



ELSEVIER

European Journal of Operational Research 122 (2000) 501–508

EUROPEAN  
JOURNAL  
OF OPERATIONAL  
RESEARCH

www.elsevier.com/locate/orms

# Utilization of artificial neural networks and the TD-learning method for constructing intelligent decision support systems

N. Baba <sup>\*</sup>, H. Suto

*Information Science, Osaka-Kyoiku University, Asahiga-Oka, 4-698-1, Kashiwara City, Osaka Prefecture 582-8582, Japan*

Received 1 October 1998; accepted 1 April 1999

---

## Abstract

The use of intelligent techniques in various decision support systems has recently become one of the most exciting and promising application areas. In this paper, we try to utilize Neural Networks (NNs) and the Temporal Difference Learning (TD-Learning) Method to construct an intelligent decision support system for dealing stocks. We also briefly touch upon an intelligent sales prediction system which has been developed in our group. © 2000 Elsevier Science B.V. All rights reserved.

*Keywords:* Neural networks; TD-Learning; Decision support system; Prediction of stock price; Sales prediction

---

## 1. Introduction

The widespread popularity of Artificial Neural Networks (ANNs) in many different fields is mainly due to their ability to build complex non-linear relationships between input variables and output variables directly from the training data [1]. ANNs can provide models for a large class of real systems which are difficult to handle using traditional approaches. Recently, because of ANNs' characteristic property, a large number of people have paid particular attention to apply ANNs to various decision support systems.

However, despite growing concerns for the utilization of ANNs, only limited success has so

far been achieved. This might be due to the reason that making an accurate forecast is a quite hard task. Even by intelligent techniques such as ANNs, GAs, etc., one would find great difficulties in producing reliable predictions.

In this paper, we will try to apply the Temporal Difference Learning (TD-Learning) Method [2] in order to enhance the performance of ANNs for several decision support systems. TD-Learning is an interesting forecasting method which was proposed by Sutton [2] in 1988. However, despite its great ability, there have so far been only few practical successes [3] concerning the application of the TD-Learning Method. We believe that one of the most important advantages of the TD-Learning Method is that it can adapt to sudden environmental changes. In this paper, we will try to apply the TD-Learning Method to the decision

---

<sup>\*</sup> Corresponding author.

support system for dealing stocks using ANNs [4–7] which has been developed in our group since the latter part of 1980s. We will also introduce our trial for constructing an intelligent sales prediction system which utilizes ANNs and the TD-Learning Method.

## 2. TD-Learning

In this section, we will briefly touch upon TD-Learning as proposed by Sutton [2] in 1988.

Let us consider the multi-step prediction problems in which experience comes in observation-outcome sequences of the form  $x_1, x_2, \dots, x_m, z$  where  $x_t$  ( $t = 1, \dots, m$ ) is a vector of observations available at time  $t$  and  $z$  is the outcome of the sequence.<sup>1</sup>

For each observation-outcome sequence, the learner produces a prediction sequence  $P_1, P_2, \dots, P_m$  each of which is an estimate of  $z$ . Here, we assume that each  $P_t$  is a function of  $x_t$  (the observation at time  $t$ ) and weight vector  $w$ :

$$P_t = P(x_t, w). \quad (1)$$

For this prediction problem, Sutton [2] proposed the following TD( $\lambda$ ) family of learning procedures in which  $\nabla_w P_t$  obtained  $k$  steps in the past is weighted with  $\lambda^k$ :

$$\Delta w_t = \alpha(P_{t+1} - P_t) \sum_{k=1}^t \lambda^{t-k} \nabla_w P_k. \quad (2)$$

By letting  $\lambda = 0$  in (2), one obtains

$$\Delta w_t = \alpha(P_{t+1} - P_t) \nabla_w P_t. \quad (3)$$

In the above update rule, called TD(0), the weight update is influenced only by the recent change of predictions.

## 3. Decision support system for dealing stocks which utilizes Neural Networks and TD-Learning

In this section, we propose a decision support system for dealing stocks which utilizes neural networks and TD-Learning. First of all, let us briefly explain the decision support system using neural networks which have been developed in our group since 1988 [4–7].<sup>2</sup>

### 3.1. Decision support system using artificial neural networks

We have developed the decision support system for dealing stocks in order to make a forecast of each stock price in Tokyo Stock Market one month in the future.

