Forecasting of Financial Time Series with a Digital Filter and a Neural Network

Susumu Saito Shintaro Kanda
School of Management
Tokyo University of Science
500 Shimokiyoku Kuki Saitama,
346-8512, Japan
ssaito@ms.kuki.tus.ac.jp

***Abstract*-- The approach to predict time series without neglecting the fluctuation in a short period is tried by using a digital FIR filter and a neural network. The differential waveform of the Nikkei average closing price is filtered by the FIR band-pass filter of 101 lengths. In applying the filter to the financial time series, one problem will arise. It is a lack of the last 50 filtered data. This is solved by adding zero as the dummy data to the original time series. It is filtered into the five frequency bands. The each filtered waveform is learned and forecasted by the neural network. By inputting the data of 20 days earlier, ten days ahead is predicted. After learning the time series of each frequency band by the neural network, the predicted data for each frequency band are obtained. The predicted waveforms of each frequency band are synthesized to obtain the final prediction. The waveform can be forecasted well as compared with the conventional method. This method is verified by the Rossler chaos data.**

1 INTRODUCTION

In signal processing, the function of a filter is to remove unwanted parts of the frequency range of the signal, such as random noises, or to extract useful parts of the signal. A digital filter has excellent characteristics showing the wide flat pass band and the steep roll-off. It is possible to separate the signals of 1000 Hz and 1001 Hz. A digital filter is programmable, and the cutoff frequencies can be easily changed. The digital filter is widely used in many types of electrical equipments for signal separation, signal compression, noise reduction and etc [1].

Although the digital filter has excellent characteristics, it has not been used for the data processing of financial time series. Usually, the financial time series has been predicted by smoothing them using moving average method, where the fluctuation in a short period has been considered to be a noise. Actually, the fluctuation of short period is important for trading financial commodities such as stock, foreign exchange etc. Therefore, it should not be neglected by the smoothing. In this study, the digital filter is not used as a noise filter but it is used in order to separate the time series data into several frequency bands.

Many papers such as [2][3] have reported forecasting methods using a neural network for financial time series. Simulations of financial time series have been reported such as [4][5] as well. These are based on the assumption that the time series is chaos (deterministic) data with noises. The neural network is suitable for this purpose because it can learn nonlinear relation while removing the noises to some extent as its characteristic.

In this study, the feasibility of the prediction without neglecting the fluctuation in a short period is investigated by the approach that uses a digital FIR band-pass filter and a neural network. But a problem arises in order to use the digital filter for the data processing of the financial time series. The problem is that it is impossible to obtain some filtered data for the last part of the financial time series. The last part of the financial time series is the most important for the prediction. To improve the problem, the special procedure is tried in this study. This method is verified using the Rossler chaos time series modified to economical one [4].

2 LEARNING AND PREDICTION OF THE TIME SERIES WITH THE BAND-PASS FILTER

2.1 Concept of the method

It is assumed that the time series data are determined by the past data.

It is expressed by the following equation.

$$\{g_n\}: g_n = f(n, g_{n-1}, g_{n-2}, \cdots, g_{n-k}) \quad \ldots\ldots (1)$$

If the structure of the function f is known or simple, the regression analysis can be used. If the function is complicated, the neural network has often been used for forecasting [2] [3]. But if the function is too complicated, it is no longer possible even to use the neural network.

In this study, such time series data are filtered by digital band pass filter into several frequencies regions and the each

filtered data $\{g_n^j\}$ is considered to be simpler structure than the original data $\{g_n\}$ and the data $\{g_n^j\}$ are assumed to be expressed by the function f_j and the function is determined by the past data of the filtered data. These are expressed as follows.

$$\{g_n\} = \{g_n^1\} + \{g_n^2\} + \cdots + \{g_n^j\} + \cdots + \{g_n^l\} \quad \ldots\ldots (2)$$

The $2m+1$ sampled data from T_{-m} to T_m are used for the filtering and the number "$2m+1$" is called the filter length. In case of filtering at the designed frequency, the characteristics become better showing steep roll-off as the filter length becomes longer.

