Chaotic and Fractal Modeling of Water-Stage Time Series in a Tidal River

Hiramatsu, Kazuaki
Laboratory of Drainage and Water Environment, Division of Regional Environment Science, Department of Bioproduction Environmental Science, Faculty of Agriculture, Kyushu University

Shikasho, Shiomi
Laboratory of Drainage and Water Environment, Division of Regional Environment Science, Department of Bioproduction Environmental Science, Faculty of Agriculture, Kyushu University

Mori, Ken
Laboratory of Bioproduction and Environment Information Sciences, Division of Bioproduction and Environment Information Sciences, Department of Bioproduction Environmental Science, Faculty of Agriculture, Kyushu University

https://doi.org/10.5109/24375

出版情報：九州大学大学院農学研究院紀要. 45 (1), pp.255-265, 2000-11. Kyushu University
バージョン：
権利関係：

J. Fac. Agr., Kyushu Univ., **45** (1), 255–265 (2000)

Chaotic and Fractal Modeling of Water–Stage Time Series in a Tidal River

Kazuaki Hiramatsu, Shiomi Shikasho and Ken Mori*

Laboratory of Drainage and Water Environment, Division of Regional Environment Science,
Department of Bioproduction Environmental Science, Faculty of Agriculture,
Kyushu University, Fukuoka 812–8581, Japan
(Received June 27, 2000 and accepted August 18, 2000)

In this paper, we examined the applicability of a state space predictor and a fractal dimension to the short–term prediction of water–stages in a tidal river. The previous researches showed that the state space predictor was an efficient tool, particularly in problems when the characteristics of process could hardly be described by only physical equations. The fractal dimension were also found an effective measure showing the possibility of predictions and to be calculated more precisely, using Higuchi's method, than those presented before. First, hourly water–stage time series was embedded into a state space using time delay coordinates. The induced mapping was then obtained from the embedded trajectory using a local approximation. This enabled us to make the short–term prediction of time series using the information based only on the past values. Second, the fractal dimension calculated by Higuchi's method was incorporated in the state space predictor to estimate the confidence limit of prediction. It was concluded that the state space predictor was a powerful tool in the short–term prediction of water–stages having a strong autocorrelation structure due to tidal motion. The maximum lead–time of prediction was efficiently determined using the fractal dimensions calculated by Higuchi's method.

INTRODUCTION

In flat low–lying areas along the lower reach of large rivers, the accurate prediction of river water–stages is indispensable for the appropriate control and operation of flood drainage systems. The present study works towards two goals. First, we examined the feasibility of using a state space predictor with a local approximation, that has been proposed in the field of chaos engineering (Farmer and Sidorowich, 1987), in order to predict the future behavior of water–stages in a tidal river. Second, we introduced the fractal dimensions calculated by Higuchi's method (Higuchi, 1988) to determine the confidence limit when predicting the future behavior of time series.

The state space predictor embeds a time series into a state space using delay coordinates. The induced mapping is then obtained from the embedded trajectories in the state space using local approximation. This allows us to make short–term predictions for the future behavior of the time series, using information based only on past values. The state space predictor has been found useful and efficient, particularly in problems

* Laboratory of Bioproduction and Environment Information Sciences, Division of Bioproduction and Environment Information Sciences, Department of Bioproduction Environmental Science, Faculty of Agriculture, Kyushu University

when the characteristics of processes can hardly be described using physical equations. For example, in the field of hydrology, the state space predictor has been successfully used by Lall *et al.* (1996) to forecast the Great Salt Lake volume time series, and by Porporato and Ridolfi (1997) to predict river flow sequences.

The fractal dimension is one of the effective measures showing the possibility to forecast a time series and is calculated more precisely, using Higuchi's method, than those presented before (Higuchi, 1988). Matsuba (1992) pointed out that the fractal dimension allowed us to determine the confidence limit in predicting a time series. Hiramatsu *et al.* (1998) utilized the fractal dimension calculated by Higuchi's method to determine the optimal network structure of an artificial neural network model when predicting the water–stages in a tidal river.