Fourteen input variables are chosen as the key factors for the changes of each stock price. In the system, training data are automatically chosen from the vast data having been accumulated for more than five years in order to make a reliable forecast of future stock [7].

A neural network model having 14 input units, one hidden layer with 50 units, and 2 output units has been trained by using the hybrid algorithm [6] which combines the modified BP method and the random optimization method.

Depending upon the outputs from the 2 output units of the trained neural network, decisions for dealing stocks are given. Our system has been tested using the data of various listed enterprises in Tokyo Stock Market. Our system has produced rather a big amount of profit even in the simulation periods when Tokyo Stock Market had suffered from the heavy decreasing trend of the average stock price.<sup>3</sup>

<sup>2</sup> We shall introduce a brief outline of our decision support system using ANNs. If you are interested in details, please read the paper [7].

<sup>3</sup> Due to limitation of space, we shall not go into details. Interested readers are kindly asked to read the paper [7].

<sup>1</sup> The components of each  $x_t$  and  $z$  are assumed to be real-valued measurements and a real-valued scalar, respectively.

3.2. Utilization of the TD-Learning Method for the decision support system in order to successfully adapt to sudden environmental changes

In the previous section, we introduced the decision support system (using neural networks) for dealing stocks which have been developed in our group. In that system, the user finds the weight vector  $w^*$  which gives the small value of the total error function of the neural network concerning the past data. Then, he carries out dealing stocks by using the future prediction which is obtained by making use of the outputs from the neural network model with the weight vector  $w^*$ . However, this approach does not update the weight vector involved in the neural network model during the period of forecasting. Therefore, it seems that this system can hardly adapt to the sudden environmental change that affects stock market heavily.

In this paper, we will try to utilize the idea of TD-Learning for our decision support system. The main objective for using TD-Learning is that incremental learning can be realized by TD-Learning.

In the following, let us explain how the TD-Learning Method has been able to be utilized in the neural network-based decision support system for dealing stocks:

1. First of all, we try to decrease the value of the total error function  $E(w)$  of the neural network as much as possible. When we have succeeded in decreasing  $E(w)$ , we stop the total calculation and obtain the corresponding weight vector  $w^*$ .
2. In the forecasting period, let  $w^*$  be the initial value of weight vector  $w$  and start forecasting the future trend of stock price by utilizing the TD-Learning Method.

We shall utilize Eq. (2) for updating the weight vector  $w$  involved in the neural network model.<sup>4</sup>  $P$  corresponds to an output from the neural network model. By the way, the neural network model has two output units.

However, in the original TD-Learning Method [2],  $P$  is assumed to be scalar.

In this paper, we shall introduce the following idea in order to solve the above problem:

Let  $w_t$  and  $x_t$  be the weight vector involved in the neural network model which has been obtained till the time  $t$  and the input vector to the neural network model at time  $t$ , respectively.

Further, let  $\tilde{x}_{1t}$  and  $\tilde{x}_{2t}$  be the output from the upper output neuron at time  $t$  and the output from the lower output neuron at time  $t$ , respectively.

Let  $\Delta\tilde{w}_{i,t}$  ( $i = 1, 2$ ) be  $(w_{t+1} - w_t)$  which is obtained by substituting  $\tilde{x}_{i,t+1}$  and  $\tilde{x}_{i,t}$  for  $P(x_{t+1}, w_t)$  and  $P(x_t, w_t)$  in (2), respectively. Then, let<sup>5</sup>

$$w_{t+1} = w_t + \Delta\tilde{w}_{1,t} + \Delta\tilde{w}_{2,t}. \tag{4}$$

By using the above calculation rule, the weight vector  $w$  is updated incrementally. However, we have to add a brief explanation concerning the derivation of  $\Delta\tilde{w}_{i,t}$  ( $i = 1, 2$ ) in the above equation.

