BP model with those of common PCA-BP model and single BP model by selecting the data of Shanghai Composite Index (SHCI). Index SHCI plays an important role in Chinese financial markets. It can reflect the activity and the trend of Chinese security markets in large degrees, so this will be helpful for us to understand the status of China macroeconomic. The database is from Shanghai Stock Exchange; see www.sse.com.cn. This paper is organized as follows. Section 2 gives a brief introduction about independent component analysis, BP neural network, and principal component analysis. The forecasting models of stock market are described in Section 3. Section 4 presents the experimental results according to the datasets of SHCI.

2. Research Methodology

2.1. Independent Component Analysis

Independent component analysis as a new statistical method is developed in the recent years; for example, see [18, 19]. The ICA model has been widely applied in signal processing, face recognition and feature extraction; see [20, 21]. Kiviluoto and Oja [22] also employed the ICA model to find the fundamental factors influencing the cash flow of the 40 stores belonging to the same retail chain. They found that the cash flow of the retail stores was mainly influenced

by holiday, season, and competitors' strategy. The purpose of ICA is that the observed data will be decomposed linearly into statistically independent components, and ICA method aims at finding several latent unobserved independent source signals. Suppose that the observed signal $x(t)$ is a zero-mean data vector observed at the time t and that the source signal $s(t)$ is a zero-mean vector with the components being mutually independent, such that $x(t) = As(t)$, where A is a full rank linear mixing matrix. The algorithm considers a linear transformation $y = w^T x(t) = w^T As(t) = z^T s(t)$ to obtain the solution of the ICA model, where $z = A^T w$, w is an estimator of a row of the matrix A^{-1} , and y is a linear combination of s_i with z_i showing the weight. Since the two independent random variables are closer to Gaussian distribution than the original variables, $z^T s$ is closer to Gaussian distribution than any s_i . Consider w as a vector that maximizes the non-Gaussian distribution of $w^T x$; it means that $w^T x = z^T s$ equals one of the independent components. A quantitative measure of the non-Gaussian of a random variable y is negentropy which is based on the concept of entropy in information theory. The entropy of random variable could be interpreted as the information degrees of a given observe variable. The negentropy J is defined by

$$J(y) = H(y_{\text{gauss}}) - H(y), \quad (2.1)$$

where $H(y)$ is the entropy, given by $H(y) = -\int p(y) \log p(y) dy$, $p(y)$ is the probability density, and y_{gauss} is a Gaussian random vector having the same covariance matrix as y . The negentropy is always nonnegative, and it is zero if and only if y follows a Gaussian distribution. For the calculation of negentropy, the method of approximate calculation is often applied, which is given as follows:

$$J(y) \approx k_i [E\{G_i(y)\} - E\{G_i(v)\}]^2, \quad (2.2)$$

where k_i are some positive constants, v is a Gaussian variable with zero mean and unit variance, y is a random variable with zero mean and unit variance, and G_i are some nonquadratic function. Even in this case, the approximation is not accurate. Here we use nonquadratic function G , and the approximation is changed into

$$J(y) \propto [E\{G(y)\} - E\{G(v)\}]^2, \quad (2.3)$$

where $G(y) = \exp\{-y^2/2\}$ or $G(y) = (1/\alpha) \log[\cos(\alpha y)]$, $1 \leq \alpha \leq 2$, and α is some appropriate constant. If the signals are repeatedly observed, the observed signals are denoted as original signals matrix X . When the matrix W , the separate matrix, is the inverse of the mixing matrix A , the independent component matrix $Y = WX$ could be used to estimate source signals matrix S , where one row of matrix Y is an independent component. In this paper, we apply the FastICA algorithm which is based on fixed-point algorithm and is applicable for any type of data to solve the separate matrix.

2.2. Artificial Neural Network Model

Neural network is a large-scale and nonlinear dynamic system [23], which has the abilities of highly nonlinear operations, self-learning, and self-organizing. Since 1987 Lapedes and Farber applied neural network technology to prediction research firstly in 1987, many

researchers have been engaged in the study of the predicting method of neural network. Azoff [24] also applied neural network to forecast time series of financial market. In the financial field, neural network is often applied to predict the closing stock price of the next trading day according to the history data. The stock data of the last trading day, including daily open price, daily closing price, daily highest price, daily lowest price, daily volume (stock trading amount), and daily turnover (stock trading money), are very important indicators. We can apply the history indicators as the input of neural network, the closing price of the next trading day as the output to predict stock price.