METHODS AND DATA

State space reconstruction

Now, we are considering the use of a finite set of river water–stage time series $\{x(t)\}$. From the given time series, we first reconstructed the state space trajectory by using the method of time–delay coordinates (Grassberger and Procaccia, 1983) in order to check the existence of deterministic dynamics. To check the dynamics, we used the state space trajectory of the time series, which plots a time series on one axis, versus the same series with a time delay on the other axis. A state space vector $X[m, \tau, t]$ is constructed, or embedded from m consecutive values of the time series into a state space whose coordinates are described by:

$$X[m, \tau, t]=[\, x(t)\ x(t-\tau) \cdots\cdots x(t-(m-1)\tau)]^{\mathrm{T}}. \qquad (1)$$

The time delay is τ, and the dimension m of the vector is known as the embedding dimension. Following equation (1), a new time series of the state space vectors $X[m, \tau,$

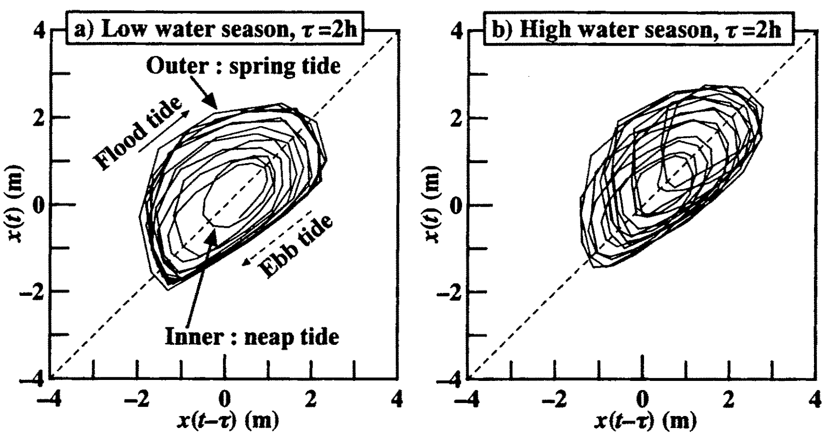

Fig. 1. Examples of two–dimensional state space trajectory for the river water–stages in a) low water season and b) high water season.

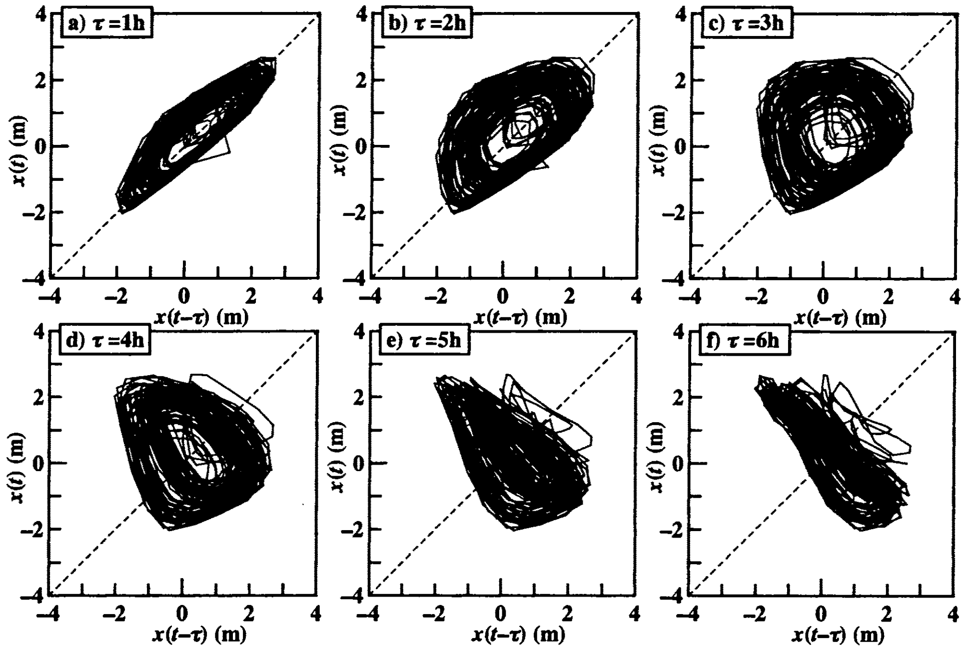

Fig. 2. Two-dimensional state space trajectories for the river water-stages with different delay times.