In seeking  $\Delta\tilde{w}_{i,t}$  ( $i = 1, 2$ ), we have to calculate  $\nabla_w P$ , that is,  $\nabla_w \tilde{x}_{o,t}$  ( $o = 1, 2; o = 1 (o = 2)$  corresponds to the upper (lower) output).

We can easily obtain  $\nabla_w \tilde{x}_{o,t}$  by using an idea introduced by Rumelhart et al. [1].

(Due to limitation of space, we shall not go into details. Interested readers are kindly asked to read the paper [8].)

3.3. Computer simulation results

In order to investigate whether the proposed approach using the TD-Learning Method can be successfully utilized in the stock market, we have carried out several computer simulations under the following realistic conditions:

1. A decision maker can commence selling before buying.
2. A decision maker can buy only one unit of each stock during a day. Further, the total amount of each stock which he is able to hold should

<sup>4</sup> Sutton suggested that one had better utilize the following update rule instead of (2):  $w_{t+1} = w_t + \alpha(P(x_{t+1}, w_t) - P(x_t, w_t)) \sum_{k=1}^t \lambda^{t-k} \nabla_w P(x_k, w_t)$ . In this paper, we will utilize the above update rule.

<sup>5</sup> The reason why (4) has been introduced for updating the weight vector  $w$  can be easily explained. However, due to limitation of space, we will omit to explain it. Interested readers are kindly asked to consider the reason.

be kept below three units. In this simulation, we have assumed that one unit =  $10^3$  shares.

3. A decision maker can sell only one unit for each stock during a day.
4. A decision maker cannot sell if (Total volume sold – Total volume bought) exceeds three units for each stock.
5. All of the transactions having been done should be cleared off at the end of the simulation period. That is to say, all of the stocks having been held should be sold out and all of the stocks having been sold should be bought at the end of the simulation period.

From the various posts in Tokyo Stock Market, the following 16 representative listed enterprises have been chosen:

- (1) Teikoku Sekiyu (1601)
- (2) Taisei Kensetsu (1801)
- (3) Nihon Seifun (2001)
- (4) Toure (3402)
- (5) Asahi Kasei (3407)
- (6) Mitsubishi Kasei (4010)
- (7) Sankyoku (4501)
- (8) Nihon Sekiyu (5001)
- (9) Shin Nihon Seitetsu (5401)
- (10) Fujitsu (6702)
- (11) Matsushita Densan (6752)
- (12) Sony (6758)
- (13) Toyota Jidousha (7203)
- (14) Sumitomo Bank (8318)
- (15) Nomura Shouken (8604)
- (16) Nihon Yusen (9101)

*Simulation 3.1.* By utilizing the data from the beginning of March in 1988 to the end of February in 1993, training of the weight vector in the neural network model has been done. Computer simulations of dealing stocks utilizing the TD-Learning Method (TD(0)) have been carried out during the period from the beginning of March in 1993 to the end of February in 1994. (Fig 1 shows the changes of the average stock price in the Tokyo Stock Market in the simulation period.) Table 1 compares the total profits obtained by utilizing the TD-Learning Method (TD(0)) with those obtained by using the neural network model. The proposed system utilizing the TD-Learning Method has

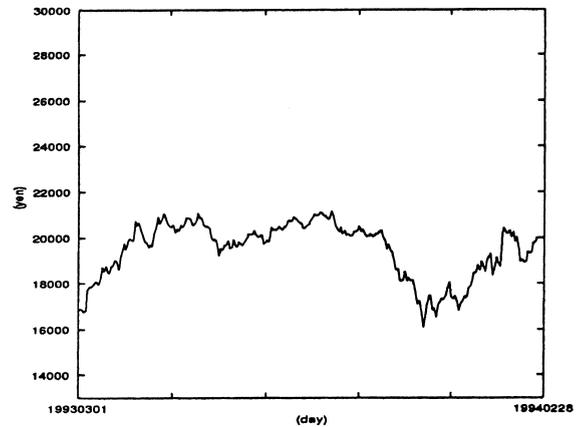


Fig. 1. Changes of the average stock price in Tokyo Stock Market (March 1993–February 1994).

succeeded in producing a big amount of profit (total: 8098 thousand yen). This comes from the fact that total profit after subtracting charge has greatly increased not only due to the growth of profit obtained by dealing stocks but also due to the decrease of charge for dealing stocks.

