*Algorithms* **2014**, 7, 650-662; doi:10.3390/a7040650

OPEN ACCESS

*algorithms*

**ISSN 1999-4893**

www.mdpi.com/journal/algorithms

*Article*

# Time Series Prediction Method of Bank Cash Flow and Simulation Comparison

**Wen-Hua Cui, Jie-Sheng Wang * and Chen-Xu Ning**

School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan 114044, Liaoning, China; E-Mails: cwh@julong.cc (W.-H.C.); NCX710754167@163.com (C.-X.N.)

* Author to whom correspondence should be addressed; E-Mails: wang_jiesheng@126.com or wangjiesheng@ustl.edu.cn; Tel.: +0412-253-8355; Fax: +0412-253-8244.

External Editors: Kenji Suzuki

*Received: 25 September 2014; in revised form: 5 November 2014 / Accepted: 14 November 2014 / Published: 26 November 2014*

**Abstract:** In order to improve the accuracy of all kinds of information in the cash business and enhance the linkage between cash inventory forecasting and cash management information in the commercial bank, the first moving average prediction method, the second moving average prediction method, the first exponential smoothing prediction and the second exponential smoothing prediction methods are adopted to realize the time series prediction of bank cash flow, respectively. The prediction accuracy of the cash flow time series is improved by optimizing the algorithm parameters. The simulation experiments are carried out on the reality commercial bank's cash flow data and the predictive performance comparison results show the effectiveness of the proposed methods.

**Keywords:** time series prediction; moving average prediction; exponential smoothing prediction

## 1. Introduction

In recent years, the time series modeling and prediction have been one of the most active research topics in the academic research and engineering practice [1–4]. The time series is usually a chronological series of observed data (information) according to the time sequence, whose values are sampled at the invariable time intervals. Researchers often predict future changes based on the historical data. For example, according to the situation in the past or the current period of the market sales, the changes of stock prices, the population growth and the bank's deposit and withdrawal, the changes of the market sales, the changes of stock prices, the population growth and the bank's deposit and withdrawal in the future are predicted. The time series forecasting affects the life of people everywhere, so it has an important practical significance and research prospects in every field of today's society, which is also an important direction in the computer application field [5–8].

The bank cash flow forecasting management information system is designed to create a system management platform for the prediction and analysis of the commercial bank cash flow. It will realize the cash flow data statistics summary, the cash flow short-term and long-term prediction, the summary and statistical analysis comprehensively and scientifically of the business information, the operational information and the management information related to the commercial bank cash flow under three levels: secondary branches (Cash Operation Center), branch (Business Library) and Network. Its purpose is to provide effective data all levels of organization to analyze and assess cash business operation conditions. Also it will provide effective system management means for the cash operation managers and decision-making people at all levels.

Aiming at the existed problem in the analysis of commercial bank cash flow, four time series prediction methods are used to set up the prediction models. The simulation results show the effectiveness of the proposed methods. The paper is organized as follows. In Section 2, the time series prediction methods of bank cash flow are introduced. The simulation experiments and results analysis are introduced in details in Section 3. Finally, the conclusion illustrates the last part.

## 2. Time Series Prediction Methods of Bank Cash Flow

The time series prediction method analyzes the predicted target's changes with the time on the times series composed of the historical data according to the chronological order. The prediction method is quantitative and the related mathematical model is established for extrapolation [9–15]. Based on the circulation data of bank cash, four prediction methods of the various cash operating specific modules are designed for providing the decision-making basis of the business plan in commercial bank. The flowchart of the time series prediction methods of bank cash flow is shown in Figure 1. The simulation environment is described as follows: *Windows* 7 operating system, *Intel* processor (2.5 GHz, 4G memory) and Matlab 2010 simulation software.

### *2.1. The First Moving Average Predictive Method*

The moving average method is a predictive technique developed on the basis of the arithmetic average. The arithmetic average method can only reflect the average of a set of data and cannot reflect the change trend of the data. However, the basic idea of the moving average method is item by item based on time series data by taking a certain number of cycles of data to be averaged every time, successive advancing in chronological order. In each propulsion cycle, the forward cycle data is rejected, a new cycle data is added and the average value is calculated [16]. Set $X_t$ is the actual value of the cycle $t$. The first moving average value is calculated by the following equation.

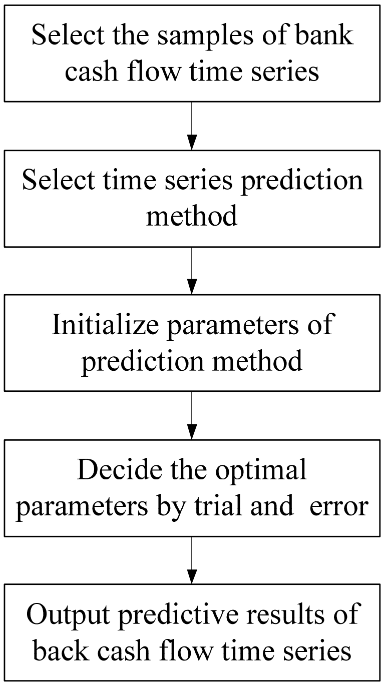
$$M_t^{(1)}(N) = \frac{X_t + X_{t-1} + \ldots + X_{t-N+1}}{N} = \sum_{i=0}^{N-1} X_{t-i} / N \tag{1}$$

So the predictive value of the next period ($t + 1$) is:

$$\hat{X}_{t+1} = M_t^{(1)} \tag{2}$$

where N is the number of selected data for calculating the first moving average value and $\hat{X}_{t+1}$ is the predictive value of the next period ($t + 1$).