1], $X[m, \tau, 2]$... are generated. Each state space vector $X[m, \tau, t]$ describes a point in the m–dimensional state space. The sequence of these vectors thereupon defines a trajectory.

Fig. 1a) and Fig. 1b) show examples of the two-dimensional state space trajectories for the water-stages in low water season and high water season respectively. In a tidal river, the important characteristic contained in the variations of the water-stages, is the strong time correlation. The water-stages have clockwise trajectories according to flood tide and ebb tide. Outer and inner trajectories correspond to spring tide and neap tide, respectively. In high water season, the trajectories tend to drift in the direction of increasing the water-stages. As shown in Fig. 1, the trajectories indicate the tidal and flood characteristics generated by the deterministic dynamics in a tidal river. Fig. 2 shows the two-dimensional state space trajectories for the water-stages with different delay times. The above-mentioned dynamical structure of water-stage time series in a tidal river is well represented in the delay times τ of 2 h and 3 h. This typical structure gradually disappears in other delay times and is practically indistinguishable in τ=5 h and 6 h.

State space predictor

Now, we are considering the short-term prediction problem when forecasting the water-stages in a tidal river. The future time series $\{x(t+T); T=1, 2, 3,...\}$ are predicted by the mapping induced from m–dimensional state space trajectories with a time delay of τ. The trajectories are reconstructed by only using known data at the present time

$\{x(t-p);\ p=0, 1, 2, 3,...\}$. The induced mapping is obtained from a local approximation. The state space predictor is a useful and efficient tool, particularly in problems when the characteristics of processes can hardly be described by using only physical equations.

The first step is to predict $x(t'+\tau)$ using the state space vector:

$$X[m, \tau, t']=[\,x(t')\ x(t'-\tau)\ \cdots\cdots x(t'-(m-1)\tau)]^{T}, \quad t'\leq t. \tag{2}$$

In this step, a lead–time of the prediction is assumed to be equal to a time delay of τ. The relationship between the present time t and t' will be explained later. In order to predict $x(t'+\tau)$, we introduced a vector norm on the m–dimensional state space, denoted by $\|\ \|$, and found the n_k nearest neighbors of $X[m, \tau, t']$, i.e. the n_k state space vectors $\{X[m, \tau, t_i], i=1, \cdots, n_k\}$ with $t'\leq t$ that minimized $\| X[m, \tau, t']-X[m, \tau, t_i] \|$.

$$X[m, \tau, t_i]=[\,x(t_i)\ x(t_i-\tau)\ \cdots\cdots x(t_i-(m-1)\tau)]^{T}, \quad i=1,\cdots, n_k. \tag{3}$$

In this paper, we tried two forms of a predictor, i.e. the zeroth–order approximation and the first–order approximation. Farmer and Sidorowich(1987) reported that the results using higher–order polynomials were not significantly better than those obtained with the first order.

In the zeroth–order approximation, $x(t'+\tau)$ is predicted by the average over the n_k past values $\{x(t_i+\tau), i=1,\cdots, n_k\}$ obtained from the nearest neighbors of $X[m, \tau, t']$.

$$x_p(t'+\tau)=\frac{1}{n_k}\sum_{i=1}^{n_k} x(t_i+\tau). \tag{4}$$

In the first–order approximation, a linear regression model is introduced to predict $x(t'+\tau)$.

$$x(t'+\tau)=a_0+a_1x(t')+a_2x(t'-\tau)+\cdots\cdots+a_mx(t'-(m-1)\tau). \tag{5}$$

The parameters $\{a_i;\ i=0,..., m\}$ are determined by equation (6), obtained from substituting the n_k nearest neighbors of $X[m, \tau, t']$ into equation (5).

$$x(t_i+\tau)=a_0+a_1x(t_i)+a_2x(t_i-\tau)+\cdots\cdots+a_mx(t_i-(m-1)\tau), \quad i=1,\cdots, n_k. \tag{6}$$

Here, Y is defined as the parameter vector consisting of $\{a_i;\ i=0,..., m\}$.

$$Y=[a_0\ a_1\cdots\cdots a_m]^{T}. \tag{7}$$

Equation (6) is simplified into the following equation.