In practice, feed-forward neural network, which can be thought of as highly nonlinear mapping from the input to the output, is usually adopted to predict. Since the three-layer feed-forward neural network possesses the capability that it can be approximated to any complicated continuous function, it is suitable for time series prediction. BP neural network that is characterized by the error backpropagation is a kind of multilayer feed-ward neural network. A three-layer BP neural network which contains input layer, one hidden layer, and output layer is chosen in this study. Figure 1 shows the corresponding topological structure. The training of BP neural network is as follows: for the neuron j , its input I_j and output O_j are calculated with the following formula:

$$I_j = \sum_i w_{ij} O_i, \quad O_j = f(I_j + \theta_j), \quad (2.4)$$

where w_{ij} is the weight of the connection from the i th neuron in the previous layer to the neuron j , $f(x) = (2/1 - e^{-2x}) - 1$ is the activation function of the neurons, and θ_j is the bias input to the neuron. The error E in the output is calculated with the following formula:

$$E = \frac{1}{2n} \sum_n \sum_l \sqrt{(T_{nl} - O_{nl})^2}, \quad (2.5)$$

where n is the number of training set, l is the number of output nodes, O_{nl} is the output value, and T_{nl} is the target value. When the error E falls below the threshold or tolerance level, the training will end. The error δ_l in output layer and the error δ_j in hidden layer are calculated according to the following formula:

$$\delta_l = \lambda(T_l - O_l) f'(O_l), \quad \delta_j = \lambda \sum_i \delta_l w_{ji} f'(O_j), \quad (2.6)$$

where T_l is the expected output of the l th output neuron, O_l is the actual output in the output layer, O_j is the actual output value in the hidden layer, and λ is the adjustable variable in the activation function. The weights and biases in both output and hidden layers are updated with back propagation error. The weights w_{ji} and biases θ_i are adjusted as follows:

$$w_{ji}(k+1) = w_{ji}(k) + \eta \delta_j O_i, \quad \theta_i(k+1) = \theta_i(k) + \eta \delta_i, \quad (2.7)$$

where k is the number of the epoch and η is the learning rate.

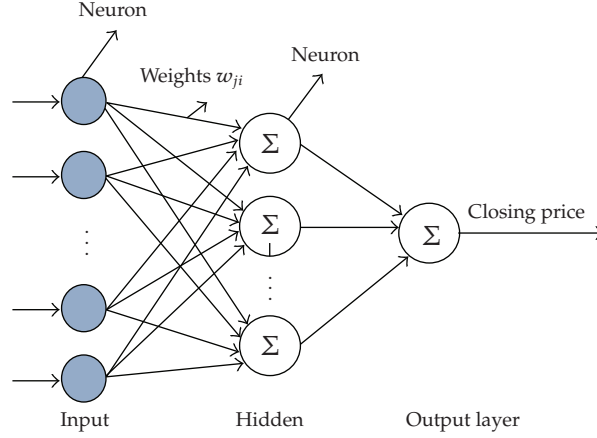


Figure 1: A three-layer BP neural network.

2.3. Principal Component Analysis

PCA is a well-established technique for feature extraction and dimensionality reduction. The basic concept of PCA is to use fewer indexes to replace and comprehensively reflect the original more information, and these comprehensive indexes are the principal components. Ouyang [25] used PCA method to evaluate the ambient water quality monitoring stations located in the main stem of the LSJR. The outcome showed that the number of monitoring stations can be reduced from 22 to 19. Yang et al. [26] built a prediction model for the occurrence of paddy stem borer based on BP neural network, and they applied the PCA approach to create fewer factors to be the input variables for the neural network. Because the essence of PCA is the rotation of space coordinates that does not change the data structure, the obtained PCs are the linear combination of variables, reflect the original information to the greatest degree, and are uncorrelated with each other. The specific steps are as follows: assume the data matrix with m variables, X_1, X_2, \dots, X_m , n times observations

$$X = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1m} \\ X_{21} & X_{22} & \cdots & X_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{nm} \end{bmatrix}. \quad (2.8)$$

Firstly, we normalize the original data by using the following method:

$$Y_{ti} = \frac{(X_{ti} - \bar{X}_t)}{S_t}, \quad (2.9)$$

where $\bar{X}_t = (1/n) \sum_{i=1}^n X_{ti}$ and $S_t = \sqrt{(1/n-1) \sum_{i=1}^n (X_{ti} - \bar{X}_t)^2}$. For convenience, the normalized Y_{ti} is still denoted as X_{ti} . Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 0$ be the eigenvalues of covariance matrix of normalized data. Also let $\alpha_1, \alpha_2, \dots, \alpha_m$ be the corresponding

Table 1: Pearson's correlation coefficient of six time series.

	x_1	x_2	x_3	x_4	x_5	x_6
x_1	1.000	1.000	1.000	0.999	0.599	0.773
x_2	1.000	1.000	1.000	1.000	0.595	0.770
x_3	1.000	1.000	1.000	1.000	0.600	0.774
x_4	0.999	1.000	1.000	1.000	0.596	0.771
x_5	0.599	0.595	0.600	0.596	1.000	0.944
x_6	0.773	0.770	0.774	0.771	0.944	1.000

Table 2: The PCA result of six time series.

