

## Nonlinear Prediction of River Water-Stages by Feedback Artificial Neural Network

Hiramatsu, Kazuaki

Laboratory of Land-Drainage and Reclamation Engineering, Faculty of Agriculture, Kyushu University

Shikasho, Shiomi

Laboratory of Land-Drainage and Reclamation Engineering, Faculty of Agriculture, Kyushu University

Mori, Ken

Laboratory of Land-Drainage and Reclamation Engineering, Faculty of Agriculture, Kyushu University

<https://doi.org/10.5109/24316>

---

出版情報：九州大学大学院農学研究院紀要. 44 (1/2), pp.137-147, 1999-11. Kyushu University  
バージョン：  
権利関係：

## **Nonlinear Prediction of River Water–Stages by Feedback Artificial Neural Network**

**Kazuaki Hiramatsu, Shiomi Shikasho and Ken Mori**

Laboratory of Land–Drainage and Reclamation Engineering,  
Faculty of Agriculture, Kyushu University 46–05,  
Fukuoka 812–8581, Japan

(Received June 11, 1999 and accepted August 24, 1999)

The feedback artificial neural network model (FBANNM) was applied to the prediction of the water–stages in a tidal river. The difference between a feed forward artificial neural network model and a FBANNM was investigated. A simple genetic algorithm (SGA) was then incorporated into a FBANNM to help search for the optimal network structure, especially the unit numbers of an input layer and a hidden layer. It was concluded that the FBANNM was a useful tool in the short–term prediction of the water–stages that had a strong autocorrelation due to tidal motion. The optimal network structure of the FBANNM was effectively determined by the SGA incorporating the fitness defined by Akaike's Information Criterion.

### INTRODUCTION

For the water–environmental conservation of the flat low–lying area along the lower reach of large rivers, the optimal control of irrigation and drainage systems depends on the accurate understanding of flow regimes. During the high water season, a quantitative understanding of flood runoff phenomena is required in particular for optimal control of drainage system that depends on the accurate prediction of the river water–stages. In the present paper, we examine the feasibility of using an artificial neural network model (ANNM), especially a feedback artificial neural network model (FBANNM), to predict the river water–stages in a tidal river. The FBANNM is an effective model to enable us to predict the time series having a strong autocorrelation structure such as the river water–stages in a tidal river. Next, we use a simple genetic algorithm (SGA) to determine the optimal network structure of the FBANNM.

An artificial neural network is a flexible mathematical structure that is capable of identifying complex nonlinear relationships between input and output data sets. ANNMs have been found to be useful and efficient, particularly in problems in which the characteristics of the processes can hardly be described using physical equations. For this reason, ANNMs have been used in a wide variety of applications. For example, in the field of water resources, ANNMs have been used by Hiramatsu *et al.* (1994) to predict streamflow in a small mountainous basin and water quality in a tidal river, by Hiramatsu *et al.* (1995a, 1995b) to estimate the chlorinity fluctuations in a tidal river.

A major difficulty encountered in applying an ANNM to a variety of prediction problems is that there is no definite method to determine its optimal network structure. In the present paper, a SGA is incorporated into a FBANNM to help search for the structure, especially the unit numbers of an input layer and a hidden layer in the FBANNM.

The genetic algorithm (GA) is a search technique that is designed to mimic some of the most salient features of natural selection and natural genetics. In recent years, a number of researchers have successfully applied the GA technique to certain aspects of the optimization in water resources problems. For example, the GA has been used by Hiramatsu *et al.* (1995b) to determine the parameters in the ANNM for estimating the chlorinity fluctuations in a tidal river.

In the present paper, a brief review of ANNMs and GAs is given, and the difference between a feed forward artificial neural network model (FFANNM) and a FBANNM is described. Next, a new method for incorporating a SGA into a FBANNM to determine the optimal network structure of the FBANNM is described. The FBANNM is then applied to the prediction of river water-stages in a tidal river.