**Figure 1.** Flowchart of the time series prediction methods of bank cash flow.




Based on the collected data in a commercial bank from January to April in 2011and 2012, the first moving average method is adopted to realize the time series prediction of bank cash flow. The cycle number of the moving average in 2011 is 2–15 and the continuous sample points are 100. The optimal moving average value for the selected data is N = 6 and the simulation results are shown in Figure 2. The cycle number of the moving average in 2012 is 2–15 and the continuous sample points are 100. The optimal moving average value for the selected data is N = 2 and the simulation results are shown in Figure 3.

It can be seen from Figures 2 and 3 that the random fluctuations of bank actual cash flow values is larger, but the random fluctuations are reduced significantly after the first moving average calculations. The more used months by calculating the average, that is to say the greater N, the stronger of the smoothing degree and the more small fluctuation. However, in this case, the reaction velocity on this change trend for the actual data is slower. Conversely, if N is set lower, the reaction velocity on this change trend for the actual data is more sensitive, but the smoothing degree is less and it is easy to reflect the random interference as the trend. Therefore, the choice of N is very important, which should be selected according to the specific situation.

When N is equal to the cycle of change, the influence of periodic change can be eliminated. In practice, the mean square error S of prediction for past data is generally adopted to be the criterion for choosing N. However, because of the accumulation of errors, the predictive error is bigger for the prediction of more distant periods. Therefore, the first moving average method is usually adapted to the stationary model and only a period prediction (*i.e.*, predicting the $t + 1$ period). So when the basic mode of predicted variables changes, the adaptability of the first moving average method will be relatively poor.

**Figure 2.** Predictive results of the first moving average method in 2011.

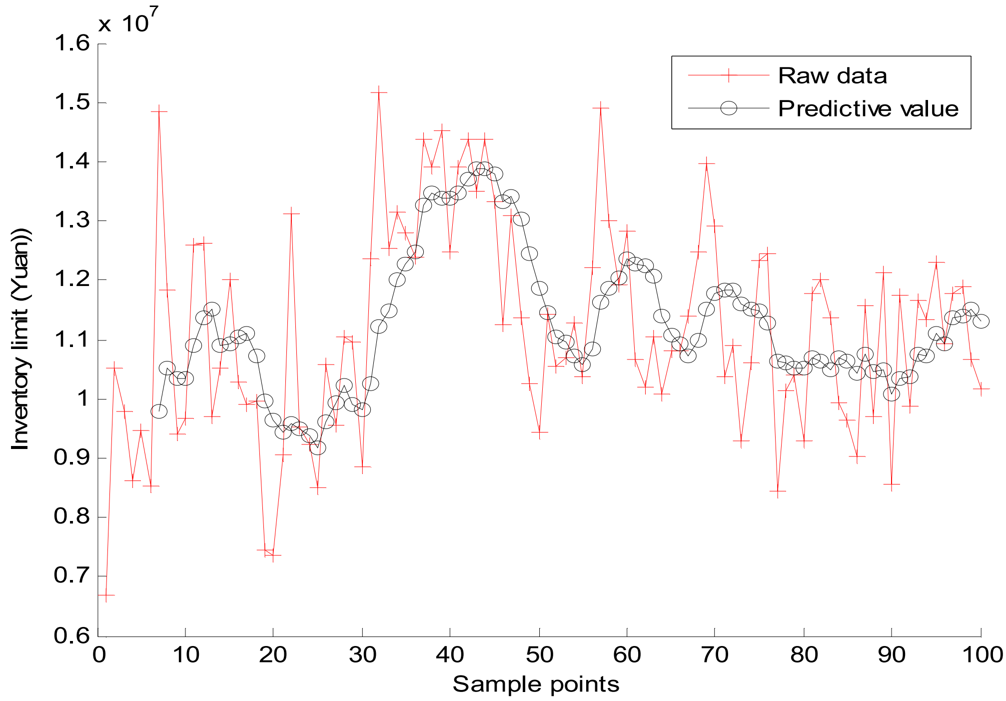
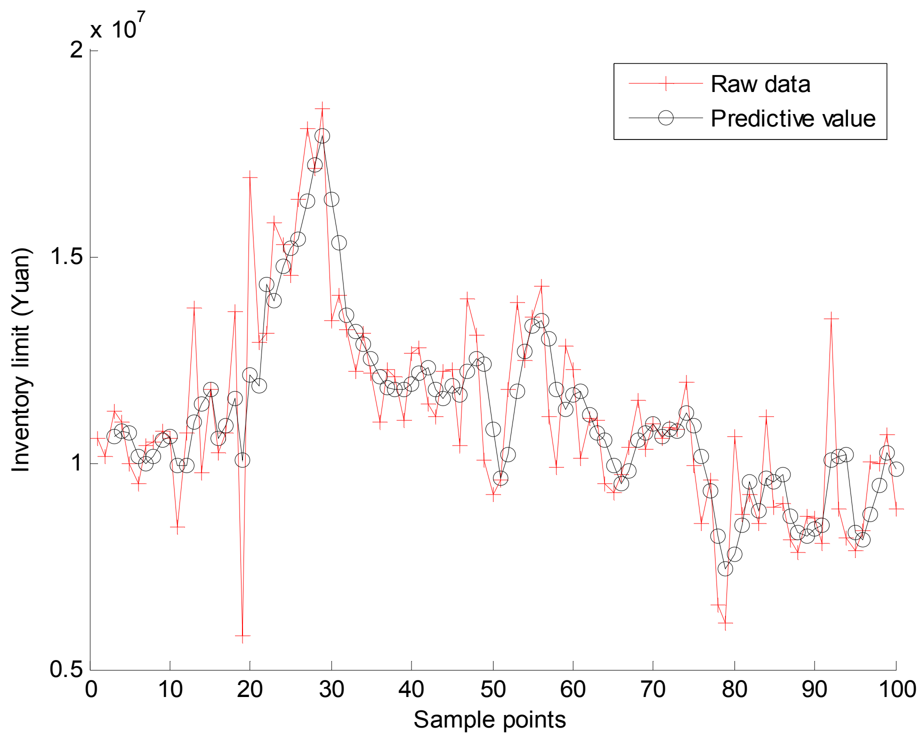