In order to investigate what kind of concrete effect has been incurred by the introduction of the TD-Learning Method, we have compared the simulation results concerning the stocks (1801, 6702). Figs. 2 and 4 show the simulations of dealings by the neural network. On the other hand, Fig. 3 and Fig. 5 show the simulations of dealings by the TD-Learning Method. Although dealing stocks have been carried out quite frequently in the simulations using neural networks, comparatively large amount of fruitless dealings have been omitted in the simulations utilizing the TD-Learning Method (Figs. 3 and Fig. 5). From the changes of each stock price during a year after the simulation period, we can easily understand the adequateness of the dealings done by the utilization of the TD-Learning Method.

**Remark.** The above simulation results suggest that utilization of the TD-Learning Method could be a powerful approach in order to adapt successfully to sudden environmental changes. However, this has been confirmed only by several examples. In order to show the effectiveness of the utilization of

Table 1

Comparison between total profits obtained by the TD-Learning Method (TD(0)) and neural network and those obtained by the neural network <sup>a</sup> (simulation period: March 1993–February 1994)

Stock	Neural network			TD(0) + Neural network		
	Profit	Charge for dealing	Profit – charge	Profit	Charge for dealing	Profit – charge
1601	0	0.000	0.000	0	0.000	0.000
1801	308	326.850	-18.850	147	60.165	86.835
2001	-238	50.190	-288.190	-149	33.315	-182.315
3402	-101	76.995	-177.995	-35	18.495	-53.495
3407	0	0.000	0.000	0	0.000	0.000
4010	159	129.015	29.985	105	38.565	66.435
4501	0	0.000	0.000	0	0.000	0.000
5001	-27	68.355	-95.355	233	64.455	168.545
5401	0	0.000	0.000	0	0.000	0.000
6702	-778	230.790	-1008.790	602	81.870	520.130
6752	1090	140.250	949.750	1810	129.450	1680.550
6758	4420	489.900	3930.100	4420	489.900	3930.100
7203	1800	265.800	1534.200	1720	156.000	1564.000
8318	710	181.950	528.050	490	185.250	304.750
8604	1110	446.250	663.750	70	57.450	12.550
9101	-202	68.250	-270.250	0	0.000	0.000
Total	8251	2474.595	5776.405	9413	1314.915	8098.085

<sup>a</sup> Thousand yen.

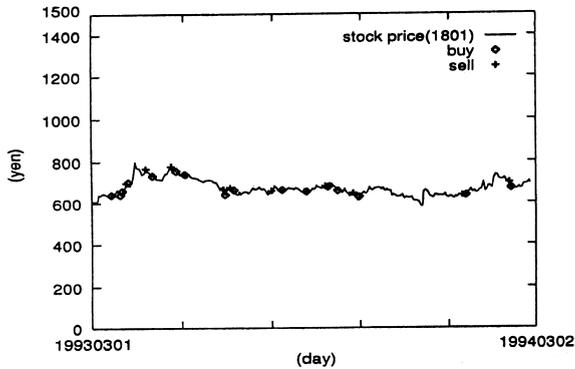


Fig. 2. Dealings having been done by ANNs (1801).

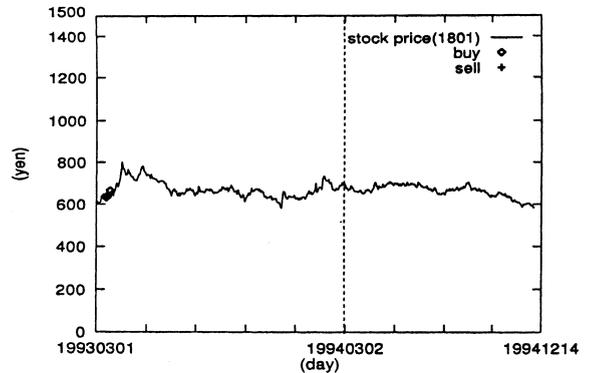


Fig. 3. Dealings having been done by ANNs and the TD-Learning Method (1801).

the TD-Learning Method, we have to investigate our conjecture by applying the approach to various examples. We have also to concentrate ourselves to give theoretical assurance.