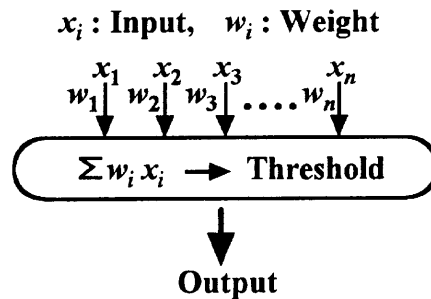
## METHODS AND DATA

### Multi-layer perceptron model

The human brain has the structure that many basic neurons contacted mutually. The function of a neuron is so complicated that there are many obscure parts. However, the basic function is clearly definite. A neuron generates a pulse and transmits it to neighboring neurons when the total amount of the weighted inputs that were received from the neighboring neurons reaches a threshold value. The model that extracted this basic function is a perceptron model shown in Fig. 1. A multi-layer perceptron model is the model made by connecting perceptron models mutually. A three-layer perceptron model (TLPM) is shown in Fig. 2.

The modeling of the artificial neural network began with the proposition of the neuron model by McCulloch and Pitts (1943). Rumelhart *et al.* (1986) proposed the calibration algorithm of the multi-layer perceptron model. The multi-layer perceptron model has been used as the practical model because of the development of this algorithm. This calibration algorithm is called generalized delta law.

In a TLPM shown in Fig. 2 that is composed of an input layer ( $n_1$  units), a hidden layer ( $n_2$  units) and an output layer ( $n_3$  units), the outputs  $\{x_{3k}; k=1, n_3\}$  are computed by:



**Fig. 1.** Schematic of a perceptron model.

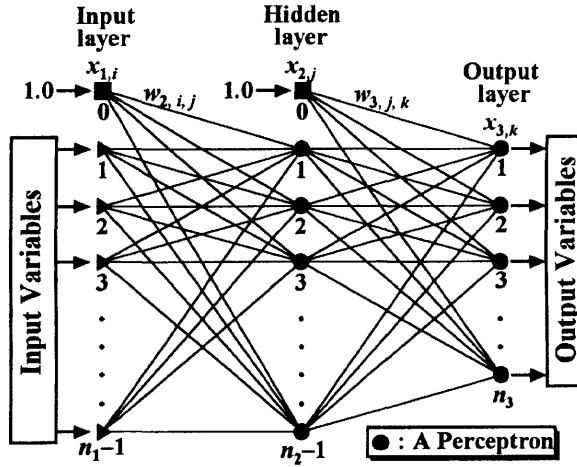


Fig. 2. Typical three-layer perceptron model.

$$x_{2,j} = f \left[ \sum_{i=0}^{n_1-1} w_{2,i,j} x_{1,i} \right] ; j = 1, n_2 - 1, \quad (1)$$

$$x_{3,k} = f \left[ \sum_{j=0}^{n_2-1} w_{3,j,k} x_{2,j} \right] ; k = 1, n_3, \quad (2)$$

where  $x_{1,0} (\equiv 1)$  is the constant that is used for output computation of the hidden layer,  $\{x_{1,i} ; i=1, n_1-1\}$  are the inputs to the hidden layer,  $\{w_{2,i,j} ; i=1, n_1-1, j=1, n_2-1\}$  and  $\{w_{2,0,j} ; j=1, n_2-1\}$  are the input-hidden weights and thresholds to use for output computation of the hidden layer,  $x_{2,0} (\equiv 1)$  is the constant that is used for output computation of the output layer,  $\{x_{2,j} ; j=1, n_2-1\}$  are the inputs to the output layer, and  $\{w_{3,j,k} ; j=1, n_2-1, k=1, n_3\}$  and  $\{w_{3,0,k} ; k=1, n_3\}$  are the hidden-output weights and thresholds to use for output computation of the output layer. Equations (1) and (2) are used in the computations for the hidden layer outputs and the output layer outputs respectively. The nonlinear transfer function  $f[y]$  is the sigmoid function defined by:

$$f[y] = \frac{1}{1 + \exp(-\kappa y)}, \quad (3)$$

where  $\kappa$  is the parameter that prescribes the response of  $f[y]$  as shown in Fig. 3.

