NONLINEAR PREDICTION OF THE WATER FLOW IN AN INTERCONNECTED MULTI-RESERVOIR POWER SYSTEM


* The Institute of Statistical Mathematics, 4-6-7, Minami Azabu, Minato-ku, Tokyo 106, Japan
** IMCC - Universidade Estadual de Campinas, CP6065 - Campinas, Brazil
+ Faculty of Engineering, Osaka University, 2-1 Yama-da-oka, Suita, Osaka 565, Japan

Abstract. The prediction of water flow in a season of heavy rainfall is considered for the purpose of control of an interconnected multi-reservoir power system in Kyushu, Japan. The identification of several nonlinear prediction models is discussed and their prediction performance is examined through numerical studies.

Keywords. Nonlinear prediction, input-output system, dynamical system, stochastic system, time series, time domain analysis, water resources, stochastic hydrology.

1. INTRODUCTION

In order to be able to control an interconnected multi-reservoir power system adequately, it is very important to be able to predict water flow coming into the reservoir from rivers accurately. This is particularly important in times of heavy rainfall, especially in countries which have rainy seasons, such as Japan.

In a multi-reservoir system, water flows into the system from rivers, and also flows between the reservoirs of the system. If a reservoir has a large enough capacity, the water flow out of the reservoir is well controlled and the flow into the system. However, there are often cases where flow into a reservoir comes directly from a river, or where the capacity of the preceding reservoir, if any, in the system is very small. In a rainy season, the water in a small capacity reservoir easily overflows down to the next reservoir in the system.

Since Japan is an island and a large percentage of the land area is covered by mountains, Japanese rivers are relatively short and drop steeply. Because of the mountainous terrain, it is difficult to construct large capacity reservoirs. Many water power systems in Japan use interconnected multi-reservoir systems with fairly small capacity reservoirs. Japan also has two very definite rainy seasons, one from June to mid-July and another in September which is usually accompanied by typhoons. Water management in these periods is crucial for the rice farming around the area below the reservoirs. In such situations it is essential to predict the amount of water flow to a reservoir accurately in rainy seasons. Since the capacity of many reservoirs in the upper part of such systems is often small and water overflows from such small reservoirs in rainy seasons, it is impossible to measure the amount of water flowing from such reservoirs down to the rivers which serve the next reservoirs in the system. However, the water level of these overflowing small reservoirs seems to contain useful information for the prediction of the water flow to following reservoirs, and this can easily be measured.

In the present paper, we introduce some statistical nonlinear models for the prediction of water flow to a dam at the bottom of an interconnected multi-reservoir power system using as inputs measurements of rainfall and the water level of overflowing reservoirs in the upper region of the interconnected multi-reservoir power system.

In section 2, we see a typical example of an interconnected multi-reservoir power system on the Chikugo river in Kyushu, Japan. In section 3, we consider the problem of obtaining a prediction model for real reservoir systems and see why a statistical approach is appropriate. In section 4, we consider statistical nonlinear input-output models for the prediction of water flow into a reservoir (the Yoake dam) which is at the bottom of the Chikugo river interconnected multi-reservoir power system. In section 5, numerical results for the identification of the statistical models are discussed. A brief conclusion is given in section 6.
Fig. 1a Yoake water flow

Fig. 1b Kusu water height

Fig. 1c Ohl water height

Fig. 1d Yuyama water height

Fig. 1e Mannen water height

Fig. 1f Yoake water height

Fig. 1g Onagohata rainfall

Fig. 1h Hoshi rainfall

Fig. 1i Mizobe rainfall

Fig. 1j Mori rainfall

Fig. 1k Nogami rainfall

Fig. 1l Jizukuma rainfall

Fig. 1 Time series record of water flow, water height and rainfall in Chikugo River interconnected multi-reservoir system
The present prediction result is applied to the control of the water level of the Yoake dam in the Chikugo river interconnected multi-reservoir power system, which is also discussed in another paper of ours (Tamura, H. and others, 1988).