$$AY=B. \tag{8}$$

A is a coefficient matrix with the dimension of $[n_k\times(m+1)]$ and B is a constant vector with the dimension of $[n_k\times 1]$. Because of using the linear regression model, n_k needs to be greater than and equal to $m+1$. The parameter vector Y can be obtained using the least–square procedure:

$$Y=(A^{T}A)^{-1}A^{T}B. \tag{9}$$

Using the parameters $\{a_i;\ i=0,..., m\}$ calculated from equation (9), $x(t'+\tau)$ can be predicted by:

$$x_p(t'+\tau)=a_0+a_1x(t')+a_2x(t'-\tau)+\cdots\cdots+a_mx(t'-(m-1)\tau). \tag{10}$$

Fig. 3 shows an example of the relationship between the above–mentioned time t' and

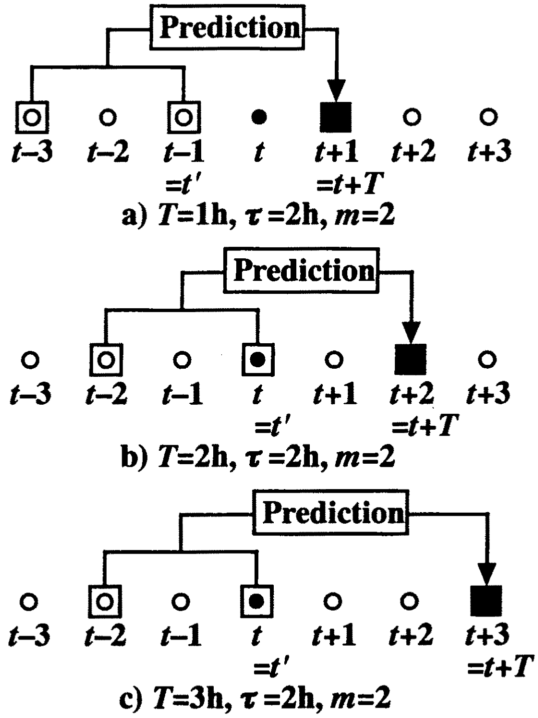

Fig. 3. Schematic diagrams of prediction of the future values using the state space predictor with m=2 and τ=2 h.

the present time t for m=2 and τ=2 h. In the case of T=1 h shown in Fig. 3a), the time (t–1) is put as t' and $x(t+T)$ is predicted by the predictors using $x(t-1)$ and $x(t-3)$ because of the delay time τ=2 h. The same can be said in the case of T=2 h shown in Fig. 3b). On the other hand, in the case of T=3 h shown in Fig. 3c), the time t is put as t' and $x(t+T)$ is predicted using $x(t)$ and $x(t-2)$, because the time $(t+1)$ is the future time and $x(t+1)$ is not known.

In the following, the model structures of the zeroth–order approximation and the first–order approximation are represented by the notations ZOA(m, τ, n_k, T) and FOA(m, τ, n_k, T), respectively.

Fractal dimension

Next, we introduce the fractal dimensions calculated by Higuchi's method to determine the confidence limit when predicting the future behavior of the water–stage time series. The fractal dimension is calculated more precisely, using Higuchi's method, than those presented before. In calculating the fractal dimension, we consider a finite set of the residual time series $\{z(1), z(2), z(3), \cdots, z(n_t)\}$, which was obtained from subtracting long–term periodic components from the time series observations. The number of time steps in the residual time series is n_t. From the given time series, we first constructed a new time series, $Z_t(k)$, defined by:

$$Z_t(k)\ ;z(t),\ z(t+k),z(t+2k),\cdots\cdots,z\left(t+\left[\frac{n_t-t}{k}\right]k\right),\quad t=1,2,\cdots,k. \tag{11}$$

The Gauss' notation is denoted by [ ]. Integers t and k indicate the initial time and the interval time, respectively. For a time interval equal to k, we got k sets of new time series. In the case of k=3 and n_t=100, three time series obtained by the above process are described by:

$$\begin{aligned} &Z_1(3)\ ;z(1),\ z(4),\ z(7),\cdots\cdots,z(97),\ z(100),\\ &Z_2(3)\ ;z(2),\ z(5),\ z(8),\cdots\cdots,z(98),\\ &Z_3(3)\ ;z(3),\ z(6),\ z(9),\cdots\cdots,z(99).\end{aligned} \tag{12}$$