In the present paper, the generalized delta law is used for the determination of the weights and the thresholds. In this algorithm, the weights and thresholds are calculated recursively until a sum of square errors between the computed outputs  $\{x_{3,k} ; k=1, n_3\}$  and the calibration data  $\{t_{p,k} ; k=1, n_3\}$  is minimized. Because of the use of sigmoid functions,  $\{t_{p,k} ; k=1, n_3\}$  must be normalized to the range  $[0,1]$ .

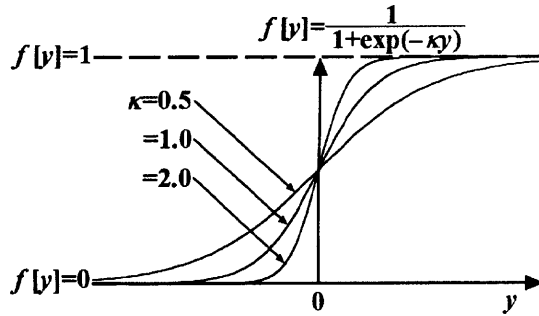


Fig. 3. The sigmoid function most often used for ANNs.

**Feedback artificial neural network model**

Now, we consider the real-time prediction problem to forecast the river water-stages in a tidal river: the time series  $\{h_{t+k}; k=1, N\}$  are predicted by a TLPM using  $\{h_{t-i+1}; i=1, n\}$ . In a tidal river, the important characteristic contained in the variations of the river water-stages is the strong time correlation. When using the FFANNM shown in Fig. 4(a),  $\{h_{t-i+1}; i=1, n\}$  and  $\{h_{t+k}; k=1, N\}$  are allocated to the input layer and the output layer, respectively. In this model, however, a characteristic in the time series is lost because of the discontinuity between  $\{h_{t-i+1}; i=1, n\}$  and  $\{h_{t+k}; k=2, N\}$ .

On the other hand, as shown in Fig. 4(b), a FBANNM feeds back the residuals  $\{|h_{t+k} - \hat{h}_{t+k}|; k=1, N-1\}$  into the input layer and gains the time correlation by these feedback terms, where  $\{\hat{h}_{t+k}; k=1, N-1\}$  are the computation outputs of the network. The feedback terms exist and gain the time correlation in case of the calibration that all the data can be used. In case of the prediction, the feedback terms are cut off because  $h_{t+k} = \hat{h}_{t+k}$  after the calibration of the network. To normalize the residuals  $(h_{t+k} - \hat{h}_{t+k})$ , the absolute values are used. The FBANNM seems to have the same structure as the FFANNM, but the two

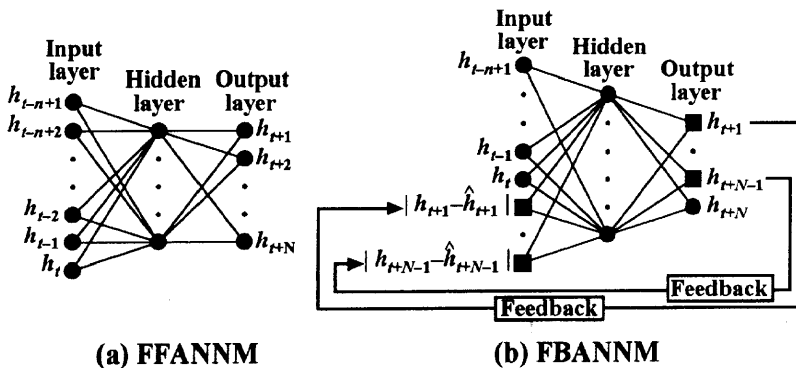


Fig. 4. Assignment of input and output variables to (a) FFANNM and (b) FBANNM.

models are different in the essential function. Matsuba (1992) suggested that the hidden layer perceptrons of the FBANNM extract the characteristic of autocorrelation structure in time series.