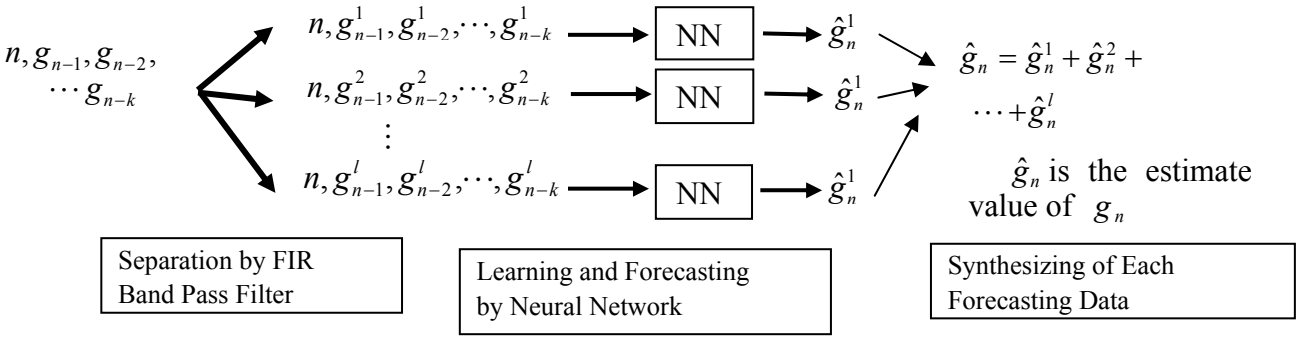

Fig.1 Concept of the forecasting method

$$g_n^j = f_j(n, g_{n-1}^j, g_{n-2}^j, \cdots, g_{n-k}^j) \quad \ldots\ldots (3)$$
$$(j = 1, 2, \cdots, l)$$

Eq (2) is the characteristic of the digital FIR filter as well.

The procedure of the learning and the forecasting by the neural network is shown in Fig.1.

2.2 Digital filter (FIR band-pass filter)

The band-pass filter passes the input waveform from the lower cutoff frequency f_{BL} (Hz) to the upper cutoff frequency f_{BH} (Hz) by constant amplitude ratio 1 and it is zero outside the pass-band.

The equation to obtain the filtered value F_0 from the sampled value T_m of the waveform is

$$F_0 = \sum_{m=-n}^{n} T_m \cdot H_m \quad \ldots\ldots (4)$$

where H_m is the filter coefficient that is calculated by the sampling frequency f_s (Hz) and the cutoff frequencies f_{BH} and f_{BL}.

$$H_m = 2\cos\left(m\frac{\omega_{BH} - \omega_{BL}}{2}\right)\left(\sin\left(m\frac{\omega_{BH} + \omega_{BL}}{2}\right)\Big/ m\pi\right)$$

where $\ldots\ldots (5)$

$$H_0 = (\omega_{BH} + \omega_{BL})/m\pi$$

and

$$\omega_{BH} = 2\pi\frac{f_{BH}}{f_s}, \quad \omega_{BL} = 2\pi\frac{f_{BL}}{f_s}.$$

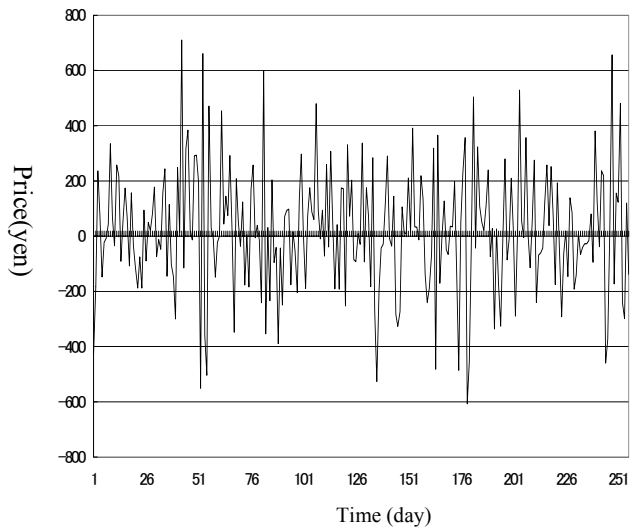

Fig.2 Time series of Nikkei Stock average (1999/1/4~2000/1/21)