We defined the length of the curve, $Z_t(k)$, in the form:

$$L_t(k)=\frac{1}{k}\left(\sum_{i=1}^{\left[\frac{n_t-t}{k}\right]}|z(t+ik)-z(t+(i-1)k)|\right)\frac{n_t-1}{\left[\frac{n_t-t}{k}\right]k}. \tag{13}$$

The term, $(n_t-1)/([(n_t-t)/k]\cdot k)$ represents the normalization factor for the curve length of subset time series. We defined the length of the curve for the time interval k, $L(k)$, as the average value over k sets of $L_t(k)$.

$$L(k)=\frac{1}{k}\sum_{t=1}^{k}L_t(k). \tag{14}$$

If $L(k)$ is in proportion to k^{-D}, then the curve is fractal with the dimension D. The fractal dimension has a real number ranging from 1 to 2 and expresses how a time series covers a plane. In the case of $D\cong 2$, a time series shows the tendency to vary strongly and the

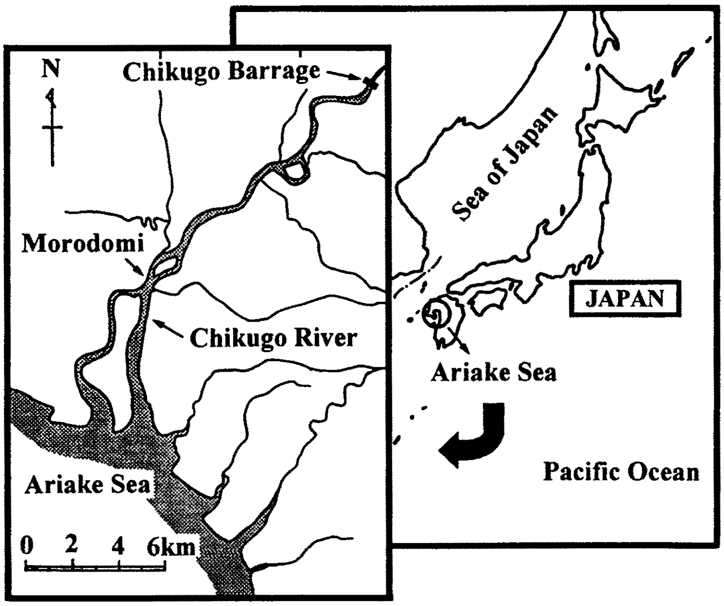

Fig. 4. Outline of lower reach of the Chikugo River.

future behavior of the time series can be hardly predicted. Therefore, we can estimate the maximum lead–time to predict the time series with sufficient accuracy by calculating the fractal dimension (Matsuba, 1992).

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 compartment of the Chikugo River. The lower reach of the Chikugo River is shown in Fig. 4. 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

The state space predictors were applied to the prediction of river water–stages at Morodomi in the Chikugo River shown in Fig. 4. Two sets of hourly water–stage data were selected for model testing. We used the induced mapping obtained from the embedded trajectories of one data set (547 day hourly water–stage data), and the other data set (75 day hourly water–stage data) was predicted by the state space predictors. With the numbers of the nearest neighbors n_k equal to 10, 20, 30, 100, the ZOA(m, τ, n_k, T) and FOA(m, τ, n_k, T) were tested by varying the embedding dimension m from 2 to 7, the delay time τ from 1 h to 7 h and lead–time T from 1 h to 3 h, respectively. Prediction performance was evaluated using the root–mean–square error E_r defined by:

Table 1. The best models that minimize the root–mean–square error E_r(m) for each n_k and T