2. THE CHIKUGO RIVER INTERCONNECTED MULTI-RESERVOIR SYSTEM

Fig. 2 shows the Chikugo river interconnected multi-reservoir system and the locations where the water level of reservoirs or rainfall are measured. The aim is to predict the water flow into the Yoake dam, which is at the bottom of the multi-reservoir system, in order to be able to control the water level of Yoake dam within a safe level in rainy seasons. The reservoirs on the right hand side of the Kusu river are all large, so that the water of these reservoirs is well under control and hardly overflows even in a very heavy rainy season accompanied by typhoons. On the other hand, the Ohī river on the left hand side of the Kusu river has no reservoir and large amounts of water flow directly into the Yoake dam when it rains. Also the reservoirs on the Kusu river, the Kurusawa reservoir, the Yuyami reservoir and the Mannen reservoir, are all very small and in the rainy season water overflows down the river into the Yoake dam.

In the rainy season, the gate of Yoake dam is left open to prevent overflowing, which could cause serious damage to the power system of the Yoake dam. The amount of water flowing out of the dam could be recorded since the dam is large and was not overflowing. Fig. 1. a shows the time series data of the water flow leaving the Yoake dam. If we compare the water flow with the rainfall and the water level of the upper reservoirs in Fig. 1, we can see that the amount of water leaving the Yoake dam increases and decreases depending on the water level of the upper reservoirs and the rainfall. Since the water flow of Yoake dam is completely open and no human intervention is applied, the amount of water leaving the Yoake dam may be considered to be approximately equivalent to the amount of water flowing into the dam. This approximate equivalence of water flowing in and out of Yoake dam helps a great deal in tackling the very difficult problem of identifying a prediction model for a real hydrology system.

Time series data of the water height of the Yoake dam is also available. However, this data contains only approximate information about the amount of water flowing into the dam because the area of the dam is , unlike the upper small reservoirs, very big. To record the water flow into the dam accurately, we would need to be able to measure the height using microcentimeters as units, which is impossible.

3. WHY TAKE A STATISTICAL APPROACH?

The calculation of the discharge (water flow) from rivers is often studied by a deterministic approach (Sugawara, 1962). A simple example of a water discharge system is given, as shown in Fig. 3, by a tank with a hole at the bottom. The discharge, \( x \), from the hole is proportional to the height \( z \) of the water in the tank as \( x = cz \), where \( c \) is some constant. The dynamics of \( z \), i.e. the rate of change of the height \( z \), will be described by

\[
\dot{z} = -az
\]

where \( a \) is a positive constant.

Fig. 3 A simple tank model.

If more water \( y \) is added to the tank, the height of the water in the tank will increase in proportion to \( y \) and the equation of dynamics of the water height will be given by

\[
z = -az + by
\]

where \( a \) and \( b \) are constants. Then the discharge \( x \) may be calculated from the
amount of input water $y$ using the state-space representation model,

$$ z = -az + by $$

$$ x = cz $$

(1)

where $z$ is the state variable.

For a more complex river discharge system, the discharge prediction problem may be studied fairly satisfactorily in the framework of deterministic system theory, provided we know precisely how the rainfall flows from the ground to the reservoirs, and how water flows from reservoir to reservoir. In other words, we must know the impulse response function of the system. In this case, we could identify a discrete time multi-dimensional state-space representation model,

$$ z(t+1) = A(t)z(t) + By(t) $$

$$ x(t) = Cz(t) + w(t) $$

(2)

where $w(t)$ is an observation error, $z(t)$ is a state vector and $y(t)$ is an input vector and $A$, $B$ and $C$ are matrices.

However, a real riverflow system is very complex indeed, and it is impossible to know the exact flow pattern from rainfall to reservoir and between reservoirs. Since in this case, the impulse response function is unknown, deterministic system theory cannot be used to identify a model of form (1). It is essential to use statistical methods to identify an appropriate model on the basis of data samples of rainfall and water levels of rivers and reservoirs. As we saw in Fig. 1, the time series data which seems to be most relevant as input for the Yokae dam water flow system is not the amount of water flowing into rivers entering the Yokae dam, as one might expect, but the water level of such rivers or of overflowing reservoirs above the Yokae dam. In fact, in practice, transformations of these data may be more appropriate as input to the system.