**Figure 3.** Predictive results of the first moving average method in 2012.



*2.2. The Second Moving Average Predictive Method*

The second moving average method is to carry out the moving average again based on the first moving average method, that is to say the prediction model is established on the basis of the first moving average values and the second moving average values to calculate the predicted value. As mentioned above, the average value calculated by the first moving average method has a lag deviation. Especially when there is a linear trend in the time series data, the first moving average value is always behind the observation value. The second moving average method is proposed to revise this lag deviation by establishing the mathematical model having the linear time relationship of the predictive target to obtain the predictive values.

The second moving average predictive method can solve the contradiction of the predicted value lagging behind the actual observed values, which are suitable for forecasting the time series with the phenomenon having the obvious change trend. At the same time, it also retains the advantage of the first moving average method. Set $M_t^{(2)}$ is the second moving average value in period $t$, $Y_t$ is the actual value of the cycle $t$. The second moving average method is described as follows.

$$M_t^{(1)} = \frac{Y_t + Y_{t-1} + \ldots + Y_{t-n+1}}{n} \tag{3}$$

$$M_t^{(2)} = \frac{M_t^{(1)} + M_{t-1}^{(1)} + \ldots M_{t-n+1}^{(1)}}{n} \tag{4}$$

where $M_t^{(1)}$ is the first moving average value in period $t$, $M_t^{(2)}$ is the second moving average value in period $t$, $n$ is the number of selected data for calculating the first moving average value.

In order to eliminate the influence of lag error on the prediction results, the linear trend model with the hysteresis error rule on the basis of the first and second moving average values is established. $M_t^{(1)}$ and $M_t^{(2)}$ are used to estimate the intercept $\hat{a}_t$ and slope $\hat{b}_t$ of the linear trend model. The calculation formulas are described as follows.

$$\begin{cases} \hat{a}_t = 2M_t^{(1)} - M_t^{(2)} \\ \hat{b}_t = \dfrac{2}{N-1}(M_t^{(1)} - M_t^{(2)}) \end{cases} \tag{5}$$

Then the linear trend prediction model is established as:

$$\hat{y}_{t+\tau} = \hat{a}_t + \hat{b}_t \tau \tag{6}$$

where $t$ is the current period, $\tau$ is from $t$ to forecast period number, $\hat{y}_{t+\tau}$ is predictive value in period $t + \tau$, $\hat{a}_t$ is the estimated value of the intercept, $\hat{b}_t$ is the estimated value of the slope.

In order to carry out the research on the second moving average method for predicting the bank cash flow time series, the same data in the first moving average method are adopted. The cycle number of the moving average in 2011 and 2012 is 2–15 and the optimal moving average value for the selected data is N = 3. The simulation results are shown in Figures 4 and 5. It can be seen from Figures 4 and 5 that the predicted effect of the second moving average method is better than the first moving average method. So it is more suitable for the prediction the time series with linear trend changes.