In Fig. 4, the first units of the input layer and hidden layer in Fig. 2 are omitted, and  $h_{t+k}$  and  $\hat{h}_{t+k}$  in Fig. 4 correspond to  $t_{p,k}$  and  $x_{3,k}$  in Fig. 2, respectively. In the following, these model structures are represented by the notations FFANNM ( $n, n_m, N$ ) and FBANNM ( $n, n_m, N$ ), where  $n_2=(n_m+1)$  and  $n_3=N$ .

### Optimal network structure

Matsuba (1992) pointed out that the calibration process of an ANN was interfered with by the network elements that had no correlation with the outputs if the network structure became redundant. To exclude the redundancy of network structure and to search for the optimal network structure efficiently, we introduced a GA. The GA is an optimization technique modeled after the biological processes of natural selection and evolution. A GA operates on a population of decision variable sets. Through the application of three specialized genetic operators: selection, crossover, and mutation, a GA population evolves toward an optimal solution. The GAs have been used for the calibration of ANNMs and the determination of the network structure of ANNMs. For example, the GA has been used by Hiramatsu *et al.* (1995a) to determine the optimal network structure of the ANN for estimating the chlorinity fluctuations in a tidal river. In the present paper, a SGA is used to determine  $n$  and  $n_m$  in FBANNM ( $n, n_m, N$ ). The basic outline of a SGA is shown in Fig. 5. A detailed description of the steps in using a SGA to search optimal  $n$  and  $n_m$  is as follows.

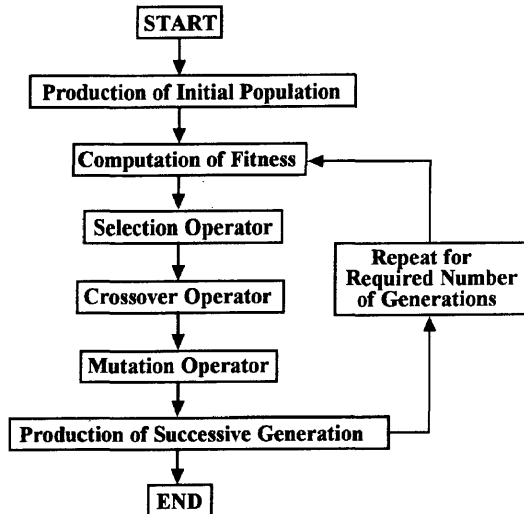


Fig. 5. Flow diagram of a SGA.

### (Step 1) Generation of initial population

An initial population of coded strings representing the network structures is randomly generated with population size  $N_p=30$ . A coded string consists of 2 coded substrings each of 5 binary bits. This coded string of 10 binary bits  $\{b_i; i=1, 10\}$  represents  $n$  and  $n_m$  in FBANNM  $(n, n_m, N)$  as follows:

$$n=2^4b_1+2^3b_2+2^2b_3+2^1b_4+2^0b_5+1, \quad (4)$$

$$n_m=2^4b_6+2^3b_7+2^2b_8+2^1b_9+2^0b_{10}+1. \quad (5)$$

Using equations (4) and (5),  $n$  and  $n_m$  are varied over the range 1 to 32 respectively.

