
The University of Manchester
Institute of Science and Technology

"ON A MODEL OF NON-LINEAR FEED-BACK SYSTEM FOR RIVERFLOW PREDICTION"

by

T. OZAKI

Department of Mathematics
(Statistics)

Technical Report No. 84
May, 1978
ON A MODEL OF NON-LINEAR FEED-BACK SYSTEM FOR RIVERFLOW PREDICTION

T. OZAKI

(University of Manchester Institute of Science and Technology)

SUMMARY

A non-linear system with feed-back is proposed as a dynamic model for the hydrological system, whose input is the rainfall and whose output is the discharge of riverflow. Parameters and orders of the model are estimated using Akaike's Information Criterion. Its application to the prediction of daily discharges of Kanna River and Bird Creek is discussed.

Keywords: NON-LINEAR SYSTEM; PREDICTION; HYDROLOGICAL SYSTEM; IDENTIFICATION; AKAIKE'S INFORMATION CRITERION.
1. INTRODUCTION

Although various stochastic models for riverflow time series have been introduced (See e.g. Lawrence et al (1977)), many of them discussed only the univariate riverflow time series. However it is obvious that the typical behaviour of sharp ascension in the riverflow time series, called the Noah effect, is preceded by a heavy rainfall. It is also well-known that the input-output relation of the rainfall and the riverflow is very non-linear. For small rainfalls such as 5 to 10 mm/day the river stage usually does not rise at all, while a heavy rainfall such as 200 mm/day can cause flooding twice as or much seriously than the case of 100 mm/day (See Sugawara (1961)). Therefore it is important to incorporate the rainfall time series in the modelling of riverflow time series for the prediction of the latter.

As a measure of the goodness of fit of the model, a most reasonable choice is the prediction performance of the model. On the basis of this choice, Sugawara ((1961), (1967), (1974), (1975)) has devised a tank model for the prediction of riverflow, using the rainfall data, and successfully applied it to the prediction of many riverflow time series.

A simple tank model is shown in Fig. 1 which shows a model of the Bikin River given by Sugawara (1974). In this figure, the input rainfall is put into the top tank $T_{11}$. The stored water in the tank $T_{11}$ comes out of the bottom exit $E_1$ of the tank and flows into the tank $T_{21}$. The stored water in the tank $T_{11}$ also comes out from the side exit $E_2$ and flows into the tank $T_{42}$. If the height of the second water in the tank
$T_{11}$ exceeds the height of the second side exit $E_3$, the water comes out from the exit and flows into the tank $T_{h2}$. Finally, the riverflow discharge is the amount of the output of the tank $T_{h2}$. A general tank model consists of series of tanks $T_{ij}$ ($i=1,\ldots,m$, $j=1,\ldots,n$) in two directions, vertical and horizontal, and each tank may have more than one exit, such as shown in the tank $T_{11}$ of Fig. 1. This series storage structure, described by such parameters of the tank as the heights and widths of the exits, is determined by a try and error method. Fig. 2 shows a typical tank model given by Sugawara (1974) for Bird Creek.

Fig. 2 should be inserted here.

Although the tank model has been applied successfully to many rivers by Sugawara and his co-workers, it requires some expertise to get a good model, because the structure and parameters of Sugawara's tank model must be determined subjectively by the analyst.

In this paper, we present a mathematical formulation of Sugawara's tank model and propose a non-linear feed-back system as a discrete version of Sugawara's model. In addition, a statistical identification procedure of the model using Akaike's Information Criterion is described. By using this criterion, the model is identified automatically without too much subjective judgement nor analysts' expertise.

2. NON-LINEAR MODELS

Let $Q_t$ and $P_t$ be a riverflow time series and a rainfall time series respectively. The $Q_t$ can be regarded as an output of a dynamic system whose input is the rainfall $P_t$. Our non-linear dynamic system model for riverflow predictions is given by
\[ Q_t = f(z_{t-1}) + \varepsilon_t \]  \hspace{1cm} (2.1)

where
\[ z_{t-1} = \phi_1 p_{t-1} + \ldots + \phi_p p_{t-p} + \theta_1 q_{t-1} + \ldots + \theta_q q_{t-q}, \]  \hspace{1cm} (2.2)

\( f(\cdot) \) is a non-linear function, usually specified by polynomials, and \( \varepsilon_t \) is a white noise assumed to be distributed as \( N(0,\sigma^2) \).