For the prediction of stochastic Gaussian time series, a stochastic realization of the linear state space model (2) is available (Akaishi, 1974a). However, as we see in Fig. 2, the most hydrological time series have a highly asymmetric and non-Gaussian character. What is needed in this situation is a nonlinear model which approximates the stochastic behavior of the Yokae dam as closely as possible, using relevant input information such as that given in Fig. 1.

If we are trying to approximate some unknown probabilistic structure with density $g(.)$ by an estimated statistical model $f(.)$, then the most natural measure of goodness-of-fit is the entropy function

$$ I(f, g) = -\int f(x) \log \frac{f(x)}{g(x)} dx $$

The best statistical model is that which maximizes the entropy function $I(f, g)$. An unbiased estimate of $I(f, g)$, the measure of the entropy, is given by AIC (Akaishi, 1974), where

$AIC = \frac{1}{2} \log(\text{maximum likelihood}) - \frac{1}{2} \text{(number of parameters)}$,

which can be used to compare the goodness-of-fit of several candidate statistical models for the time prediction of the discharge of the Yokae dam, the best model being that which has minimum AIC value.

Using AIC, we can try to extend the basic linear model, which can easily be estimated using a standard method, to a nonlinear model by introducing various types of nonlinearity and examining how much the fit of the model has been improved. However, an obvious problem with this kind of statistical approach is that there can be too many different possible nonlinear models, and it is unrealistic and almost impossible to calculate all the possible candidate nonlinear models and compare them. Therefore, in the following section, we demonstrate some natural methods of narrowing the set of candidate nonlinear models for the prediction of water flow.

4. NONLINEAR MODELING

Let $p(t)$ be the time-averaged amount of water flow into the Yokae dam at time $t$. $p(t)$ is measured at 15 minute time intervals, and represents the average water flow over the past 15 minute interval. For a positive-valued stochastic process like $p(t)$, it is often the case that the innovation variance is greater for large values of $p(t)$ than it is for small values of $p(t)$. Therefore, it is common practice to transform the water flow $p(t)$ by either a square root or logarithmic transformation before considering a dynamic model. This type of transformation has also been justified for the identification of non-Gaussian stochastic processes (Okaishi, 1985). In the present study, we take a square root transformation of $p(t)$ and try to identify a model for $Q(t) = \sqrt{p(t)}$ by statistical methods, using $p(t)$, the water level of the Kusugawa reservoir, $p(t)$, the water level in the Ohtsu river, and the 15 minute amount of rainfall at Onogohara (see Fig. 1).

A general nonlinear prediction model for $Q(t)$ may be written as

$$ Q(t) = f(p(t-1), \ldots, p(t-1), \ldots, p(t-1), \ldots, \eta(t), \ldots, t=1, 2, \ldots, N \ldots (3) $$

where $N$ is the number of observations, and $\eta(t)$ is an innovation term which can be considered to be the prediction error of $Q(t)$. The problem is to find $f(.)$. One approach would be to consider polynomial families as suggested by Wiener (1949) and try to find the best combination of finite polynomial terms using AIC. However, the number of possible combinations of polynomial terms is huge, and it is not realistic to calculate the likelihood and compare AICs for so many models. It is necessary to reduce the number of potential models by introducing a reasonable family of nonlinear models, based on physical considerations.
Instantaneous transformation of inputs

One idea for simplifying (3) could be to write

\[ Q(t) = a_0 f_1(p(t-s_1)) + b_0 f_2(p(t-s_2)) + c_0 f_3(p(t-s_3)) + \eta(t) \] (4)

where \( f_1(\cdot), f_2(\cdot), \) and \( f_3(\cdot) \) are polynomials of the form

\[ f_1(z) = a_1 z^2 + ... + a_1 z^{s_1} \] (5)
\[ f_2(z) = b_1 z^2 + ... + b_1 z^{s_2} \] (6)
\[ f_3(z) = c_1 z^2 + ... + c_1 z^{s_3} \] (7)

Maximum likelihood estimation of these parameters, which is asymptotically equivalent to least squares estimation, may easily be performed by solving a set of linear equations. The polynomial orders \( l, m, n \) and lag orders \( s_1, s_2 \) and \( s_3 \) are chosen by AIC.