**Figure 4.** Predictive results of the second moving average method in 2011.

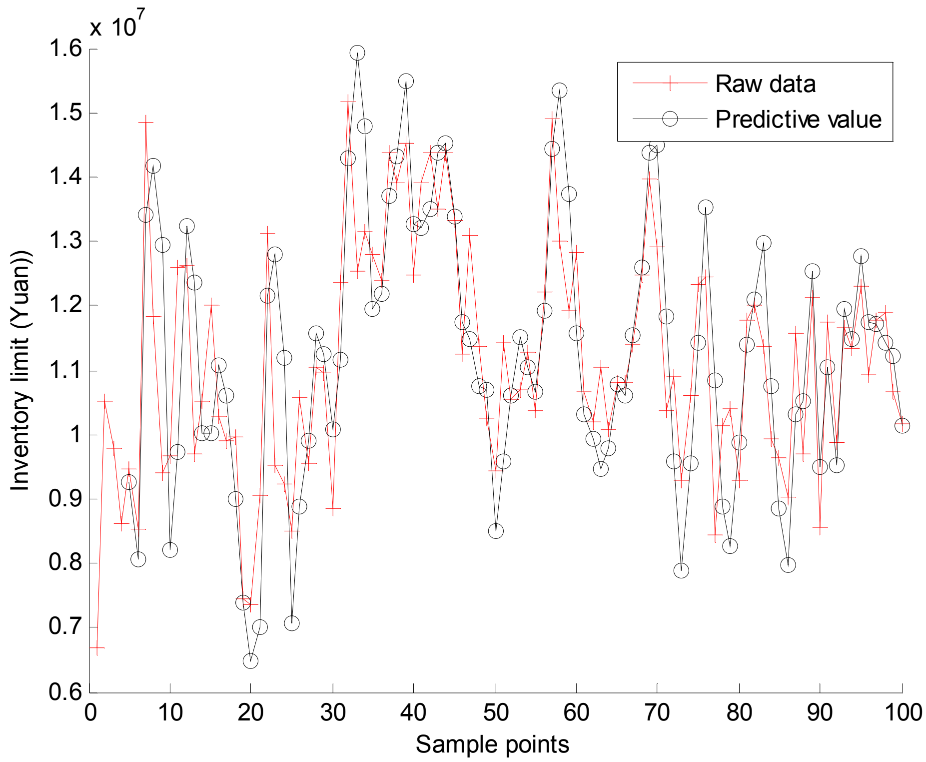
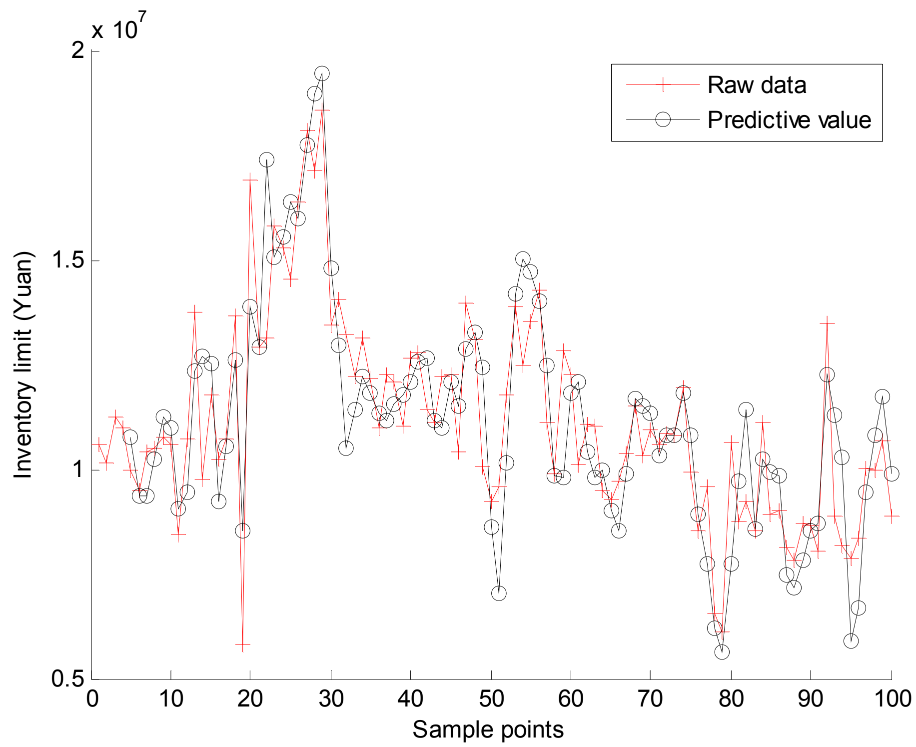



**Figure 5.** Predictive results of the second moving average method in 2012.




### *2.3. The First Exponential Smoothing Prediction Method*

The exponential smoothing method is a time sequence analysis prediction method developed from the moving average method. By calculating the exponential smoothing value, the future of phenomenon is predicted with the certain time series prediction model. Its principle is to calculate the exponential smoothing values in any period with the style of weighted average by combining the actual observed value and the exponential smoothing value in the previous period. The first exponential smoothing method is a weighted prediction, whose weight coefficient is α. It need not store all history data, which can greatly reduce the data storage. Sometimes, only a new observation value, the latest predicted value and the weight coefficient α can be used to realize the time series prediction.

When the time series has no obvious trend change, the first exponential smoothing prediction method is effective. Set $X_0, X_1, \ldots, X_n$ are the observations of time series and $S_1^{(1)}, S_2^{(1)}, \ldots, S_n^{(1)}$ are the exponential smoothing values of observations in time *t*. So the first exponential smoothing value is calculated by the following equation.

$$S_t^{(1)} = \alpha X_t + \alpha(1-\alpha)X_{t-1} + \alpha(1-\alpha)^2 X_{t-2} + \ldots \tag{7}$$

where α is smoothing coefficient, $0 < \alpha < 1$.

It can be seen form Equation (7) that the weight coefficients of the actual values $X_t$, $X_{t-1}$ and $X_{t-2}$ are α, $\alpha\cdot(1-\alpha)$ and $\alpha\cdot(1-\alpha)^2$, respectively. The data is farther from the moment, with the smaller weight coefficient. Because the weight coefficient is an index geometric series, it is called the exponential smoothing method. Equation (7) can be changed slightly as:

$$\begin{aligned} S_t^{(1)} &= \alpha X_t + (1-\alpha)\left[\alpha X_{t-1} + \alpha(1-\alpha)X_{t-2} + \ldots\right] \\ &= \alpha X_t + (1-\alpha)S_{t-1}^{(1)} \end{aligned} \tag{8}$$

Equation (8) can be rewritten as:

$$S_t^{(1)} = S_{t-1}^{(1)} + \alpha(X_t - S_{t-1}^{(1)}) \tag{9}$$

Then the predictive value of next cycle is obtained:

$$\hat{X}_{t+1} = S_t^{(1)} \tag{10}$$

$$\hat{X}_{t+1} = \hat{X}_t + \alpha(X_t - \hat{X}_t) \tag{11}$$

Seen from the Equation (11), the exponential smoothing method solves a problem existing in the moving average method, *i.e.*, threre is no longer a need to store the historical data in past N cycles. One only needs the most recent observed value $x_t$, the most recent predicted value $\hat{X}_t$ and the weight coefficient α to calculate a new predictive value.