Approximation	n_k	Lead Time		
		T=1 h	T=2 h	T=3 h
Zeroth–order	10	E_r=0.155 by ZOA(2, 1, 10, 1)	E_r=0.241 by ZOA(6, 4, 10, 2)	E_r=0.241 by ZOA(6, 4, 10, 3)
	20	0.150 ZOA(2, 1, 20, 1)	0.247 ZOA(6, 4, 20, 2)	0.247 ZOA(6, 4, 20, 3)
	30	0.150 ZOA(2, 1, 30, 1)	0.255 ZOA(6, 4, 30, 2)	0.255 ZOA(6, 4, 30, 3)
	100	0.157 ZOA(2, 1, 100, 1)	0.285 ZOA(6, 4, 100, 2)	0.285 ZOA(6, 4, 100, 3)
First–order	10	0.133 FOA(3, 1, 10, 1)	0.241 FOA(5, 2, 10, 2)	0.274 FOA(5, 3, 10, 3)
	20	0.121 FOA(6, 1, 20, 1)	0.176 FOA(7, 2, 20, 2)	0.208 FOA(7, 3, 20, 3)
	30	0.117 FOA(6, 1, 30, 1)	0.165 FOA(7, 2, 30, 2)	0.193 FOA(7, 3, 30, 3)
	100	0.110 FOA(6, 1, 100, 1)	0.170 FOA(7, 2, 100, 2)	0.191 FOA(7, 3, 100, 3)

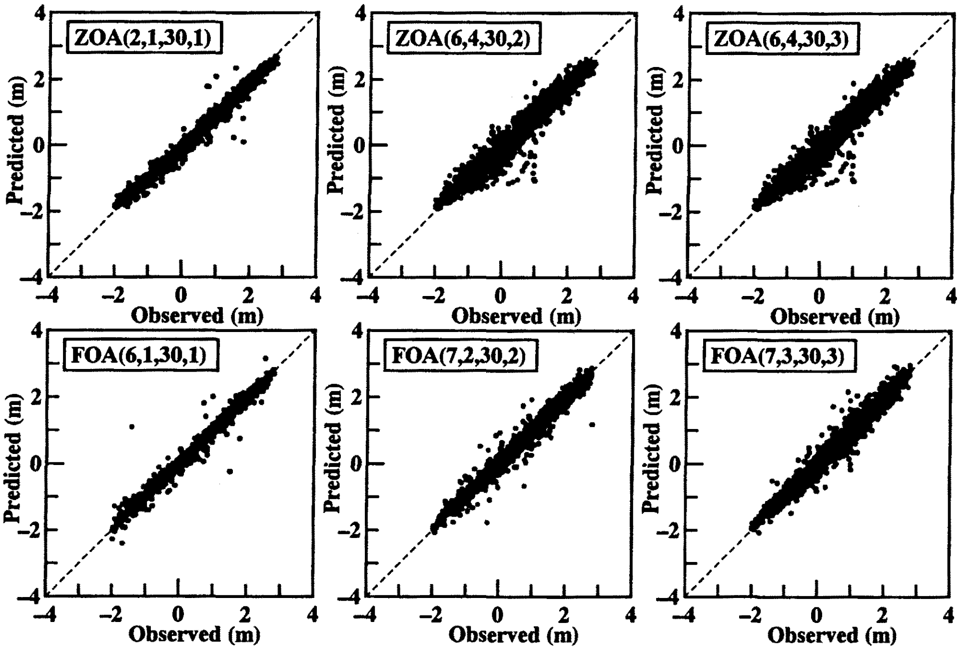

Fig. 5. Scatterplots comparing the predicted and observed water-stages for the ZOA(m, τ, n_k, T) and FOA(m, τ, n_k, T).

$$E_r=\sqrt{\frac{1}{n_t}\sum_{t=1}^{n_t}(x(t+T)-x_p(t+T))^2},\ T=1,2,3. \tag{15}$$

The best models that minimize E_r for each n_k and T, and their E_r are summarized in Table 1. Fig. 5 shows scatterplots comparing the predicted and observed water-stages for the ZOA(m, τ, n_k, T) and FOA(m, τ, n_k, T). Fig. 6a) and Fig. 6b) show the ability of the FOA(m, τ, n_k, T) to match the water-stages during low water and high water seasons, respectively. In Fig. 5 and Fig. 6, the water-stages were predicted by using the best models in n_k equal to 30 shown in Table 1. As shown in Fig. 5, Fig. 6 and Table 1, the ZOA(m, τ, n_k, T) and FOA(m, τ, n_k, T) show good performance in predicting the water-stages both during low water and high water seasons. Table 1 indicates that the E_r-values are almost smaller for the FOA(m, τ, n_k, T) then for the ZOA(m, τ, n_k, T) in each n_k and T. For the data sets used in this study, the necessity would be felt to use a n_k-value of 30, but the optimal n_k-value can be considered to be dependent on the accuracy of water-stage data.