### (Step 2) Computation of the fit to the calibration data

For each  $N_p$  of the strings, a coded string is decoded into  $n$  and  $n_m$  by equations (4) and (5). The weights and thresholds of FBANNM  $(n, n_m, N)$  are determined by the generalized delta law with 100 iterations. The fit to the calibration data is evaluated using two residual statistics: the mean-square error  $E_1$  and the Akaike's Information Criterion AIC (Akaike, 1976).  $E_1$  and AIC are computed using the equations:

$$E_1 = \frac{1}{n_t} \frac{1}{N} \sum_{t=1}^{n_t} \sum_{k=1}^N \{h_{t+k} - \hat{h}_{t+k}\}^2, \quad (6)$$

$$n_f = n + n_m, \quad (7)$$

$$E_2 = \frac{1}{N} \sum_{t=1}^{n_t} \sum_{k=1}^N \{h_{t+k} - \hat{h}_{t+k}\}^2, \quad (8)$$

$$\text{AIC} = n_t \log_e E_2 + 2n_f, \quad (9)$$

where  $n_t$  is the number of time steps in the calibration data and  $n_f$  is the total number of units in the input layer and the hidden layer to be identified. Notice that while the mean-square error statistics are expected to progressively improve as more units are added to the model, the AIC statistics penalize the model for having more units and therefore tend to result in more parsimonious models.

### (Step 3) Computation of the fitness

The fitness of each string must be determined to select which strings in the population get to reproduce. To compare AIC with the mean-square error  $E_1$ , two forms of fitness function are tried:

$$F_1 = \frac{1}{E_1} - \frac{1}{E_1^{\text{Max}}}, \quad (10)$$

$$F_2 = \frac{1}{\text{AIC}} - \frac{1}{\text{AIC}^{\text{Max}}}, \quad (11)$$

where  $( )^{\text{Max}}$  is the maximum value in the current population. To adjust the calculated raw fitness of strings in the population and maintain appropriate levels of competition between the strings, the scaling of fitness is introduced in equations (10) and (11).

### (Step 4) Generation of a new population using the selection operator

New members of the next generation are generated by a proportionate selection method with the elite preservation strategy.

**(Step 5) The crossover operator**

One-point crossover is generally used in a SGA. However, Hiramatsu *et al.* (1995a, 1995b) have used uniform crossover because of its good convergence. The uniform crossover is performed with the probability of crossover  $p_c=0.5$  for each pair of parent strings selected in step 4, that is, on average 100  $p_c$ % of the pairs will undergo crossover.

**(Step 6) The mutation operator**

Mutations are introduced into the population on a bit-by-bit basis at every generation to prevent the algorithm from converging to a local optimum. The mutation is performed with the probability of mutation  $p_m=0.05$  for each bit in the strings.

**(Step 7) Production of successive generations**

Using steps 2 to 6 described above, a new generation is produced. Steps 2 to 6 are repeated until the number of generations exceeds the preselected number of generations  $N_g=20$ .

**(Step 8) Final calibration of FBANNM**

The optimal string that has the maximum fitness in the final population obtained after  $N_g$  generations is decoded into  $n$  and  $n_m$  by equations (4) and (5). The weights and the thresholds of FBANNM ( $n, n_m, N$ ) is finally determined by the generalized delta law with 60,000 iterations.

**River water-stage data**

The data used in this study are two-year hourly water-stage data obtained at Morodomi Gauging Station, Saga Prefecture, Japan. Morodomi is situated on the tidal