The exponential smoothing method is used to realize the unequal weights data handle with different times by using the smoothing coefficient α. So the problem of chosing α is discuss as follows. (1) $0 < \alpha < 1$. (2) The selection of the smoothing coefficient α is an important problem, which affects the prediction results directly. It is selected according to the characteristics of the actual time series and expert experiences.

(1) When the fluctuation of time series is not big and the long-term trend is relatively stable, it is better to select a smaller α. The weight of the past predictive values can then be aggravated. Its value generally locates the scope [0.05, 0.2].

(2) When the fluctuation of time series is big and has an obvious tendency of rapid change, it is better to select the big α, which can aggravate the weight of the new predictive values. Its value generally locates the scope [0.3, 0.7].

In order to carry out the research on first exponential smoothing prediction method for predicting the bank cash flow time series, the same data in the above two methods are adopted. When the weight coefficient is 0.9, the mean square error is minimum and the simulation results are shown in Figures 6 and 7. It can be seen from Figures 6 and 7 that the first exponential smoothing model prediction method need to find the best value α through the repeated experiments. However, even with a best value α, the accuracy of predicted results also is not ideal for the time series with larger trend.

**Figure 6.** Predictive results of the first exponential smoothing prediction method in 2011.

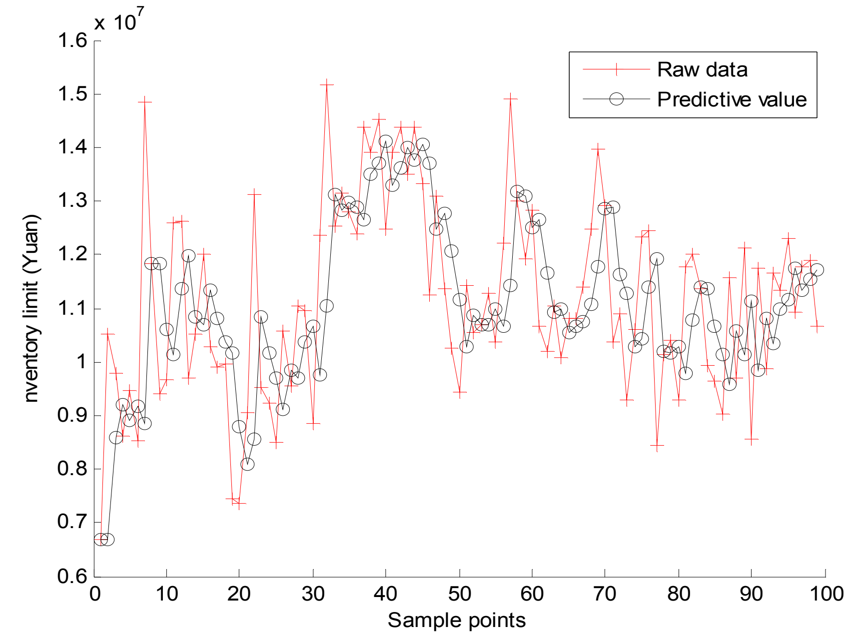
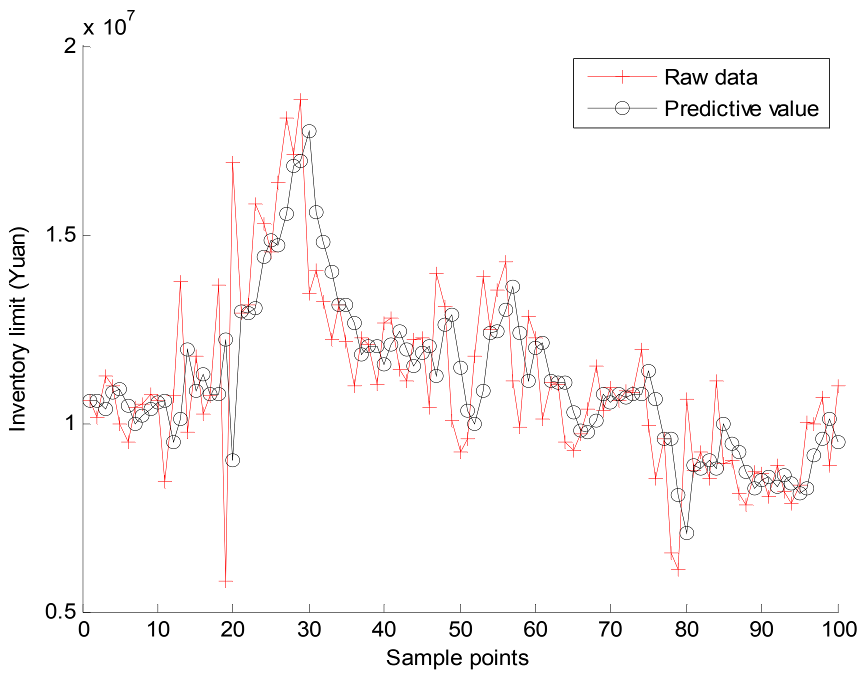



**Figure 7.** Predictive results of the first exponential smoothing prediction method in 2012.



*2.4. The Second Exponential Smoothing Prediction Method*

Because the first moving average method and second linear moving average method all need store a large amount of historical data, in order to compensate for this limitation, the second linear exponential smoothing model is developed, which only need to use three data values and a smoothing coefficient α. At the same time, it also can make the weight of past observations small. In most cases, the use of second linear exponential smoothing model is more convenient.