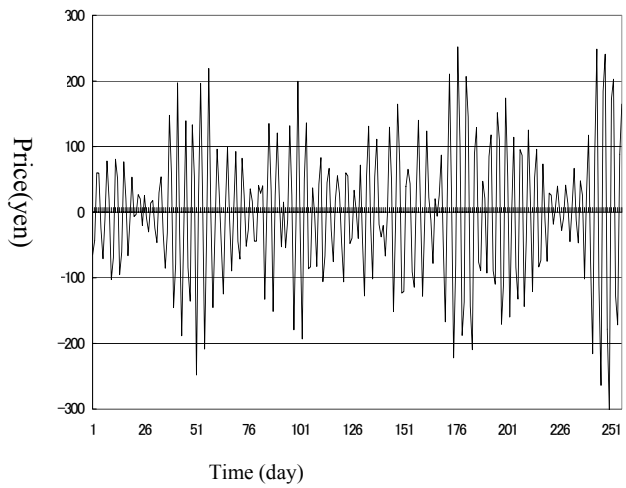

Fig.3 Filtered time series of the stock

The FIR band-pass filter of 101 lengths with Hamming window filters the differential waveform of the Nikkei stock average closing price. As the length of FIR filter becomes longer, the characteristics of the filter become better. The FIR filter of 101 lengths shows an excellent characteristic with a flat pass band and a sharp edged cutoff frequency.

In this study, the time series is filtered into the five frequency bands of 0-1Hz, 1-2Hz, 2-3Hz, 3-4Hz and 4- 5Hz by setting the sampling frequency 10Hz. (the actual sampling period: 1 day) The divided waveforms can be returned to the original waveform by synthesizing them.

Fig.2 shows the time series of Nikkei stock average from January 4,1999to January 21, 2000.

Fig.3 shows the time series of the filtered data of Fig. 2 by the band pass filter where f_{BL} is 2Hz and f_{BH} is 3Hz as the sampling frequency f_s is 10Hz (the actual sampling period: one day).

The filtered waveform is more sinusoidal as compared with the original one. Therefore, it is considered that the learning with the neural network becomes easier.

Fig.4 is the result of the FFT analysis of Fig.2 and Fig.5 is that of the Fig.3.which is obtained by filtering the above data with the 1-2 Hz band-pass filter of length 101.

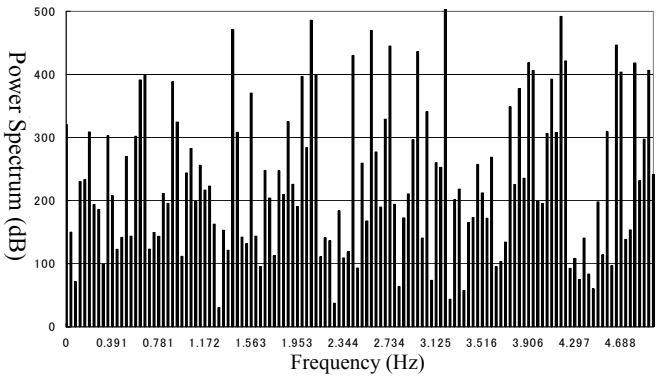

Fig.4 FFT analysis of Fig. 2

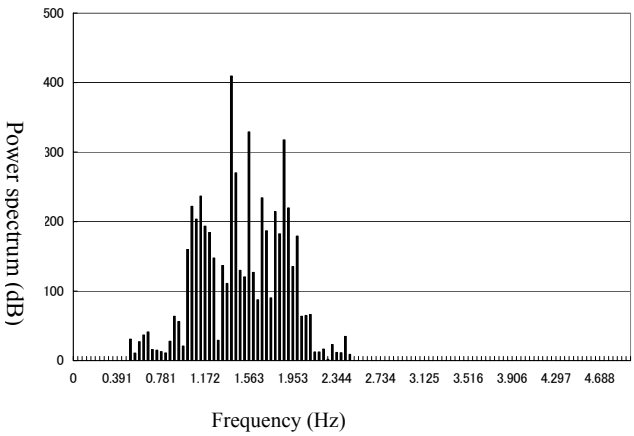

Fig.5 FFT analysis of Fig.3

2.3 The Rosslea Chaos Time Series

The deterministic complicated data without random data are made by the following Eq (6). The equation is a type of the Rossler chaos and is modified in order to apply to the economical simulation by Goodwin [4]. The data are used as the time series to confirm the validity of the method developed in this study.