Linear system

Another simplification is to take several lags for each of the input variables \( P_1(t), P_2(t), P_3(t) \) without using a nonlinear transformation. The model can then be written

\[ Q(t) = a_1 f_1(p(t-s_1)) + ... + a_1 f_1(p(t-l_1)) + b_1 f_2(p(t-s_2)) + ... + b_1 f_2(p(t-l_2)) + c_1 f_3(p(t-s_3)) + ... + c_1 f_3(p(t-l_3)) + \eta(t) \] (8)

Here, the identification problem is reduced to the determination of the lags \( s_1, s_2, s_3 \) and the maximum orders \( l_1, l_2 \) and \( l_3 \). Even if we restrict the number of lags to, say, 2, there will still be too many possible combinations for realistic comparison.

It is well-known in system theory that state-space modeling has a smaller number of parameters and is more appropriate for a linear system of type (8). Suppose \( Q(t) \) is observed from a state-space model as follows,

\[ \begin{align*}
    \dot{x}(t) &= A x(t-1) + B u(t) \\
    Q(t) &= C x(t)
\end{align*} \]

where the state vector \( x(t) \) has dimension \( k \), \( p(t) \) is a three dimensional input vector \( p(t) = [p(t-s_1), p(t-s_2), p(t-s_3)] \)' and \( A, B \) and \( C \) are matrices of dimension \( k \times k, k \times 3 \) and \( 3 \times k \) respectively. Then from the state-space representation we obtain,

\[ Q(t) = \phi(1) Q(t-1) + \phi(2) Q(t-2) + ... + \phi(k) Q(t-k) + C (A^{k-1} \phi(1) A^{k-2} + ... \phi(k-1) A + \phi(k) I) B u(t-k+1) + C (A^{k-1} \phi(1) A^{k-2} + ... \phi(k-1) A + \phi(k) I) B u(t-k+2) + ... + C \phi(1) + C B g(t) \]

where \( \phi(1), \phi(2), ..., \phi(k) \) are the coefficients of the characteristic equation,

\[ A^k - \phi(1) A^{k-1} - ... - \phi(k-1) A - \phi(k) I = 0 \]
of the matrix \( A \). Using this state-space modeling idea, we can consider the following type of model,

\[ Q(t) = \phi(1) Q(t-1) + ... + \phi(k) Q(t-k) + \psi(1)p(t-s_1) + ... + \psi(k)p(t-s_1-k+1) + \xi(1)p(t-s_2) + ... + \xi(k)p(t-s_2-k+1) + \eta(t) \] (9)

where \( \phi(1), \psi(1), ... \xi(1), ... \eta(1) \) are scaled parameters. Model (9) can be expressed by means of the backward shift operator \( z^{-1} \) as follows,

\[ Q(t) = \psi(z^{-1}) \phi(z^{-1}) p(t-s_1) + \theta(z^{-1}) \phi(z^{-1}) p(t-s_2) + \pi(z^{-1}) \phi(z^{-1}) p(t-s_3) + \eta(t) \] (10)

where \( \theta(z^{-1}) = 1 - \psi(z^{-1}) - ... - \psi(k) z^{-k} \)
\( \pi(z^{-1}) = \psi(1) + \psi(2) z^{-1} + ... + \psi(k) z^{-k+1} \)
\( \phi(z^{-1}) = \theta(1) + \theta(2) z^{-1} + ... + \theta(k) z^{-k+1} \)
\( \pi(z^{-1}) = \pi(1) + \pi(2) z^{-1} + ... + \pi(k) z^{-k+1} \)

The identification problem now becomes the determination of the system order \( k \) and the lags \( s_1, s_2, s_3 \), and \( s_4 \) for each input variable. The possible lags of some of the input variables are restricted by the geographical features of the area to a fairly small range, so that the real problem of identification is the determination of the system order \( k \). For fixed \( s_1, s_2, s_3 \), and \( s_4 \), the least squares estimation of the model parameters leads to solving a set of linear equations, which is computationally simple. The number of model parameters to be estimated altogether, including the residual variance, is \( 4k+1 \).