The second exponential smoothing prediction method is to carry out another exponential smoothing based on the first exponential smoothing prediction method, which has the advantages of simple calculation, less sample requirement, strong adaptability and the relatively stable predicted results. It cannot predict lonely and must cooperate with the first exponential smoothing prediction method to build the time series predictive mathematical model. The weighted average of historical data is as the predicted value of future time. The second exponential smoothing prediction method is described as follows.

$$S_t^{(1)} = aY_t(1-\mathrm{a})\mathrm{S}_{(\mathrm{t}-1)}^{(1)} \tag{12}$$

$$S_t^{(2)} = aS_t^{(1)} + (1-\mathrm{a})\mathrm{S}_{(\mathrm{t}-1)}^{(2)} \tag{13}$$

where, $S_t^{(1)}$ is the first exponential smoothing value in period $t$, $S_t^{(2)}$ is the second exponential smoothing value in period $t$ and α is the smoothing constant. So the prediction model of the second exponential smoothing method is described as follows.

$$F_{t+T} = a_t + b_t T \tag{14}$$

$$a_t = 2S_t^{(1)} - S_t^{(2)} \tag{15}$$

$$b_t = (a/1-\mathrm{a})(\mathrm{S}_t^{(1)} - \mathrm{S}_t^{(2)}) \tag{16}$$

where $F_{t+T}$ is the predictive value in period $t + T$, $T$ is the number of the future forecast period, $a_t$ and $b_t$ are the model parameters.

Equation (15) is used to add the difference between the first exponential smoothing value and the second exponential smoothing value on the first exponential smoothing value. Equation (16) is adopted to add the trend change value. When predicting the period $t + 1$, a trend change value $b_t$ is added to the $a_t$. When predicting the period $t + T$, the trend change value $T \times b_t$ is added to the $a_t$.

In order to carry out the research on the second exponential smoothing prediction method for predicting the bank cash flow time series, the same data in the above three methods are adopted. When the weight coefficient is 0.5, the mean square error is minimum and the simulation results are shown in Figures 8 and 9.

It can be seen from Figures 8 and 9 that because the second exponential smoothing method considers the linear parameter changes in different periods of the time series, the extent of the predicted values fitting with the original time series is very good, which reflects the variation trends of the original time series in different time periods. So it is more suitable for the prediction of time series with a linear trend.

**Figure 8.** Predictive results of the second exponential smoothing method in 2011.

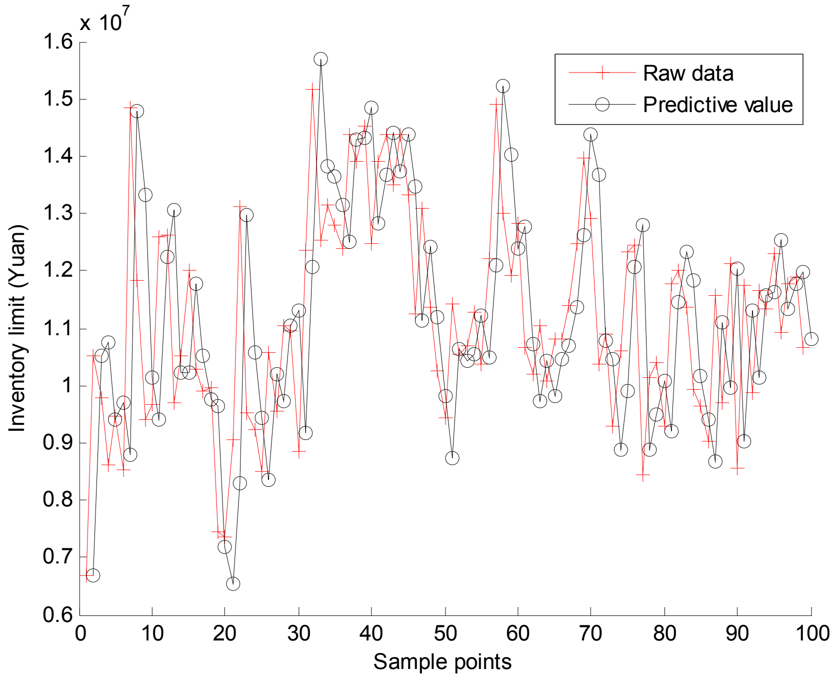
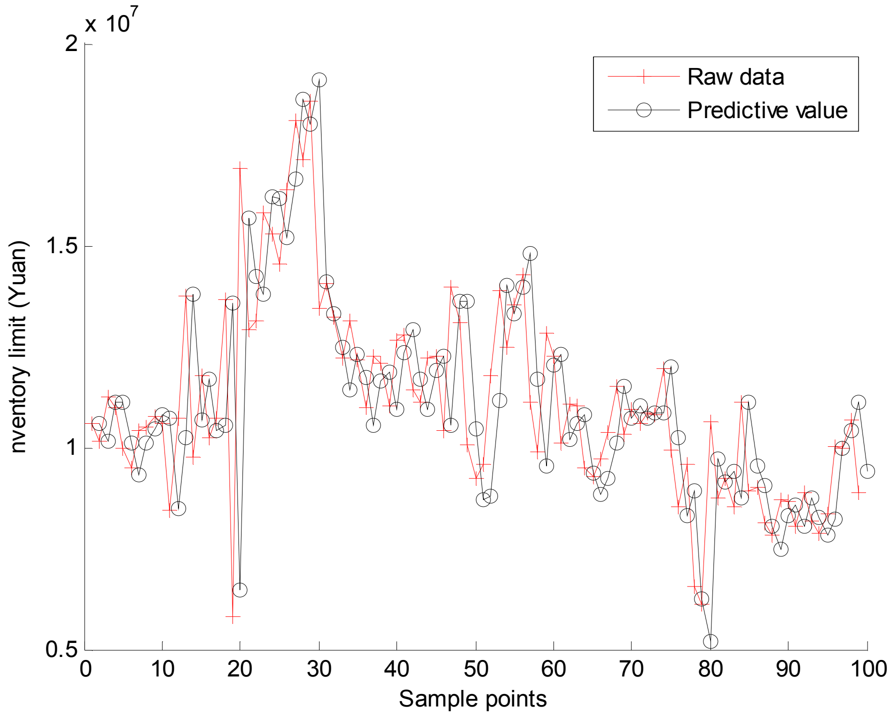