*Simulation 3.2.* We have also carried out computer simulations during the same period assuming that  $\lambda$  takes one of the six values ( $\lambda = 0, 0.2, \dots, 0.8, 1.0$ ) and  $\alpha$  takes one of the eleven values ( $\alpha = 0, 0.1, \dots, 0.9, 1.0$ ). Fig. 6 shows the total profits having been obtained by TD( $\lambda$ ).

From the above simulation results,<sup>6</sup> we can observe that the TD-Learning Method has played an important role for improving our decision support system using a neural network model [4–7].

<sup>6</sup> We have also carried out several computer simulations during other periods. However, due to limitation of space, we will omit them.

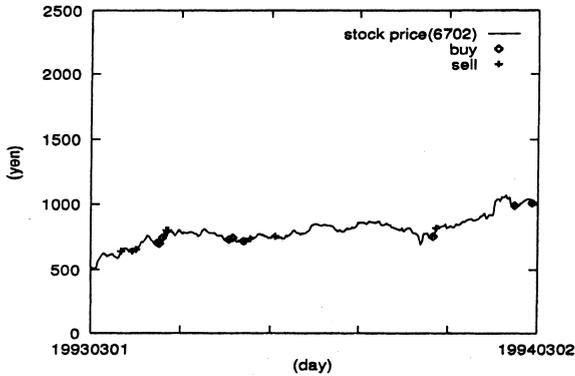


Fig. 4. Dealings having been done by ANNs (6702).

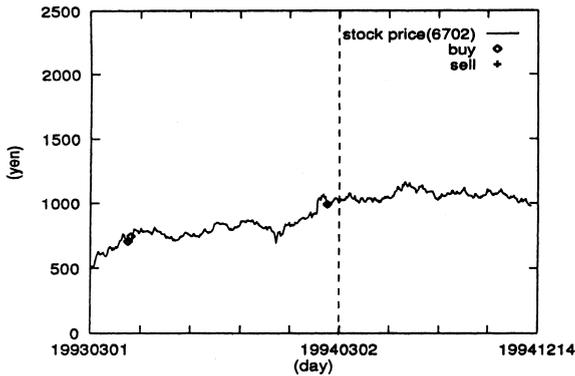


Fig. 5. Dealings having been done by ANNs and the TD-Learning Method (6702).

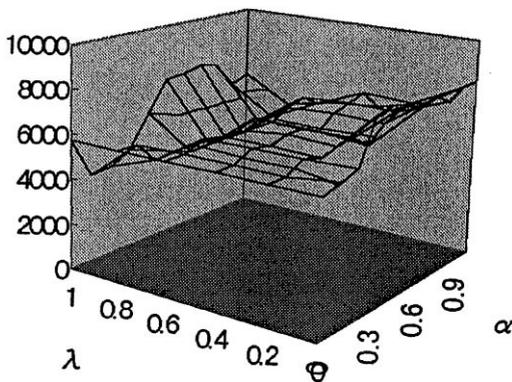


Fig. 6. Total profit for 16 stocks versus parameters  $\alpha$  and  $\gamma$  in  $TD(\lambda)$  (March 1993–February 1994) (thousand yen).

#### 4. A sales prediction system which utilizes NNs and TD-Learning

Making a good prediction concerning sales of each manufactured product is quite important for an enterprise. If an appropriate prediction cannot be given, the enterprise will be heavily damaged due to surplus or shortage of stocks.

In this section, we propose a system which utilizes artificial neural networks for making a prediction regarding sales of an electronic part three months ahead.

Further, we suggest that the proposed system becomes quite helpful when it is supported by the TD-Learning Method.

##### 4.1. Neural network model

Our neural network model consists of 24 input units, two hidden layers having 20 units and 15 units, and one output unit.