System nonlinearity

We can extend the basic linear system model (10) into a nonlinear system model by incorporating some kind of system nonlinearity. It is shown in Oishi (1985) that the damping characteristics of riverflow are nonlinear; the water level decreases quickly when the water level is high and decreases slowly when the water level is low. This kind of system nonlinearity may be described by the following nonlinear dynamical system with input \( p(t) \),

\[ \dot{Q} = -a_0 \dot{Q} + b p \]

For \( Q(t) > 1 \), \( Q(t) \) for this nonlinear model decreases faster than \( Q(t) \) represented by the linear model

\[ \dot{Q} = -a_0 \dot{Q} + b p \]

but for \( Q(t) < 1 \), \( Q(t) \) decreases slower in the nonlinear case. It is shown in Oishi (1985) that this kind of nonlinearity may be incorporated in a discrete time model with \( Q(t) \)-dependent coefficients \( \phi_{k-1} \) and \( \phi_{k-1} \) thus

\[ Q(t) = \phi_{k-1} Q(t-1) + \theta_{k-1} p(t-1) \]
A reasonable parameterization of $\theta_{e-1}$ and $\theta_{s-1}$ is a continuous function which approaches a constant as $Q(t-1) \to \infty$ (Ozaki, 1985). One example is

$$
\phi_{e-1} = a + b \exp(-\gamma Q(t-1)^2)
$$

which gives

$$
\phi_{e-1}(t) = a + b \exp(-\gamma Q(t-1)^2)
$$

$\gamma$ is a constant which may be adjusted by taking into account the scale of $Q(t)$.

This kind of system nonlinearity is easily extended to a more complicated model of high order, such as model (10) by making the coefficients $\phi(1), \psi(1), \theta(1)$ and $\xi(1)$ of the polynomials $\Phi(\cdot), \Psi(\cdot), \Theta(\cdot)$ and $\Xi(\cdot)$ functions of $Q(t-1)$ as follows

$$
\phi_{e-1}(1) = a_1(1) + b_1(1) \exp(-\gamma Q(t-1)^2)
$$

$$
\psi_{e-1}(1) = a_2(1) + b_2(1) \exp(-\gamma Q(t-1)^2)
$$

and

$$
\theta_{s-1}(1) = a_3(1) + b_3(1) \exp(-\gamma Q(t-1)^2)
$$

where $a_j(1), b_j(1), j=1, \ldots, 4, i=1, \ldots, k$ are constants.

The model may then be expressed as follows,

$$
Q(t) = \phi_{e-1}(1)(Q(t-1) + \cdots + \phi_{e-1}(k)Q(t-k)) + \psi_{e-1}(1)(p(t-s_1) + \cdots + \psi_{e-1}(k)p(t-s_1-k+1)) + \theta_{s-1}(1)(p(t-s_2) + \cdots + \theta_{s-1}(k)p(t-s_2-k+1)) + \xi_{s-1}(1)(p(t-s_3) + \cdots + \xi_{s-1}(k)p(t-s_3-k+1)) + \eta(t)
$$

$\gamma$ is a positive constant appropriately chosen so that for a small $\varepsilon > 0$ and for $Q_0 = \max(Q(t))$, $\exp(-\gamma Q^2) = \varepsilon$.

Then $\exp(-\gamma Q(t-1)^2) > \varepsilon$ for any $Q(t-1)$. $\gamma$ is fixed to an appropriate value before we apply the least squares estimation method to the model. The number of model parameters to be estimated, including the residual variance, is $8k+1$.