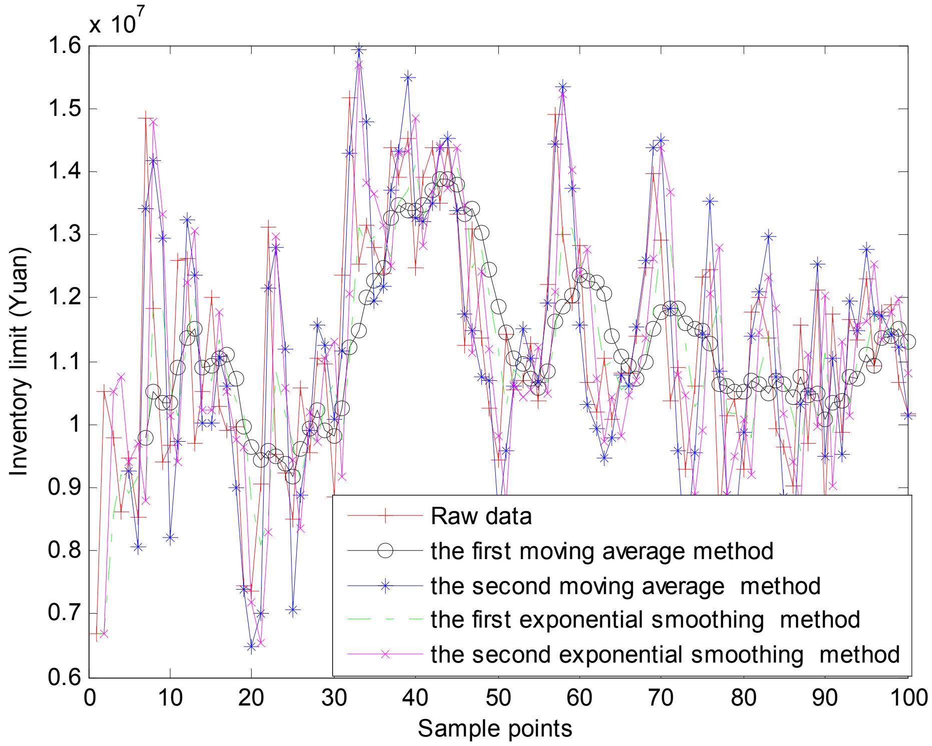

**Figure 9.** Predictive results of the second exponential smoothing method in 2012.




## 3. Simulation Comparison

In order to carry out the performance comparison under the above mentioned four time series predictive methods for bank cash flow, the simulation curves are shown in Figure 10. It can be seen clearly that the predicted results of the second exponential smoothing method is better than other methods.

**Figure 10.** Performance comparison results.




The prediction error is the deviation between the predicted results and the actual results, which determines the prediction accuracy. The accuracy of the quantitative prediction methods has a lot of measurable indicators. Suppose $y_1, y_2, \ldots, y_n$ are the actual observed values of the predicted target and $\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_n$ are the predicted values.

The absolute error of predicted points is

$$a_t = y_t - \hat{y}_t, \quad t = 1, 2, \cdots, n \tag{17}$$

where $a_t$ is the absolute error at the point $t$. Obviously, $a_t$ is the most direct measure index of prediction error, but it is affected by the measurement unit of predicted object. So it is unsuitable as the final measure indicator of prediction accuracy.

Relative error of predicted points

$$\hat{a}_t = \frac{a_t}{y_t} = \frac{y_t - \hat{y}_t}{y_t}, \quad t = 1, 2, \cdots, n \tag{18}$$

where $\hat{a}_t$ is the relative error at the point $t$, which is usually expressed as a percentage and measures the accuracy of the predicted values relative to the observed value at the predicted point $t$.

Prediction accuracy of the prediction points

$$A_t = 1 - |y_t - \hat{y}_t| / y_t\,, \; 0 \le |y_t - \hat{y}_t / y_t \le 1 \tag{19}$$

$$A_t = 0\,, \; |y_t - \hat{y}_t| / y_t > 1 \tag{20}$$

where $A_t$ is the prediction accuracy at the prediction point $t$.