The time series data of the equation are calculated by the finite difference method.

$$\begin{cases} \dot{v} = -du + fv - ez \\ \dot{u} = hv \\ \dot{z} = b + gz(v - c) \end{cases} \qquad \ldots\ldots (6)$$

Where v :employment ratio, u :unit labor cost, z :policy variable.

The parameters are $d = 0.50$, $f = 0.15$, $e = 0.30$, $h = 0.50$, $b = 0.01$, $g = 85.0$, $g = 85.0$, $c = 0.05$.

The initial conditions are $v = 0$, $u = 0.03$, $z = 0.02$.

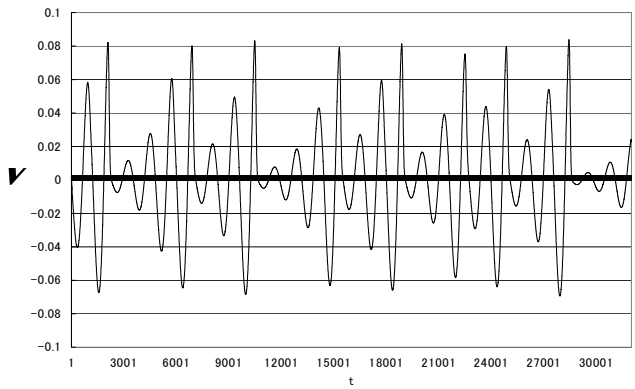

Fig.6 Time series of v

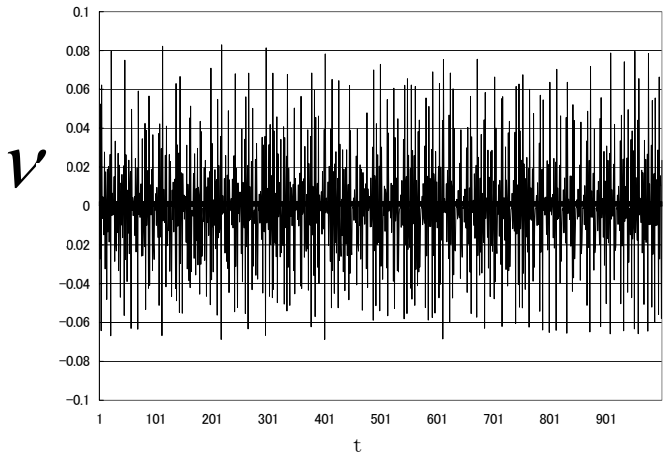

Fig.7 Time series of every 500 data of Fig. 6

Fig. 6 shows the time series of v . The Euler difference method is applied while the time difference $\Delta t = 0.01$.

It is a problem that the data in Fig.6 are quite different from the Nikkei stock average, which is used in this study. Fig.7 is every 500 data of Fig.6. The time series is similar to that of Nikkei stock average in this case. Therefore the data of Fig.3 are used to verify the effectiveness of the method in this study.