Fig. 7 shows an example of the residual time series obtained from subtracting a M_2 tidal component of the 12.42 h period from the water-stage observations by the moving average method (Suzuki, 1973). The moving average used here is a weighted running mean. In Fig. 8, the logarithm of the length $L(k)$ calculated by equation (14) for the residual time series, is plotted as a function of the logarithm of the time interval k. The straight line is fitted according to the least-square procedure. The graph breaks at the value of k_c=2 h. In the range of the time scale above k_c=2 h, the curve is linear with a slope equal to −2.0, i.e. the fractal dimension D=2.0. In the lower range, a slope of −1.5 is seen.

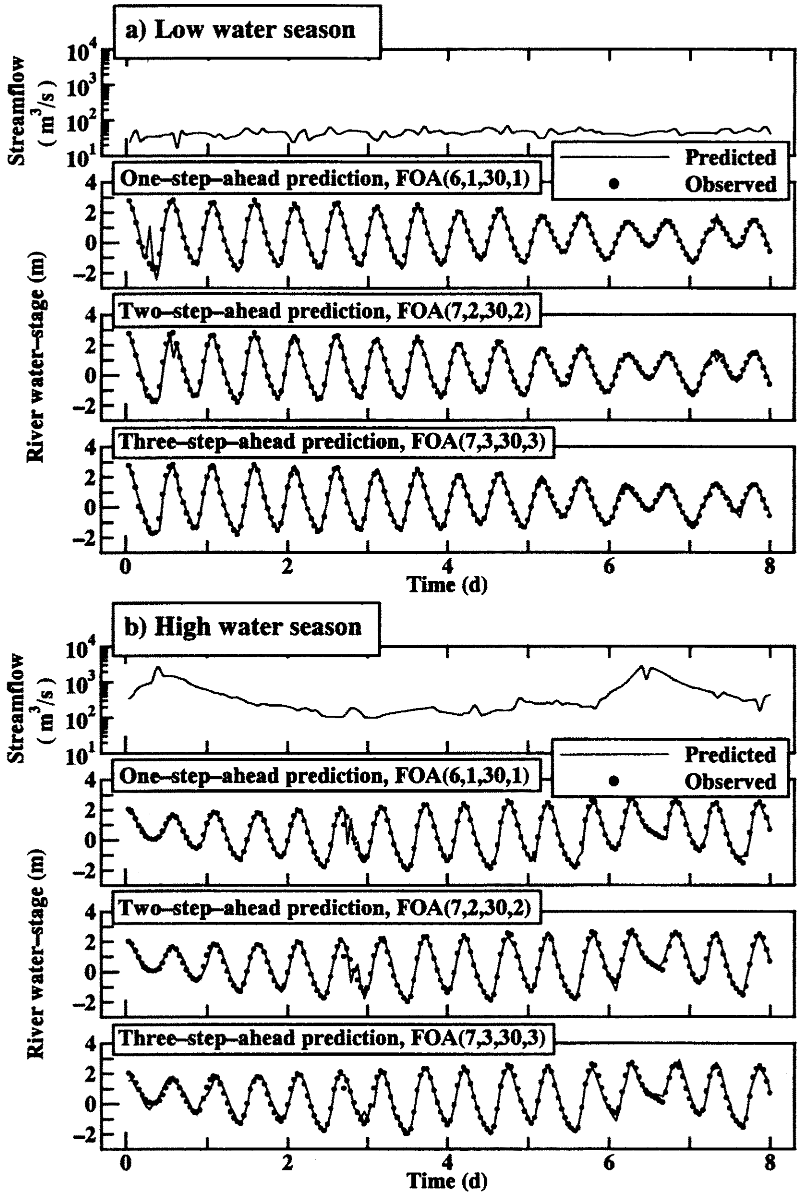

Fig. 6. Comparison of the predicted and observed water-stage time series for the FOA(m, τ, n_k, T) in a) low water season and b) high water season.