Our model is introduced basing on the following observation of Sugawara's model. One of the main characteristics of Sugawara's model is its non-linear structure which is graphically shown in Fig. 3 and Fig. 4.

---

Fig. 3 and Fig. 4 should be inserted here.

When the tank has three exits as in Fig. 3, the output discharge shows a piece-wise linear increase as a function of the height of the stored water as in Fig. 4. Thus the multi-exit tank of Sugawara's model can be regarded as a piece-wise linear approximation of a non-linear function of the height, while our model uses a polynomial approximation to the function. On the other hand, the height of the water in the tank increases linearly with the input to the tank and decreases linearly with the output from the tank. This shows that the "intermediate variable" \( z_t \) is determined by a linear combination of the past input and the past output as is in (2.2).

The model (2.1) and (2.2) can be regarded as a feed-back system which is characterized by the "state variable" \( z_{t-1} \) and the non-linear function \( f(\cdot) \) of the state variable (See Fig. 5). When \( f(\cdot) \) is a linear function,

---

Fig. 5 should be inserted here.

the model is the ordinary linear feed-back system.
When we employ a polynomial function

\[ f(z) = a_1 z + a_2 z^2 + \cdots + a_r z^r \]

our model structure is specified by the orders \( p, q \) and \( r \), and the parameters \( \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q, a_1, \ldots, a_r \) of the model are estimated by the Maximum Likelihood Method, which will be discussed in the next section.

3. IDENTIFICATION OF THE MODEL

To identify the model, we must estimate both the orders \( p, q \) and \( r \) of the model and the coefficients \( \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q, a_1, \ldots, a_r \). For this purpose, Akaike's Information Criterion (AIC) is used (Akaike (1973)). AIC is defined by

\[ AIC = (-2) \log \text{(maximum likelihood)} + 2 \text{(number of independent parameters)} \quad (3.1) \]

It is introduced as a consistent and unbiased estimate of the expected neg-entropy \( E_x [-B(g; f(\cdot | \hat{\theta}))] \) where

\[ B(g; f(\cdot | \hat{\theta})) = \int g(y) \log \left( \frac{g(y)}{f(y | \hat{\theta})} \right) \, dy, \quad (3.2) \]

and \( f(\cdot | \hat{\theta}) \) is the distribution defined by the model, whose parameter \( \hat{\theta} \) is estimated from the sample \( x_1, \ldots, x_N \), and \( g(\cdot) \) is the true distribution. The neg-entropy \( -B(g; f(\cdot | \hat{\theta})) \) is a measure of distance between the true model \( g(\cdot) \) and the estimated model \( f(\cdot | \hat{\theta}) \). Consequently, the model which gives the minimum value of AIC is employed as a best approximate model in the sense of entropy maximization (Akaike (1977)). The identification procedure is called Minimum AIC Estimation (MAICE) procedure and has been applied to various statistical modelling problems (Akaike (1972), Tanabe

To use the AIC for the identification of our model, we need the maximum likelihood estimation of the parameters. Suppose the input data \(x_1, \ldots, x_N\) are given and let the likelihood of the output data \(y_1, \ldots, y_N\) be

\[
L = f(y_1, \ldots, y_N \mid x_1, \ldots, x_N, \phi, \theta, a)
\]

where \(\phi = (\phi_1, \ldots, \phi_p)\), \(\theta = (\theta_1, \ldots, \theta_q)\) and \(a = (a_1, \ldots, a_r)\), then

\[
\log f(y_1, \ldots, y_N \mid x_1, \ldots, x_N, \phi, \theta, a)
= \log f(y_1, \ldots, y_q \mid x_1, \ldots, x_N, \phi, \theta, a)
+ \log f(y_{q+1}, \ldots, y_N \mid x_1, \ldots, x_N, y_1, \ldots, y_q, \phi, \theta, a)
\]

\[(3.3)\]