Component	Eigenvalue	Contribution rate (%)	Cumulative contribution rate (%)
1	4.456	74.270	74.270
2	1.522	25.369	99.639
3	0.013	0.212	99.851
4	0.006	0.100	99.950
5	0.003	0.044	99.994
6	0.000	0.006	100.000

eigenvector; the i th principal component is such that $F_i = \alpha_i^T X$, where $i = 1, 2, \dots, m$. Generally, $\lambda_k / \sum_{i=1}^m \lambda_i$ is called the contribution rate of the k th principal component and $\sum_{i=1}^k \lambda_i / \sum_{i=1}^m \lambda_i$ is called the cumulative contribution rate of the first k principal components. If the cumulative contribution rate exceeds 85%, the first k principal components contain the most information of m original variables.

3. The Forecasting Models of Stock Market

3.1. Common PCA-BP Forecasting Model

The BP neural network model requires that the input variables should have worse correlation because the better correlation between input variables implies that they carry more repeated information, and it may lead to increasing the computational complexity and reducing the prediction accuracy of the model. The concept of the common PCA-BP forecasting model is explained as follows; for more details, see [14, 26]. Firstly, use PCA method to extract the principal components from the input data of BP neural network, and then conduct the principal components as the input of the BP neural network. The following example is to illustrate how to extract the principal components from the input data using the method of PCA. Six financial time series are denoted as x_1, x_2, x_3, x_4, x_5 , and x_6 , the size is or each 1×400 . Table 1 exhibits the correlation which is measured by Pearson that correlation coefficient. From Table 1, we can clearly see that the correlation between the six time series is obvious; it means that they contain more repeated information.

Table 2 shows the PCA result on six time series. It indicates that the cumulative contribution rate of the first two PCs exceeds 99%, namely, the first two PCs contain 99% information of the original data. The two PCs are, respectively, recorded as F_1 and F_2 which are conducted as the input of the PCA-BP model instead of the original data.

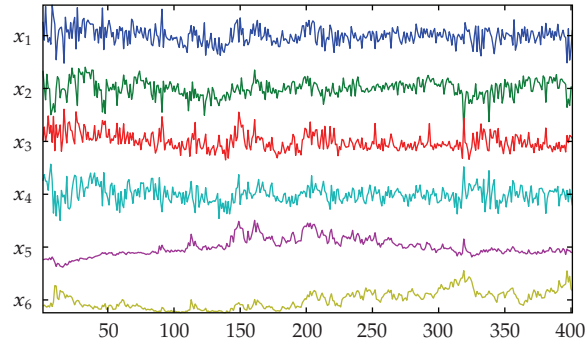


Figure 2: Six time series data the size of each is 1×400 .

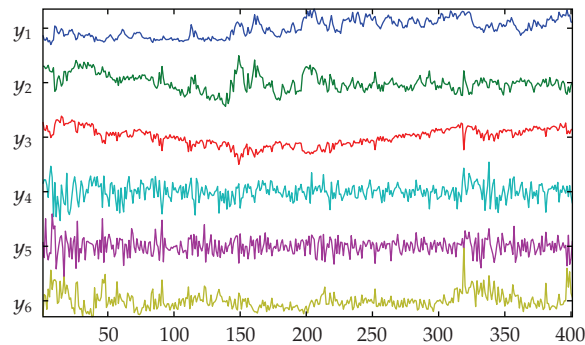


Figure 3: Six ICs of time series data in Figure 2.

3.2. ICA-BP Forecasting Model

In the proposed ICA-BP model, ICA method is firstly used to extract the independent components from the original signals. The feature of original signals is contained in the ICs; each IC represents a feature. The IC including the least effective information of original signals is the noise IC. After identifying and removing the noise IC, the rest of ICs are conducted as the input of BP model. Here the observed time series represent the original signals. The obtained PCs according to PCA method are only eliminated by the correlation, but the obtained higher-order statistics with the ICA method are also independent of each other. In statistical theory, independent is a stronger condition than uncorrelated. The key of the model is to identify the noise IC after obtaining the latent ICs. The testing-and-acceptance (TnA) method is used to solve the problem in this study; see [27].