Continuous time nonlinear modelling

We can also fit a continuous time nonlinear dynamical system model,

$$
\dot{Q} = f(Q) + b_1 p_1(t-s_1) + b_2 p_2(t-s_2) + \cdots + b_8 p_8(t-s_8)
$$

by using a maximum likelihood method. Ozaki (1985) introduced the following discrete time model derived from (12),

$$
Q(t+1) = \Lambda_0 Q(t) + B_1 b_1 p_1(t-s_1) + B_2 b_2 p_2(t-s_2) + B_8 b_8 p_8(t-s_8) + C_0 \eta(t-1)
$$

where,

$$
\begin{align*}
A_0 &= e^{X_1 \gamma} \\
B_1 &= (e^{X_1 \gamma} - 1)/K_1 \\
C_0 &= \sqrt{(e^{X_1 \gamma} - 1)/(2K_1)} \\
K_1 &= (1/5) \log(1 + (e^{X_1 \gamma} - 1)f(Q(t))/Q(t)) \\
J_0 &= \frac{2f(Q)}{\sigma_0} \int Q(t) \\
\end{align*}
$$

and $\delta$ is a properly chosen sampling interval. The model (13) is one kind of first order nonlinear input-output system parameterized by $a_1, a_2, a_3, b_1, b_2$ and $b_3$ in the continuous time domain. We can estimate these parameters by the maximum likelihood method. The order of the polynomial and the lags $s_1, s_2$ and $s_3$ are chosen by AIC.

5. NUMERICAL RESULTS

We applied model (11) from the previous section to the data of the Chikugo river interconnected multi-reservoir system. The minimum AIC procedure chose the model of order $n=5$ with lag orders $s_1=2$, $s_2=2$ and $s_3=1$. The estimated residual variance was $\sigma = 0.1850$ and AIC = 2259.9. The prediction performance of the identified model can be checked by observing the residuals of the model. The critical time for prediction is when the water flow increases quickly after heavy rain. A typical pattern can be seen around the 14th day from the start of the study.

**Fig. 4. Prediction using model (11)**

**Fig. 5. Prediction using model (4)**
If we use model (4) which is a linear combination of instantaneous transformations of the past observations of the input series, $p_1(t)$, $p_2(t)$, and $p_3(t)$, we obtain a model with lags $s_1=3$, $s_2=7$ and $s_3=25$ and polynomials of order 3, i.e. $l=2$, $m=2$ and $n=2$ in (5), (6) and (7) respectively. The AIC for the model is $-4101.4$ and the estimated residual variance is $\sigma^2 = 4.5819$. Fig. 5 shows the prediction performance of this model over the same period as Fig. 4. Clearly, prediction using the nonlinear dynamic model (11) seems to be much better than the instantaneous transformation model (4).

Although the prediction performance of model (4) is much inferior to model (11), the model gives us useful information about the variables of the system. The coefficients $a_0$, $b_0$, and $c_0$ of $f_1(p_1(t-3))$, $f_2(p_2(t-7))$, and $f_3(p_3(t-25))$ of the model, show the weight of the contribution from these inputs to $Q(t)$, the water flow to the Yoake dam. The ratio between $s_2$, $b_0$, and $c_0$ is 70, 10 and 1. This means the water flow pattern to the Yoake dam depends mainly on the water flow from the reservoir. The estimated polynomial of $f_1(p_1(t-3))$ for the range of given data values is shown in Fig. 6. Of course, it is impossible to estimate the form of $f_1(p_1(t-3))$ for values outside this range.

A structurally simple model often gives us useful information in understanding the relationships between variables, while a general nonlinear modelling often brings us a model which gives better prediction performance but is difficult to interpret physically. In this case, our identified model (11) also contains some structural information about the Yoake dam prediction system.

The coefficients $\phi_1(i), \ldots, \phi_n(k)$ of the model characterize the system in the same way as the transition matrix of a state-space model does. If we calculate the roots of the following characteristic equation,

$$\lambda^n - \phi_1(1)\lambda^{n-1} - \cdots - \phi_n(k) = 0$$

for each $0 < Q(t) < \max(Q(t))$, we can see how the system dynamics change depending on the level of the water flow $Q(t)$. The five characteristic roots of the equation are plotted in Fig. 7, where $x$ with a circle shows the roots when $Q=0$ and $x$ with a triangle shows the roots when $Q=\max(Q(t))$. All the roots are inside the unit circle, which means that the model is stable for the region of $Q(t)$. If we assume that $\phi_1(i), \ldots, \phi_n(k)$ stay constant for $Q(t) > Q_0$, and if the expectations of the input variables $p_1(t), p_2(t)$ and $p_3(t)$ are finite, we can say that the stochastic process $P(t)$ defined by the estimated nonlinear model is a stationary Markov process (Orzuki, 1985).