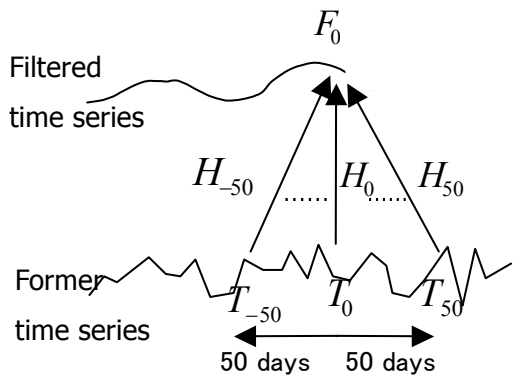

Fig. 8 Filtering with the 101 length band path filter

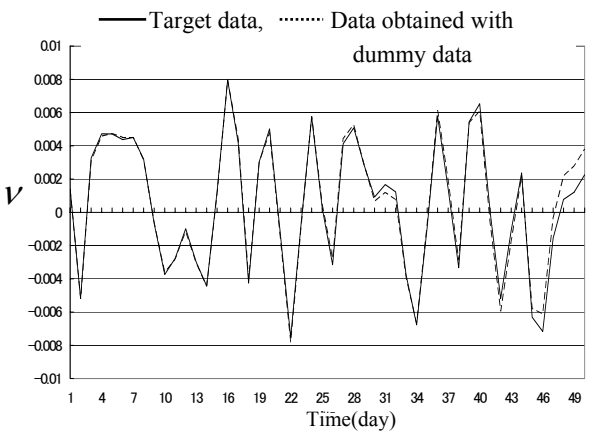

Fig. 9 The target and the filtered data

2.4 *Learning and Forecasting of Waveform by the Neural Network*

The each filtered waveform is learned and forecasted by the neural network. The neural network of the back propagation method is adopted in the learning the waveform.

After learning the time series of each frequency band by the neural network, the data of 20 days in the past are inputted and the prediction of ten days ahead is carried out. Each predicted data are combined. The combined data are compared with the actual stock data.

2.5 *The Problem of the Prediction*

As the length of the band-pass filter is 101 at the prediction with the band-pass filter, the data of the back and the forth 50 days (samples) are used for filtering. The last 50 filtered data which are the most important for the prediction can not be obtained as shown in Fig. 8.

The time series of 1000 days of Nikkei stock average is used for the learning. The data are the difference from the value of the previous day. Therefore, the average of the time series data is zero. For that reason, value zero is added as the dummy data from the time 1001 to 1050. The data that are filtered from the time 50 to 1000 can be obtained by this procedure.

Fig.9 is one of the examples of the comparison between the target data and that obtained by using dummy data. As shown in Fig.9, the differences between the target data and the data obtained with the dummy data is seen only the last two or three of them and the difference is not large.

Anyway, the last two or three data are supposed to have error to some extent. Therefore, the filtered data from time 50 to 997 are used as the learning data of the neural network and the neural network forecasts the data from time 998 to 1007. By this procedure, it is possible to estimate the data from time 1001 to 1007.

3 THE LEARNING AND FORECASTING OF WAVEFORM BY THE NEURAL NETWORK

3.1 *The Rossler Chaos*

First, to verify the effectiveness of this method, the Rossler chaos time series is predicted by this method. In this case, the simple neural network prediction that is done without using the digital filter are also carried out as the conventional one for the comparison with the method developed in this study. This is called “the simple method” and the method using the digital filter is called “the filtering method” in this report.

The neural network by the back propagation is used in learning. The construction of the neural network is composed of four layers, where the cell number of input layer is 20, the hidden layer1 17, the hidden layer2 14 and the output layers10.

The input data are that of 20 days that are obtained by filtering the chaos data for each frequency band. The number of the out put cells of the neural network is 10. However, the last three data of the filtered data are not used for the prediction because the filtered data are obtained with the dummy data ant they are supposed to have error to some extent. Therefore a substantial prediction becomes after the fourth day and therefore, it becomes seven days prediction simultaneously.

Fig. 10 shows the comparison among the target data and forecasted data by this method. Table1 shows the comparison of the correlation coefficient between the both methods. As shown in Fig.10 and Table1, the both methods show very good coincident with the target data. This is because the Rossler chaos time series is not random data but the deterministic one that is obtained from the calculation. Anyway, it is proved that the filtering method developed in this study can be used for forecasting.