Similarly as the above given six time series $x_1, x_2, x_3, x_4, x_5,$ and x_6 , Figure 2 shows the tendencies of six time series. Each of the time series is considered as a row; they can be formed by the matrix X of size 6×400 . By the ICA method, the separate matrix W and the independent component matrix Y can be obtained. Each row of Y , the y_i of size 1×400 , represents an IC. Figure 3 shows the tendencies of the six ICs. It can be seen from Figure 3 that each IC can represent different features of the original time series data in Figure 2.

Now the TnA method is applied to identify the noise IC. To introduce the algorithm of TnA, we consider the obtained m ICs. After excluding one IC for each iteration, the remaining $m - 1$ ICs are used to reconstruct the original signals matrix. Let y_k be the excluded IC and

Table 3: The RHD reconstruction errors.

The IC excluded	The ICs included	RHD value
IC1	IC2 IC3 IC4 IC5 IC6	0.6417
IC2	IC1 IC3 IC4 IC5 IC6	1.1588
IC3	IC1 IC2 IC4 IC5 IC6	1.2835
IC4	IC1 IC2 IC3 IC5 IC6	2.0466
IC5	IC1 IC2 IC3 IC4 IC6	2.7914
IC6	IC1 IC2 IC3 IC4 IC5	1.9036

X^R the reconstructed original signals matrix. X^R can be calculated according to the following equation:

$$X^R = \sum_{i=1, i \neq k}^m a_i y_i, \quad 1 \leq k \leq m, \quad (3.1)$$

where $X^R = [x_1^R, x_2^R, \dots, x_m^R]^T$, x_i^R is the i th reconstructed variable, a_i is the i th column vector of mixing matrix A which is the inverse of separate matrix W , and y_i is the i th IC. Respectively, we consider the cases $k = 1, 2, \dots, m$, that is, repeat m iterations and each IC is excluded once. The reconstruction error, which is measured by using relative hamming distance (RHD) [27], between each reconstructed matrix X^R and the original signals matrix X can be computed. The RHD can be computed as follows:

$$\text{RHD} = \frac{1}{n-1} \sum_{i=1}^{n-1} [R_i(t) - \hat{R}_i(t)]^2, \quad (3.2)$$

where $R_i(t) = \text{sign}[A_i(t+1) - A_i(t)]$, $\hat{R}_i(t) = \text{sign}[P_i(t+1) - P_i(t)]$. Here $\text{sign}[u] = 1$ if $u > 0$, $\text{sign}[u] = 0$ if $u = 0$, and $\text{sign}[u] = -1$ if $u < 0$. A_i is the actual value, P_i is the predicted value, and n is the total number of data points.

The RHD reconstruction error can be used to assess the similarity between the original variables and their corresponding reconstructed variables. When the RHD value is closer to zero, it shows that there is higher similarity between original variables and their corresponding reconstructed variables, that is, the corresponding ICs that are used to reconstruct original variables contain more features of original variables and the eliminated IC contains less effective information. On the contrary, when the RHD value is farther from zero, this means that the similarity between the original variables and their corresponding reconstructed variable is lower, that is, the eliminated IC contains more effective information of original variables. So the reconstruction in which the RHD value is the closest to zero should be found out; the corresponding eliminated IC is the noise IC. In allusion to the given six financial time series, Table 3 shows the RHD reconstruction errors of each iteration.

Table 3 reveals that the value of RHD, which is reconstructed by using IC2, IC3, IC4, IC5, and IC6 and eliminating IC1, is the smallest. It is concluded that the IC1 contains the least information and IC1 represents the noise IC. IC2, IC3, IC4, IC5, and IC6 are conducted as the input of the proposed ICA-BP model.

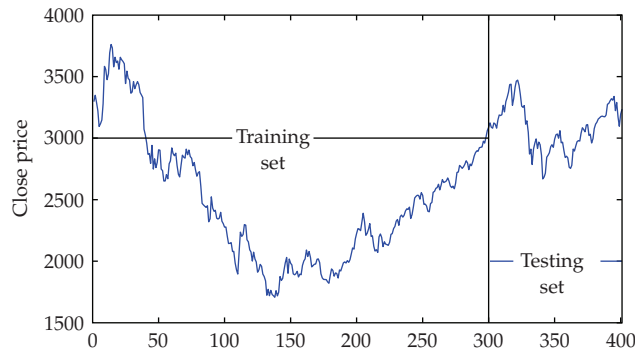


Figure 4: The daily SHCI closing price from April 11, 2008, to November 30, 2009.