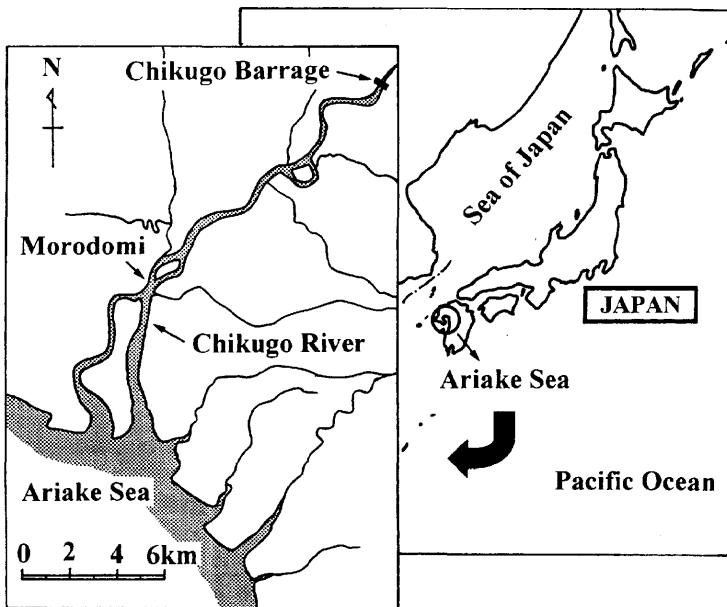


Fig. 6. Outline of the lower reach of the Chikugo River.



compartment of the Chikugo River. The lower reach of the Chikugo River is shown in Fig. 6. During spring tide, the maximum tidal variation in the Ariake Sea is about 6 m. Morodomi is located 9 km up from the mouth of the river and the Chikugo Barrage, located 14 km up from Morodomi, is the upper end of the tidal compartment.

## RESULTS AND DISCUSSION

To compare FBANNM with FFANNM, FFANNM ( $n, 9, 3$ ) and FBANNM ( $n, 9, 3$ ) are applied to the prediction of river water-stages obtained at Morodomi in the Chikugo River shown in Fig. 6. Eleven sets of hourly water-stage data observed in eleven consecutive months were selected for model testing. The models were calibrated by the data until 9 a.m. on the 26th each month and predicted the water-stages at 10, 11 and 12 a.m. on the day. Prediction performance was evaluated using the root-square error  $E_{3,T}$  defined by:

$$E_{3,T} = \sqrt{(h_{t+T} - \hat{h}_{t+T})^2}, T = 1, N. \quad (12)$$

The average values over the eleven months are shown in Table 1. By way of example, the results of the prediction using FFANNM (36, 9, 3) and FBANNM (36, 9, 3) are shown in Fig. 7. Table 1 shows that the errors are smaller for FBANNMs than for FFANNMs in all  $n$  and  $N$ . The effect of acquisition of the autocorrelation structure in the time series by the feedback terms is suggested.

The FBANNM ( $n, n_m, N$ ) is then applied to the prediction of river water-stages at

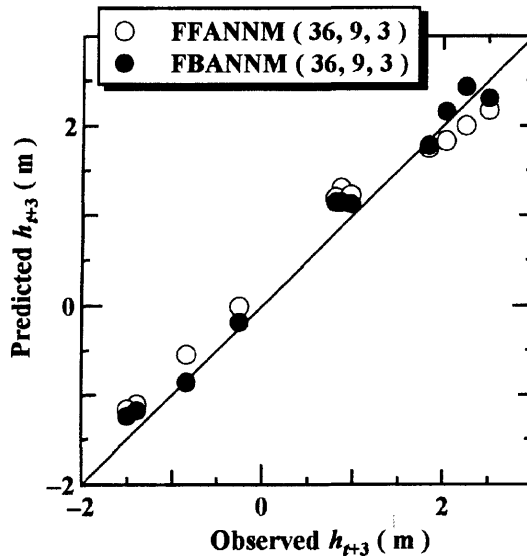


Fig. 7. Scatterplots comparing observed and predicted river water-stages

**Table 1.** Comparison of the average root-mean-square errors

Models		T=1 h	T=2 h	T=3 h
n=6	FBANNM	0.140	0.228	0.305
	FFANNM	0.152	0.239	0.313
n=12	FBANNM	0.138	0.221	0.176
	FFANNM	0.149	0.224	0.186
n=24	FBANNM	0.260	0.269	0.288
	FFANNM	0.264	0.277	0.297
n=36	FBANNM	0.263	0.156	0.170
	FFANNM	0.363	0.199	0.277