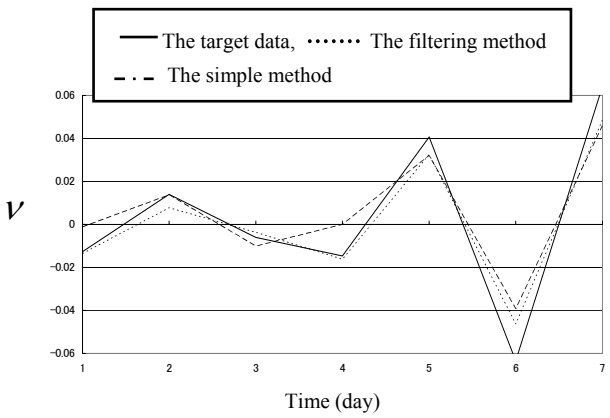

Fig. 10 Comparison among the target data and the predicted data

Table 1 Comparison of the correlation coefficient

Time	Correlation coefficient Upper: The filtering method Lower: The simple method
1～1000	0.865 0.852
501～1500	0.984 0.996

3.2 *Nikkei Stock Average*

Nikkei stock average is used for the forecasting as the actual financial time series.

The neural network used for the prediction is the same as the one used in section 3.1.

The input data are the data of 20 days that is filtered by the FIR band path filter with the dummy data to the 5 frequency bands as same as section 3.1. Although these data contain noises justly, neural network is able to learn while eliminating the noises to some extent.

The learning data are 997 samples of the filtered data. The learning period is shown in Table2.

As the result of the prediction, there are the cases that are good coincidence to the waveform filtered by length 101 and the other cases that have the shift partly. Generally, the good coincidence is obtained for the high frequency bands. Fig.11 and 12 show the examples of the predicted waveform for 1-2 Hz and 3-4 Hz in the same sequence. The predicted data for each frequency band are combined in order to predict the Nikkei stock average as well as the Rossler chaos.

Fig. 13 shows the example of the combined or final forecast. Fig. 13 also shows the data of the simple method for the comparison.

The filtering method shows the much better result as compared with the simple method. It can be said that it is almost impossible for the simple method to predict the time series.

Table2 shows the comparison between the simple method and the filtering method on the correlation coefficient. Table 2 shows the values for seven days which is obtained simultaneously, and they are on a relatively long period. The correlation coefficients of the filtering method are better than those of the simple method. It can be said to be good results as it is a value for a long period of seven days. The value for the period 1992/10/1~1996/10/16 does not show a good result among that of other period. Then the learning by the other neural network construction (number of input cell: 20, hidden layer1: 20, hidden layer2: 14, output: 10) is tried for the period. Fig.14 shows the comparison among the new construction neural network and previous methods. The forecasting is clearly improved by the new method as compared with the other methods.

Table 3 shows the comparison between the correlation coefficients for the former construction of the neural network and the new construction. It is clearly said that the new one improves the result as compared with the former construction.