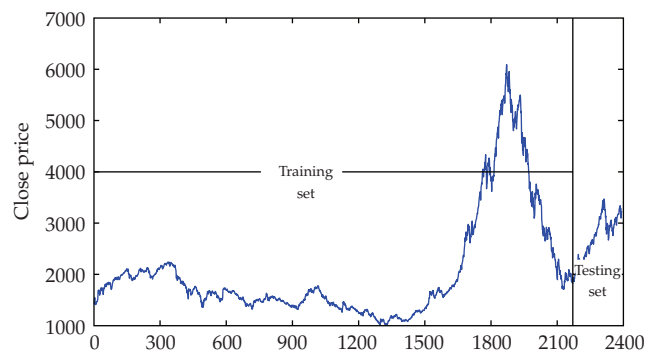


Figure 5: The daily SHCI closing price from January 4, 2000, to November 30, 2009.

4. Empirical Research

4.1. Selection of Datasets

For evaluating the performance of the proposed ICA-BP forecasting model and the common PCA-BP forecasting model, we select the data of Shanghai Composite Index to analyze the models by comparison. In the BP model, the network inputs include six kinds of data, daily open price, daily closing price, daily highest price, daily lowest price, daily volume, and daily turnover. The network outputs include the closing price of the next trading day, because, in stock markets, practical experience shows us that the six kinds of data of the last trading day are very important indicators when we predict the closing price of the next trading day at the technical level. For comparing the performance, two sets of data are used to analyzed, that is, Set 1 and Set 2. Set 1 contains relatively fewer data, that is, the data of SHCI each trading day from April 11, 2008, to November 30, 2009. Figure 4 presents the daily SHCI closing price in this period.

Set 1 includes 400 selected data in which the first 300 data points are used as training set while the rest 100 data points are used as testing set. Set 2 contains relatively more data, that is, the data of SHCI each trading day from January 4, 2000, to November 30, 2009. Figure 5 presents the daily SHCI closing price in this period. In Set 2, there are 2392 selected data in which the 2171 data points from 2000 to 2008 are conducted as training set while the remaining 221 data points are conducted as testing set.

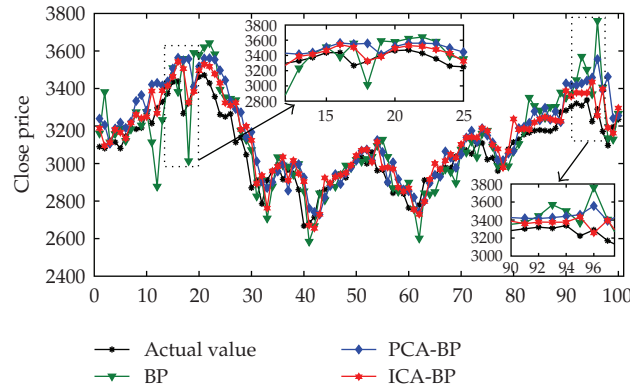


Figure 6: The actual daily SHCI closing price and the predicted values of the three models in Set 1.

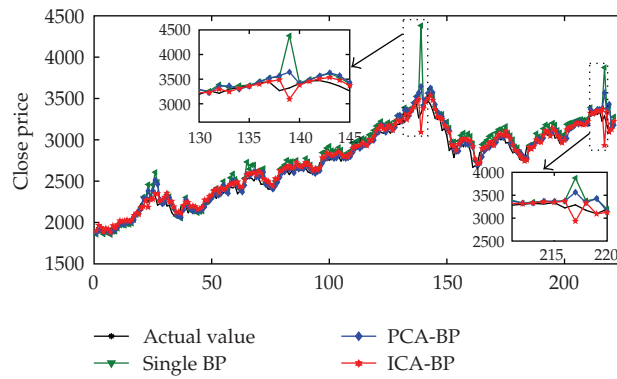


Figure 7: The actual daily SHCI closing price and the predicted values of the three models in Set 2.

4.2. Performance Criteria and Basic Setting of Model

The prediction performance is evaluated by using the following performance measures: the mean absolute error (MAE), the root mean square error (RMSE), and the correlation coefficient (R). The corresponding definitions are given as follows:

$$\begin{aligned} \text{MAE} &= \frac{1}{n} \sum_{i=1}^n |A_i - P_i|, & \text{RMSE} &= \sqrt{\frac{1}{n} \sum_{i=1}^n (A_i - P_i)^2}, \\ R &= \frac{\sum_{i=1}^n (A_i - \bar{A})(P_i - \bar{P})}{\sqrt{\sum_{i=1}^n (A_i - \bar{A})^2 \sum_{i=1}^n (P_i - \bar{P})^2}}, \end{aligned} \quad (4.1)$$

where A is the actual value, P is the predicted value, \bar{A} is the mean of the actual value, \bar{P} is the mean of the predicted value, and n is the total number of the data. The smaller