Mean square error (MSE)

Mean square error (MSE) is a kind of convenient method to measure the average error to evaluate the degree of data change, which is described as follows.

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(\mathrm{y}_t - \hat{y}_t)^2 \tag{21}$$

Four mentioned methods are used to carry out the simulation on the same pretreated experimental. The simulation results are shown in Table 1. The performance comparison results in Table 1 also showed that the second exponential smoothing method obtains the higher prediction precision of the cash flow time series.

**Table 1.** Performance comparison results.

| Performance indicators | First moving average method | Second moving average method | First exponential smoothing method | Second exponential smoothing method |
|---|---|---|---|---|
| MSE | $2.47 \times 10^{12}$ | $1.55 \times 10^{12}$ | $2.67 \times 10^{12}$ | $1.02 \times 10^{12}$ |
| Absolute error | $1.14 \times 10^{6}$ | $6.23 \times 10^{5}$ | $1.06 \times 10^{6}$ | $1.12 \times 10^{5}$ |
| Relative error of (%) | 5.2 | 2.9 | 3.4 | 1.1 |
| Prediction accuracy (%) | 94.8 | 97.1 | 96.6 | 98.9 |

## 4. Conclusions

Four time series predictive methods are adopted to realize the real-time prediction of cash flow in the commercial bank. By comparing the time series predictive performance under four algorithms and analyzing the simulation results in-depth, the second exponential smoothing prediction method is the optimal prediction method.

## Acknowledgments

This work is partially supported by the Program for China Postdoctoral Science Foundation (Grant No. 20110491510), the Program for Liaoning Excellent Talents in University (Grant No. LR2014008) and the Project by Liaoning Provincial Natural Science Foundation of China (Grant No. 2014020177).

## Author Contributions

Wen-Hua Cui participated in the concept, design, interpretation and commented on the manuscript. A substantial amount of Jie-Sheng Wang's contribution to the draft writing and critical revision of this paper was undertaken. Chen-Xu Ning participated in the data collection, analysis and algorithm simulation.

## Conflicts of Interest

The authors declare no conflict of interest.

## References


1. Bar-Joseph, Z. Analyzing time series gene expression data. *Bioinformatics* **2004**, *20*, 2493–2503.
2. Dong, Z. Study on the time-series modeling of China's per capita GDP. *Contemp. Manag.* **2006**, *11*, 15.
3. Kasabov, N.K.; Song, Q. DENFIS: Dynamic evolving neural-fuzzy inference system and its application for time-series prediction. *IEEE Trans. Fuzzy Syst.* **2002**, *10*, 144–154.
4. Martinetz, T.M.; Berkovich, S.G.; Schulten, K.J. Neural-gas' network for vector quantization and its application to time-series prediction. *IEEE Trans. Neural Netw.* **1993**, *4*, 558–569.
5. Chen, Y.; Yang, B.; Dong, J. Time-series prediction using a local linear wavelet neural network. *Neurocomputing* **2006**, *69*, 449–465.
6. Xie, J.X.; Cheng, C.T.; Chau, K.W.; Pei, Y.Z. A hybrid adaptive time-delay neural network model for multi-step-ahead prediction of sunspot activity. *Int. J. Environ. Pollut.* **2006**, *28*, 364–381.
7. Shi, Z.; Han, M. Ridge regression learning in ESN for chaotic time series prediction. *Control Decis.* **2007**, *22*, 258.
8. Li, D.; Han, M.; Wang, J. Chaotic time series prediction based on a novel robust echo state network. *IEEE Trans. Neural Netw. Learn. Syst.* **2012**, *23*, 787–799.
9. Long, W.; Liang, X.M; Long, Z.Q; Qin, H.Y. RBF neural network time series forecasting based on hybrid evolutionary algorithm. *Control Decis.* **2012**, *8*, 029.
10. Zhang, G.P. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing* **2003**, *50*, 157–175.
11. Silva, C.G. Time series forecasting with a nonlinear model and the scatter search meta-heuristic. *Inf. Sci.* **2008**, *178*, 3288–3299.
12. Sun, P.; Marko, K. The square root Kalman filters training of recurrent neural networks. In Proceedings of IEEE International Conference on Systems, Man, and Cybernetics, San Diego, CA, USA, 11–14 October 1998.
13. Martinetz, T.M.; Berkovich, S.G.; Schulten, K.J. Neural-gas' network for vector quantization and its application to time-series prediction. *IEEE Trans. Neural Netw.* **1993**, *4*, 558–569.
14. Connor, J.T.; Martin, R.D.; Atlas, L.E. Recurrent neural networks and robust time series prediction. *IEEE Trans. Neural Netw.* **1994**, *5*, 240–254.
15. Van Gestel, T.; Suykens, J.A.K.; Baestaens, D.-E.; Lambrechts, A.; Lanckriet, G.; Vandaele, B.; de Moor, B.; Vandewalle, J. Financial time series prediction using least squares support vector machines within the evidence framework. *IEEE Trans. Neural Netw.* **2001**, *12*, 809–821.
16. Aras, C.M.; Jeffrey D.M.; Roderick K.S. System for allocation of network resources using an autoregressive integrated moving average method. U.S. Patent No. 5,884,037, 16 March 1999.