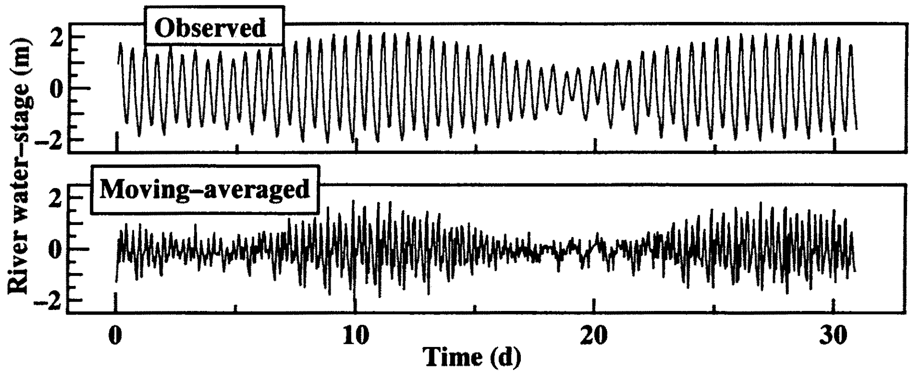

Fig. 7. An example of the observed and moving–averaged water–stage time series.

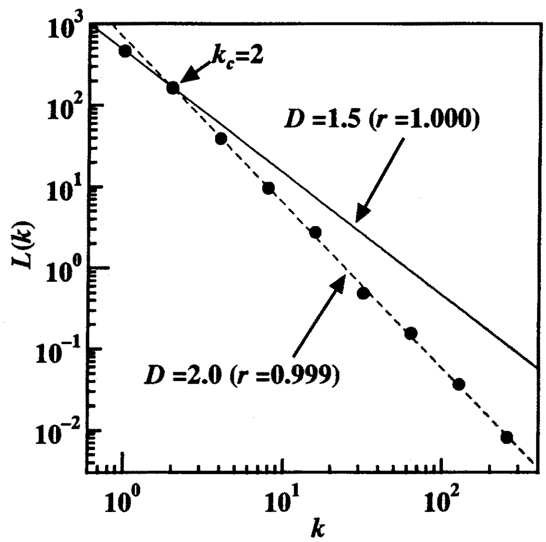

Fig. 8. The fractal dimensions D calculated by using Higuchi's method; r is a correlation coefficient.

As mentioned before, Fig. 8 shows that the confidence limit for predicting the residual time series can be considered 2 h. In a case where the residual time series is predicted, the maximum lead–time of prediction is 2 h. However, in the case of actual prediction, the water–stage time series contains long–term periodic components. By predicting the long–term periodic components along with the residual components using the state space predictors, we can predict the future values with the lead–time of the time scale above k_c=2 h.

CONCLUSIONS

In this paper, the state space predictors and the fractal dimension were used in the prediction of water-stages in a tidal river. It was concluded that the state space predictors were useful tools in the short-term prediction of water-stages that had a strong autocorrelation structure due to tidal motion. The maximum lead-time to predict the future values of the time series with sufficient accuracy was estimated by the fractal dimensions calculated using Higuchi's method.

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

Farmer, J. D. and J. J. Sidorowich 1987 Predicting chaotic time series. *Phys. Rev. Lett.*, **59**: 845–848

Grassberger, P. and I. Procaccia 1983 Measuring the strangeness of strange attractors. *Physica D*, **9**: 189–208

Higuchi, T. 1988 Approach to an irregular time series on the basis of the fractal theory. *Physica D*, **31**: 277–283

Hiramatsu, K., S. Shikasho and K. Mori 1998 The use of genetic algorithm and artificial neural network for the short-term prediction of water-stages in a tidal river. *Applied Hydrology*, **11**: 37–45 (in Japanese)

Lall, U., T. Sangoyomi and H. D. I. Abarbanel 1996 Nonlinear dynamics of the Great Salt Lake: nonparametric short-term forecasting. *Water Resour. Res.*, **32**: 975–985

Matsuba, I. 1992 Chaos and prediction. *Mathematical Sciences*, **348**: 64–69 (in Japanese)

Porporato, A. and L. Ridolfi 1997 Nonlinear analysis of river flow time sequences. *Water Resour. Res.*, **33**: 1353–1367

Suzuki, E. 1973 *Meteorological Statistics*. Chijin-Shokan, Tokyo (in Japanese)