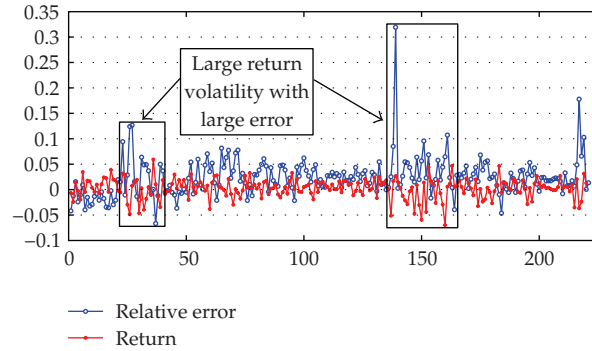


Figure 8: The daily SHCI return and the relative error of forecasting result from the single BP model.

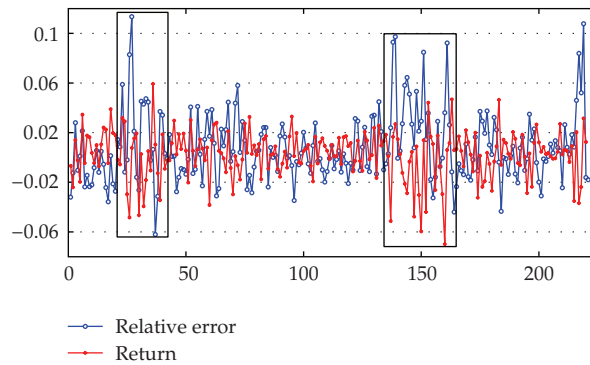


Figure 9: The daily SHCI return and the relative error of forecasting result from the common PCA-BP model.

MAE and RMSE values and the larger R value represent the less deviation, that is, the better performance of the forecasting model.

To compare the forecasting performance of the single BP model, the common PCA-BP model, and the proposed ICA-BP model, all three models contain BP neural network, so we set for the BP neural network the similar architecture and the same parameters. This can show the effect that the PCA and the ICA method process the input data of BP neural network. For the BP neural network, we only set one hidden layer. The number of neural nodes in the input layer is N (it is different for the 3 models), the number of neural nodes in the hidden layer is set to be $2N + 1$ according to the empirical formula (see [23]), and the number of neural nodes in the output layer is 1. We can use $N - (2N + 1) - 1$ to represent the architecture of the network. The threshold of the maximum training cycles is 1000, the threshold of the minimum error is 0.0001, the activation function is $f(x) = (2/1 - e^{-2x}) - 1$, and the learning rate is 0.1. In the single BP model, the number of neural nodes in the input layer is 6 which corresponds to daily open price, daily closing price, daily highest price, daily lowest price, daily volume, and daily turnover, the number of the neural nodes in the hidden layer is 13, the number of neural nodes in the output is 1 which corresponds to closing price of the next trading day, and the architecture is 6-13-1. In the common PCA-BP model, after analyzing the six original time series by using PCA method, we obtain two PCs (see Section 3.1), and the number of neural nodes in the input layer is 2 which corresponds to the

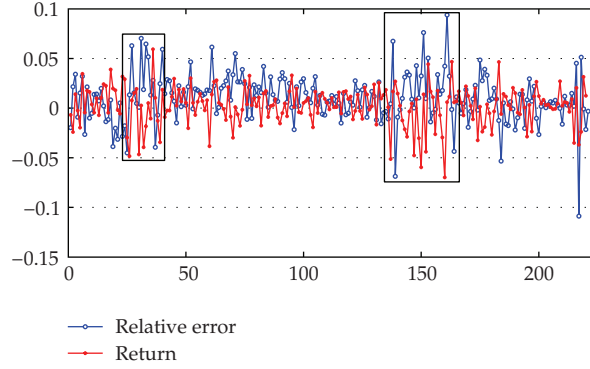


Figure 10: The daily SHCI return and the relative error of forecasting result from the ICA-BP model.

Table 4: Forecasting result of the daily SHCI closing price with three models using Set 1.

	MAE	RMSE	R
Single BP	95.8038	130.2011	0.8723
PCA-BP	84.3123	119.5324	0.9206
ICA-BP	68.5315	90.3209	0.9334

two PCs, the number of neural nodes in the hidden layer is 5, the output layer is the same as the single BP model, and the architecture is 2-5-1. In the proposed ICA-BP model, after analyzing the six time series by using ICA method and eliminating one IC that represents the noise, we obtain five ICs (see Section 3.2), and the number of neural nodes in the input layer is 5 which corresponds to the five ICs, the number of neural nodes in the hidden layer is 11, the output layer is also the same as the single BP model, and the architecture is 5-11-1. The architectures of all the three models are $N - (2N + 1) - 1$.