##### 4.2. Input variables

In order to design a system which makes appropriate predictions concerning the sales of an electronic part three months ahead, we have chosen the following 24 input variables:

$$x_1 \dots x_8,$$

sales of an electronic part ( $i$  months ago)/sales of an electronic part in the same month one year ago ( $i = 0, 1, 2, 3, 4, 6, 8, 10$ );

$$x_9 \dots x_{16},$$

volume of orders of an electronic part ( $i$  months ago)/volume of orders of an electronic part in the same month one year ago ( $i = 0, 1, 2, 3, 4, 6, 8, 10$ );

$$x_{17} \dots x_{24},$$

volume of production of portable telephones ( $i$  months ago)/volume of production of portable telephones in the same month one year ago ( $i = 0, 1, 2, 3, 4, 6, 8, 10$ ).

### 4.3. Output function

In our system, we have used the bipolar function (5) as the output function of each neuron:

$$y = \frac{1 - \exp(-\sigma x)}{1 + \exp(-\sigma x)} \tag{5}$$

### 4.4. Output from the Neural Network model and teacher signal

We have tried to construct the neural network model whose output gives an information concerning the increase (decrease) rate of (sales three months ahead/sales in the same month one year ago).

Therefore, teacher signals have been prepared in order to satisfy this objective.

### 4.5. Computer simulation results

In order to investigate whether an appropriate sales prediction can be made by the proposed system, we have carried out several computer simulations. In the following, we shall show some of these simulation results.

*Simulation 4.1.* The neural network model (24 × 20 × 15 × 1) has been trained by using the data from October 1992 to March 1995. Fig. 7 illustrates predictions (during the period from April 1995 to March 1997) having been made by the trained neural network.

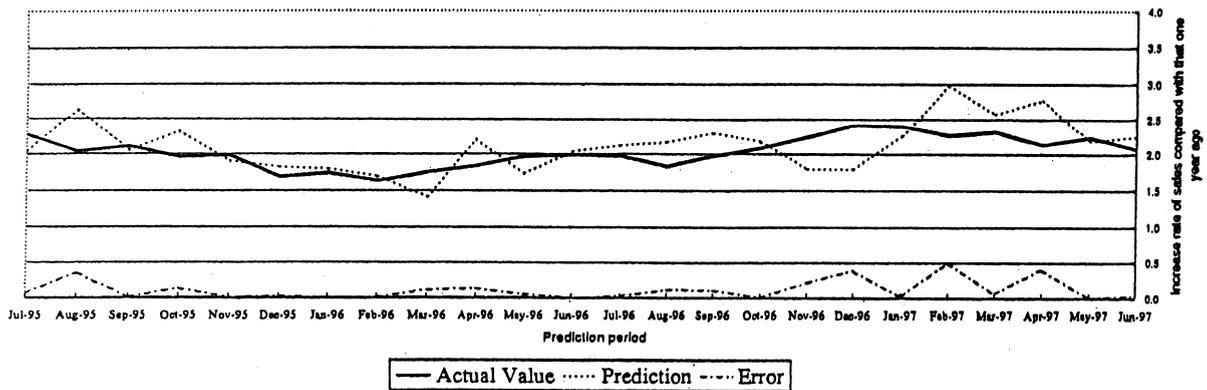


Fig. 7. Predictions having been made by ANNs.