One of the characteristic roots for $Q=0$ is very close to the unit circle and shifts closer to the center of the unit circle for larger $Q$. This means the damping speed of the water flow is small for small $Q(t)$ but is larger for large $Q(t)$ so that $Q(t)$ decreases more quickly when $Q(t)$ is large.

When we fit a linear model (9) of order $k=5$, we have a model whose characteristic roots are constant as shown in Fig. 8. If we compare them with the dynamic roots in Fig. 7 we can see that the main feature of system dynamics contained in the data is characterized by these constant characteristic roots in the linear model.

When we fit models of types (9) and (11) with order $k=5$, we obtain an AIC value of $-429.5$ and estimated residual variance $\sigma^2 = 2.002$ for model (9) and AIC value of $-435.7$ and $\sigma^2 = 0.1943$ for model (11). The AIC values show that these models are a much better fit than model (9) but not so good as model (11) with order $k=5$. The three characteristic roots of these models are shown in Fig. 9 and Fig. 10. Comparing Figs. 7, 8, 9 and 10, we notice that the...
real root near the unity consistently stays in the same position, while the other roots are different for each different model.

Fig. 9 The characteristic roots for the estimated model (9) of order k=3

Fig. 10 The characteristic roots for the estimated model (11) of order k=3

Fig. 11, which shows the prediction performance of the estimated model of type (12). The estimated nonlinear function f(Q) is

\[ f(Q) = 0.79795Q + 0.50733Q^2 - 0.21286Q^3 \] \hspace{1cm} (14)

which is shown in Fig. 12. The estimated residual variance of this model is \( \sigma^2 = 11.3212 \) and the AIC value is 6.532.9.

This rather large value of AIC also means that a model with only one system order is not good for prediction when the water flow is increasing as on the 14-th day of Yoake dam water flow, even though the dynamics of the model are made nonlinear. Certainly, the prediction performance will be improved by using a continuous time system model of higher order by making the first order differential equation in the model (12) into a second order differential equation. Statistical identification methods for such models are given in Oda, Ozaki and Yamamoto (1987) with numerical applications. Hence, apply those methods we need velocity data of the water flow dynamics, which are not available in this case.

Model (11) of order k=3 gives the minimum AIC value and the best prediction performance so far. However, if we look carefully at the detail of the residuals of the fitted model of (11), we see that residuals in some periods are much larger than those in other periods, which implies that prediction errors are much bigger for those periods. One reason for this may be because the variance of the system noise is still not constant, being larger for the period where \( Q(t) \) is large and smaller for the period where \( Q(t) \) is small. This suggests that we may be able to improve the prediction model by introducing more sophisticated variable transformation than the square root transformation we employed.

Further, some large prediction errors may occur because we don't have enough information about the water flow in the areas surrounding the places from which our data came. Considering and designing the best method of sampling the rainfall and water height around the reservoirs may also be an interesting future research topic.

6. CONCLUSION

The use of the statistical modeling approach for the identification of a prediction model for water flow is shown to be useful in a real interconnected multi-reservoir power system. It is shown that the statistical modeling approach is extendible to nonlinear modelling if a properly chosen nonlinear model family is used. By using the identified nonlinear model, we can achieve better prediction results and also we can characterize some kind of system nonlinearity in the hydrology system.

REFERENCES


ACKNOWLEDGEMENT

The authors thank Dr. V. Haggan Ozaki, for her helpful comments on the present paper. The authors also thank Dr. Y.H. Tamura for his help in editing the time series data used in the paper. The present research was supported by a cooperative research grant of the Institute of Statistical Mathematics, Tokyo.