4.3. The Comparisons of Forecasting Results

For comparing the forecasting performance of the proposed ICA-BP model with the common PCA-BP model and the single BP model, the two sets of data (Set 1 and Set 2) are, respectively, used for the empirical study.

(I) Firstly, Table 4 depicts the forecasting result of daily SHCI closing price with the three forecasting models by using Set 1 data. It can be observed that the MAE is 68.5315, the RMSE is 90.3209, and the R is 0.9334 in the proposed ICA-BP model. The MAE and the RMSE are smaller and the R is larger than those of the other two models. We can summarize that the proposed ICA-BP model outperforms the other two models and the common PCA-BP model outperforms the single BP model. From Figure 6, we can conclude the same result.

Table 5 and Figure 7 both show the forecasting result of daily SHCI closing price with the three forecasting models by using Set 2 data. The conclusion is similar to that of Set 1, that is, the proposed ICA-BP model has the best performance and the common PCA-BP outperforms the single BP model.

(II) Comparing Table 4 with Table 5, we can also see that the proposed ICA-BP model outperforms the common PCA-BP model distinctly in Set 1. The MAE values are 68.5315 and 84.3123, respectively, where the difference of the two numbers is about 16, and

Table 5: Forecasting result of the daily SHCI closing price with the three models using Set 2.

	MAE	RMSE	R
Single BP	86.0270	129.4198	0.9746
PCA-BP	56.4246	80.9682	0.9848
ICA-BP	51.5165	70.8551	0.9877

the RMSE values are 90.3209 and 119.5324, respectively, where the difference of two numbers is about 29. Nevertheless in Set 2, the MAE values are 51.5165 and 56.4246, respectively, where the difference of the two numbers is about 5, and the RMSE values are 70.8551 and 80.9682, respectively, where the difference of the two numbers is about 10. It shows that the performance of the common PCA-BP model becomes closer to the proposed ICA-BP. It means that the denoising ability of ICA method is clearly better than that of the PCA method in relatively small samples, but the denoising ability of PCA method is closer to that of the ICA method in relatively large samples. This may be because the PCA method is based on Gaussian assumption and the ICA method is based on non-Gaussian assumption. In the case of small sample, the corresponding distribution usually deviates from the Gaussian distribution.

(III) In this part, we consider the statistical behaviors of the price returns in Shanghai stock market and the relative error of the forecasting result in Set 2. The formula of stock logarithmic return and relative error is given as follows:

$$r(t) = \ln A(t+1) - \ln A(t), \quad e(t) = \frac{[A(t) - P(t)]}{A(t)}, \quad (4.2)$$

where $A(t)$ and $P(t)$, respectively, denote the actual value and the predicted value of daily closing price of SHCI at the date t , $t = 1, 2, \dots$. In Figure 8, we consider the fluctuation of the daily SHCI return and the relative error of forecasting result from the single BP model. Similarly, Figure 9 is the plot for the daily SHCI return and the relative error of forecasting result from the common PCA-BP model; Figure 10 is the plot for the daily SHCI return and the relative error of forecasting result from the proposed ICA-BP model. From Figures 8–10, it can be seen that there are all some points with large relative error of forecasting result in the three models. Through the observation, we can notice that these points appear basically in the place where there is large return volatility (marked in Figures 8–10). This indicates that the predicted results to the points where prices fluctuate violently are relatively not satisfactory by using the three models. The marked parts in Figures 6 and 7 can also support this opinion.

5. Conclusion

In the present paper, we investigate and forecast the fluctuations of Shanghai stock market. The independent component analysis method and the principal component analysis method are introduced into the neural network model to forecast the stock price. In the proposed ICA-BP model, the input data is firstly analyzed by using ICA, and we obtain several latent ICs; after identifying and eliminating, the IC represents the noise and the remaining ICs are conducted as the input of the BP model. Further, the empirical research is made to compare the actual daily SHCI closing price with the predicted values of the three models,

and the relative errors of forecasting results are analyzed. Empirical results show that the ICA-BP model outperforms the other two models.

Acknowledgment

The authors are supported in part by the National Natural Science Foundation of China Grant no. 70771006 and no. 10971010.