**Table 2.** Optimal network structures of a FBANNM obtained by the SGA processes

Data set No.	$F_1$ (by square error)			$F_2$ (by AIC)			Remarks
	$n$	$n_m$	$E_4$	$n$	$n_m$	$E_4$	
1	32	31	0.0608	25	9	0.0825	Low water season
2	32	23	0.0624	26	14	0.0742	
3	31	32	0.0671	31	19	0.0648	
4	31	24	0.0927	25	11	0.1068	High water season
5	31	16	0.1109	25	14	0.1131	
6	32	30	0.1694	25	10	0.1562	

Morodomi and the optimal network structure of the FBANNM ( $n, n_m, N$ ) is determined by the SGA described above. Six sets of hourly water-stage data with seven consecutive days were selected for model testing. The model was calibrated by one data set and validated by the same data set for the seven-day calibration period. Prediction performance was evaluated using the root-mean-square error  $E_4$  defined by:

$$E_4 = \sqrt{\frac{1}{n_t} \frac{1}{N} \sum_{t=1}^{n_t} \sum_{k=1}^N \{h_{t+k} - \hat{h}_{t+k}\}^2} \tag{13}$$

The optimal network structure and the performance of the identified models using the fitness  $F_1$  (defined by mean-square error) and the fitness  $F_2$  (defined by AIC), respectively, are summarized in Table 2. In case of the models using  $F_1$ , the maximum values (=32) of  $n$  or  $n_m$  are obtained in some data sets. On the other hand, the models using  $F_1$  and  $F_2$  result in the average prediction errors of almost the same value. As shown in Fig. 8, the identified models using  $F_2$  show good performance to predict the river water-stages both during low water season and high water season. The AIC statistics penalize the model for having more units and therefore tend to result in more parsimonious models.