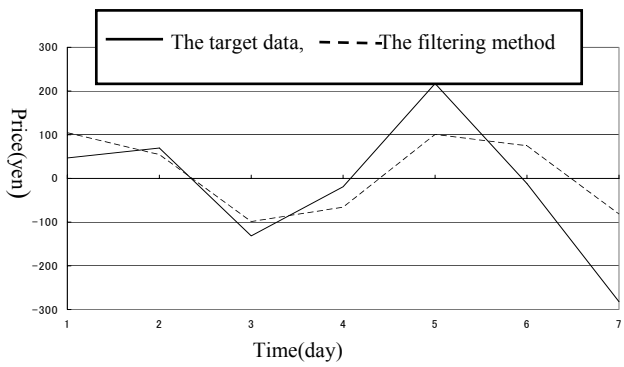

Fig.11 Forecasting of 1Hz-2Hz

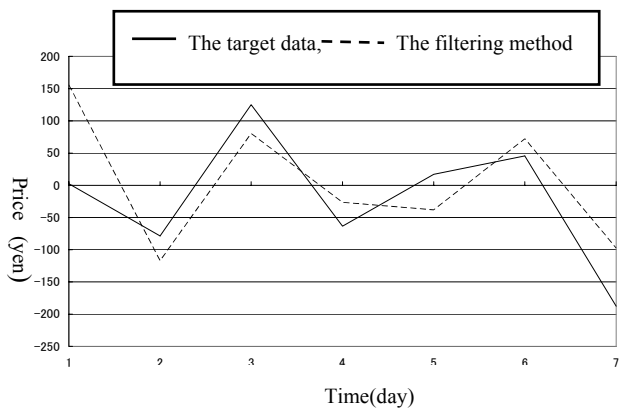

Fig. 12 Forecasting of 3Hz-4 Hz

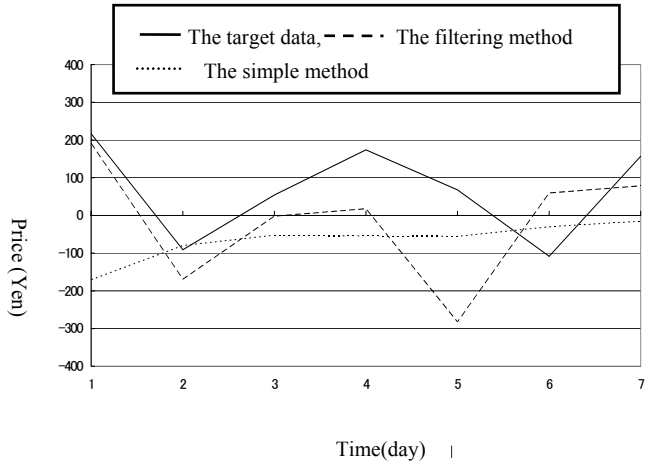

Fig. 13 Forecasting of final time series

Learning period	Upper: The filter method Lower: The simple method
1992/10/1～ 1996/10/16	0.217 0.202
1993/6/16～ 1997/6/30	0.582 -0.725
1994/8/1～ 1998/8/17	0.411 0.619
1995/1/4～ 1999/1/20	0.439 -0.515
1996/3/7～ 2000/3/31	0.649 -0.212

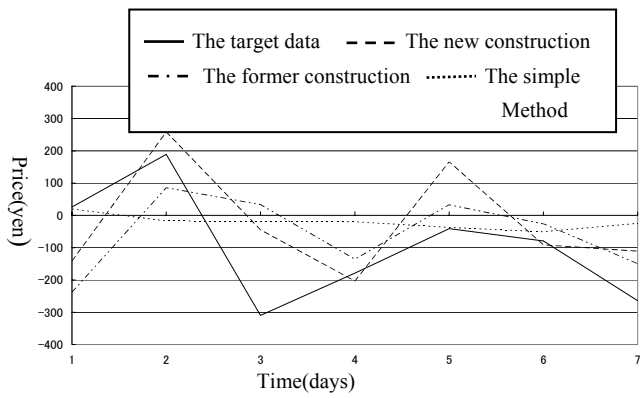

Fig. 14 Forecasting of final time series

Table 3 Comparison of the correlation coefficient

	Correlation coefficient
The former construction	0.216
The new construction	0.617

4 CONCLUSIONS

In this study, the forecast using the FIR digital band-pass filter and the neural network has carried out. FIR band path of 101 lengths is used that is relatively long length. The filter of long length shows the good characteristic though one problem will arise in applying the FIR digital filter of long length to the financial time series. It is a lack of the last part of the filtered data by the half of filter stage. This is solved by adding zero to the original time series as the dummy data.

The original time series is divided into five frequency bands and the each divided time series is learned and predicted by the neural network. As for the prediction of the each filtered waveform, the good results are obtained except for the low frequency bands. This is that the waveform in the high frequency bands which are processed with the band-pass filter is close to a sine wave, and therefore the prediction becomes easy.

The effectiveness of this method is verified by using the Rossler Chaos time series.

It can be said that the prediction of the Nikkei stock average is possible to some extent.

REFERENCES

[1] Saito, S. *A digital filter for economical time series data. The 2nd International Symposium on Frontiers of Time Series Modeling*. The institute of Statistical Mathematics. 232-233 2000

[2] Andresia, K. Vannucci, A. Cesar, E.. *Using artificial neural networks to forecast chaotic time series. Physica A* 284: 393-404 2000.

[3] Qi, M. Zhang, G, P.. *An investigation of model selection criteria for neural network time series forecasting. European Journal of Operational Research* 132: 666-680,. 2001

[4] R.M.Goodwin, *Chaotic economic dynamics, Oxford University Press*, 1990

[5] Takayasu,H. Takayasu,M. *Critical fluctuations of demand and supply, Physica A* 269,24-29,1999