References

- [1] R. Gamberini, F. Lolli, B. Rimini, and F. Sgarbossa, "Forecasting of sporadic demand patters with seasonality and trend components: a empirical comparison between Holt-Winters and (S) ARIMA methods," *Mathematical Problems in Engineering*, vol. 2010, Article ID 579010, 14 pages, 2010.
- [2] R. Gaylord and P. Wellin, *Computer Simulations with Mathematica: Explorations in the Physical, Biological and Social Science*, Springer, New York, NY, USA, 1995.
- [3] K. Ilinski, *Physics of Finance: Gauge Modeling in Non-equilibrium Pricing*, John Wiley, New York, NY, USA, 2001.
- [4] 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. 29–33, 2007.
- [5] Q. D. Li and J. Wang, "Statistical properties of waiting times and returns in Chinese stock markets," *WSEAS Transactions on Business and Economics*, vol. 3, pp. 758–765, 2006.
- [6] T. C. Mills, *The Econometric Modelling of Financial Time Series*, Cambridge University Press, Cambridge, UK, 2nd edition, 1999.
- [7] D. Enke and S. Thawornwong, "The use of data mining and neural networks for forecasting stock market returns," *Expert Systems with Applications*, vol. 29, no. 4, pp. 927–940, 2005.
- [8] A. Kanas and A. Yannopoulos, "Comparing linear and nonlinear forecasts for stock returns," *International Review of Economics and Finance*, vol. 10, no. 4, pp. 383–398, 2001.
- [9] H. F. Zou, G. P. Xia, F. T. Yang, and H. Y. Wang, "An investigation and comparison of artificial neural network and time series models for Chinese food grain price forecasting," *Neurocomputing*, vol. 70, no. 16–18, pp. 2913–2923, 2007.
- [10] T. Bollerslev, "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, vol. 31, no. 3, pp. 307–327, 1986.
- [11] E. D. McKenzie, "General exponential smoothing and the equivalent ARMA process," *Journal of Forecasting*, vol. 3, pp. 333–344, 1984.
- [12] F. Leon and M. H. Zaharia, "Stacked heterogeneous neural networks for time series forecasting," *Mathematical Problems in Engineering*, vol. 2010, Article ID 373648, 20 pages, 2010.
- [13] 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.
- [14] B. Y. Lu, Y. L. Chen, and Y. Y. Li, "The forecast of the pre-processing data with BP neural network and principal component analysis," *Science & Technology Information*, vol. 17, pp. 29–30, 2009.
- [15] D. Olson and C. Mossman, "Neural network forecasts of Canadian stock returns using accounting ratios," *International Journal of Forecasting*, vol. 19, no. 3, pp. 453–465, 2003.
- [16] V. D. A. Sánchez, "Frontiers of research in BSS/ICA," *Neurocomputing*, vol. 49, pp. 7–23, 2002.
- [17] A. Back and A. Weigend, "Discovering structure in finance using independent component analysis," in *Proceedings of 5th International Conference on Neural Networks in Capital Market*, pp. 15–17, Kluwer Academic Publishers, 1997.
- [18] A. Hyvarinen, *Independent Component Analysis*, Mechanic Industry Press, 2007.
- [19] Z. Q. Yang, Y. Li, and D. W. Hu, "Independent component analysis: a survey," *Acta Automatica Sinica*, vol. 28, no. 5, pp. 762–772, 2002.
- [20] C. F. Beckmann and S. M. Smith, "Probabilistic independent component analysis for functional magnetic resonance imaging," *IEEE Transactions on Medical Imaging*, vol. 23, no. 2, pp. 137–152, 2004.
- [21] C. J. James and O. J. Gibson, "Temporally constrained ICA: An application to artifact rejection in electromagnetic brain signal analysis," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 9, pp. 1108–1116, 2003.

- [22] K. Kiviluoto and E. Oja, "Independent component analysis for parallel financial time series," in *Proceeding of the 5th International Conference on Neural Information*, pp. 895–898, 1998.
- [23] L. Q. Han, *Theory, Design and Application of Artificial Neural Network*, Chemical Industry Press, 2002.
- [24] E. M. Azoff, *Neural Network Time Series Forecasting of Financial Market*, Wiley, New York, NY, USA, 1994.
- [25] Y. Ouyang, "Evaluation of river water quality monitoring stations by principal component analysis," *Water Research*, vol. 39, pp. 2621–2635, 2005.
- [26] L.-N. Yang, L. Peng, L.-M. Zhang, L.-I. Zhang, and S.-S. Yang, "A prediction model for population occurrence of paddy stem borer (*Scirpophaga incertulas*), based on Back Propagation Artificial Neural Network and Principal Components Analysis," *Computers and Electronics in Agriculture*, vol. 68, no. 2, pp. 200–206, 2009.
- [27] Y.-M. Cheung and L. Xu, "Independent component ordering in ICA time series analysis," *Neurocomputing*, vol. 41, pp. 145–152, 2001.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