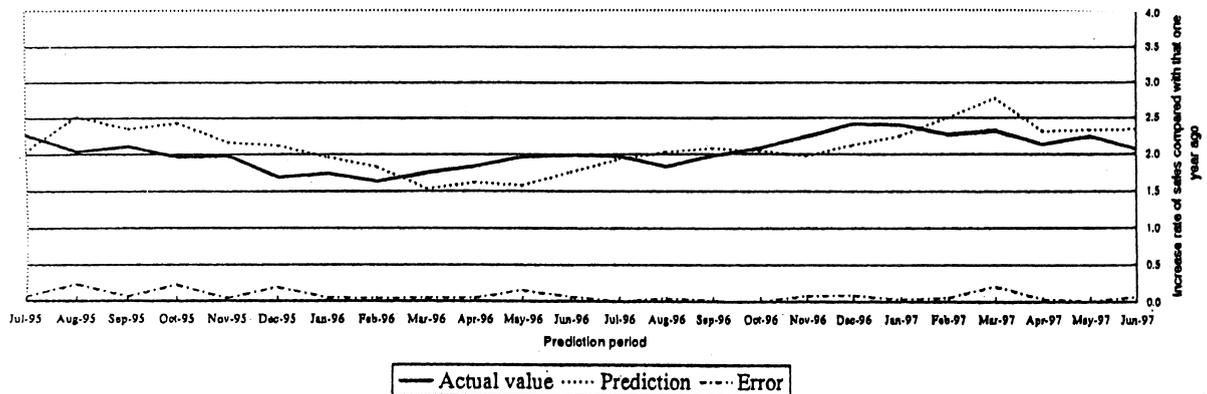


Fig. 8. Predictions having been made by ANNs and the TD-Learning Method.

*Simulation 4.2.* We have tried to utilize the TD-Learning Method in order to enhance the performance of the proposed system. Fig. 8 illustrates predictions having been made when parameters  $\lambda$  and  $\alpha$  have been chosen as  $\lambda = 0$  and  $\alpha = 0.05$ .

**Remark.** The two Figs. 7 and 8 suggest some improvement due to the utilization of the TD-Learning Method. In order to show the improvement clearly, let us compare two approaches using the measures of the total relative error and the maximum error:

*Total relative error:* Fig. 7, 0.114137; Fig. 8, 0.074233.

Here, total relative error is defined to be  $\frac{1}{n} \sum_{i=1}^n (r_i - p_i)^2$  where  $n$  is the total number of the months,  $r_i$  the actual value at the  $i$ th month, and  $p_i$  is the prediction.<sup>7</sup>

*Maximum error:* Fig. 7, 0.493053; Fig. 8, 0.229036.

The above two measures show that the proposed approach utilizing TD-Learning Method has made better predictions.

## 5. Concluding remarks

We have proposed to utilize the TD-Learning Method for constructing an intelligent decision support system for dealing stocks. By combining the TD-Learning Method with the decision support system using a neural network, a remarkable performance has been observed in the computer simulations of dealing stocks in the Tokyo Stock Market. The performance of the proposed approach has been simulated only with respect to the 16 listed enterprises in Tokyo Stock Market. Therefore, in order to show the effectiveness of the

utilization of the TD-Learning Method, we should carry out computer simulations regarding many other enterprises.

We have also briefly touched upon our recent studies concerning the sales prediction system utilizing neural networks. The computer simulation results also suggest the effectiveness of the combined system of the TD-Learning Method and the neural network model.

Future efforts are needed in order to give theoretical background to our approach.

## Acknowledgements

The authors would like to thank the referees for their kind and constructive suggestions.

## References

- [1] D.E. Rumelhart et al., *Parallel Distributed Processing*, MIT Press, Cambridge, MA, 1986.
- [2] R.S. Sutton, Learning to predict by the methods of temporal differences, *Machine Learning* 3 (1988) 9–44.
- [3] G. Tesauro, Temporal difference learning and TD-gammon, *Communications of the ACM* 38 (1995) 58–68.
- [4] N. Baba, M. Kozaki, An intelligent forecasting system of stock price using neural network, in: *Proceedings of the International Joint Conference on Neural Networks*, vol. 1, 1992, pp. 371–377.
- [5] N. Baba et al., A user friendly decision support system for dealing stocks using neural network, in: *Proceedings of 1993 IJCNN Conference*, 1993, pp. 762–765.
- [6] N. Baba et al., A hybrid algorithm for finding the global minimum of error function of neural networks, *Neural Networks* 7 (1994) 1253–1265.
- [7] N. Baba et al., Utilization of neural networks and GAs for constructing an intelligent decision support system to deal stocks, in: *Proceedings of the SPIE Conference*, vol. 2760, 1996, pp. 164–174.
- [8] Baba, TD-Learning with neural networks, *Journal of Robotics and Mechatronics* 10 (1998) 289–294.

<sup>7</sup> We could also compare by using the measure  $\frac{1}{n} \sum_{i=1}^n |r_i - p_i|$  (Fig. 7: 0.271792; Fig. 8: 0.243709).