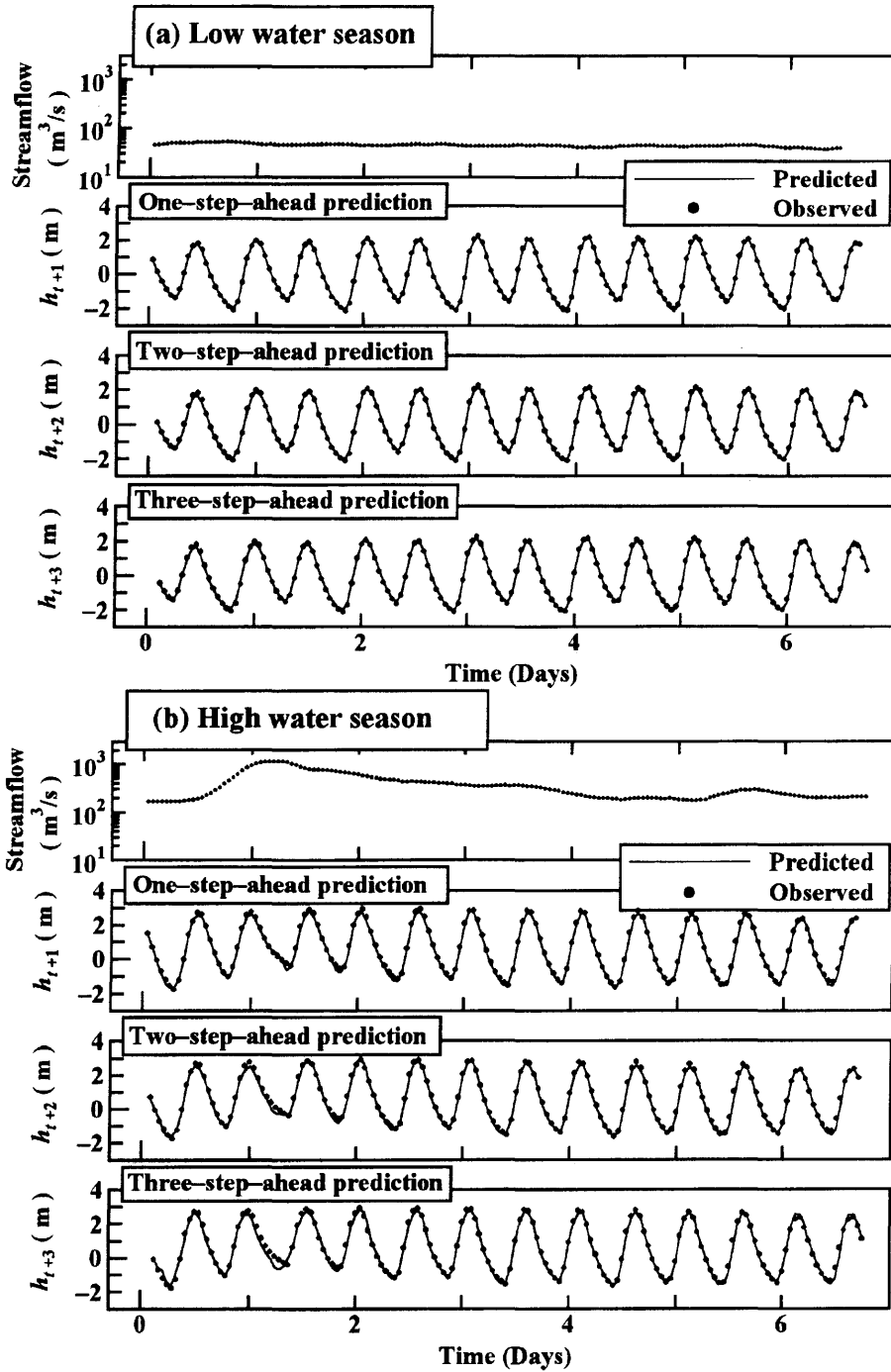


Fig. 8. Observed and predicted river water-stages during (a) low water season and (b) high water season.

## CONCLUSIONS

In the present paper, we examined the feasibility of using the FBANNM to predict the river water-stages in a tidal river and used the SGA to determine the optimal network structure of the FBANNM. It was suggested that the FBANNM was an effective model for short-term prediction of the time series having a strong autocorrelation and the optimal network structure was efficiently obtained by the SGA using the fitness defined by AIC.

## ACKNOWLEDGEMENTS

Partial financial support for this research was provided by a Grant-in-Aid for Scientific Research (B) (No. 08456121) and (C) (No. 09660262) from the fund of the Ministry of Education, Science, Sports and Culture of Japan.

## REFERENCES

- Akaike, H. 1976 An Information Criterion AIC, *Mathematical Sciences*, **153**: 5–11 (in Japanese)
- Hiramatsu, K., S. Shikasho, Y. Tohara, K. Mori and T. Ono 1994 The Predictions of Flood Runoff and Water Quality using Multi-Layered Perceptron Model. *Applied Hydrology*, **6**: 59–68 (in Japanese)
- Hiramatsu, K., S. Shikasho and K. Mori 1995a Application of Multi-Layered Perceptron Model to the Estimation of Chlorinity Variations in Tidal River –Determination of Optimal Inputs using Genetic Algorithm–. *Applied Hydrology*, **7**: 21–30 (in Japanese)
- Hiramatsu, K., S. Shikasho, and K. Mori 1995b Application of Multi-Layered Perceptron Model to the Estimation of Chlorinity Variations in Tidal River. *Transactions of the Japanese Society of Irrigation, Drainage and Reclamation Engineering*, **178**: 83–92 (in Japanese with English abstract)
- McCulloch, W. and W. Pitts 1943 A Logical Calculus of the Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Biophysics*, **9**: 127–147
- Matsuba, I. 1992 Chaos and Prediction. *Mathematical Sciences*, **348**: 64–69 (in Japanese)
- Rumelhart, D., G. Hinton and R. Williams 1986 Learning Representations by Back-Propagation Errors. *Nature*, **323**: 533–536