Since the transformation from \((y_{q+1}, \ldots, y_N)\) to \((\epsilon_{q+1}, \ldots, \epsilon_N)\) has unit Jacobian, we have

\[
\log f(y_{q+1}, \ldots, y_N \mid x_1, \ldots, x_N, y_1, \ldots, y_q, \phi, \theta, a)
= \log g(\epsilon_{q+1}, \ldots, \epsilon_N \mid x_1, \ldots, x_N, y_1, \ldots, y_q, \phi, \theta, a)
\]

\[(3.4)\]

where \(g(\epsilon_{q+1}, \ldots, \epsilon_N \mid x_1, \ldots, x_N, y_1, \ldots, \phi, \theta, a)\) is the likelihood of \(\epsilon_{q+1}, \ldots, \epsilon_N\) given \(x_1, \ldots, x_N, y_1, \ldots, y_q, \phi, \theta, a\). As \(\epsilon_t\) is assumed to be Gaussian white noise, we have

\[
(-2) L = (-2) \log f(y_1, \ldots, y_q \mid x_1, \ldots, x_N, \phi, \theta, a)
+ (n-q) \log 2\pi \sigma^2 + \sum_{t=q+1}^{N} \epsilon_t^2/(N-q) + (N-q)
\]

\[(3.5)\]
When the data length \( N \) is sufficiently large the first term of the right hand side of (3.5) can be ignored and the maximization of the likelihood \( L \) reduces to the minimization of \( \sum_{t=q+1}^{N} e_t^2/(N-q) \). Let the minimized value of \( \sum_{t=q+1}^{N} e_t^2/(N-q) \) be \( \hat{\sigma}^2 \), then the approximate (-2) log maximum likelihood is given by

\[
(N-q) \log 2\pi + (N-q) + (N-q) \log \hat{\sigma}^2
\]  

(3.6)

Thus the approximate maximum likelihood estimates are equivalent to the least squares estimates. Using the above approximate likelihood, AIC of the model

\[
y_t = a_1 z_{t-1} + \ldots + a_r z_{t-1}^r + \varepsilon_t \\
z_{t-1} = \phi_1 y_{t-1} + \ldots + \phi_p y_{t-p} + \theta_1 y_{t-1} + \ldots + \theta_q y_{t-q},
\]

(3.7)

which we abbreviate as AIC\((p,q,r)\), is given by

\[
(N-q) \log 2\pi + (N-q) + (N-q) \log \hat{\sigma}^2 + 2(p+q+r+1).
\]

To compare the AIC values of each model \((p,q,r)\) within the set, \( S \), of candidate models, it is necessary to base our information on the same sample size of the innovations. The constant terms common to each model can be ignored. From these observations, the following definition of AIC is employed in the present paper;

\[
AIC(p,q,r) = (N-q_o) \log \hat{\sigma}^2 + 2(p+q+r+1)
\]

where \( q_o = \max (q) \) and \( \hat{\sigma}^2 = \sum_{t=q_o+1}^{N} (y_t - f(z_{t-1}))^2/(N-q_o) \).

To get the (approximate) maximum likelihood estimates, we minimize the sum of squares of residuals which is a non-linear function of the parameters. For this purpose, some non-linear optimization procedures, such as the Marquardt method (Marquardt(1963)) or the Davidon-Fletcher-
Powell method (Fletcher-Powell (1963)) are used.

It must be noted, however, that polynomial type non-linear model is unstable (see Ozaki et al (1977)). This is because the estimated non-linear polynomial \( f(z) \) is valid only for

\[
\min_{0 \leq t \leq N} (z_{t-1}) \leq z \leq \max_{0 \leq t \leq N} (z_{t-1})
\]

and there is no guarantee that the simulated "intermediate variable" \( z_t \) remains within that domain under the given input and Gaussian assumption of the error term \( \varepsilon_t \). Some modification of \( f(\cdot) \) is necessary when \( z_t \) exceeds the domain although such a case does not occur so often in real applications. If we approximate the non-linear function by another type polynomials given by say

\[
f(z) = a_1 e^{-z^2} + a_2 ze^{-z^2} + a_3 z^2 e^{-z^2} + \ldots + a_r z^{-r} e^{-z^2},
\]

the above mentioned unstability and non-stationarity can be avoided.

We note that \( e^{-z^2}, ze^{-z^2}, z^2 e^{-z^2}, \ldots \) form a base of Hilbert space \( L_2(-\infty, \infty) \) of square integrable functions.

4. NUMERICAL RESULTS

Fig. 6 shows the riverflow time series \( y_t \) (1 \( \leq t \leq 366 \)) which is the daily discharge of Kanna River from 1st January 1956 to 31st December 1956 in \( \log_{10} \) scale. Fig. 7 shows the rainfall time series \( x_t \) of drainage basin of the Kanna River. The AIC values and residual variances of some models are given in Table 1.

---

Fig. 6 and Fig. 7 should be inserted here.
Table 1 AIC values and residual variances of some models

<table>
<thead>
<tr>
<th>(p,q,r)</th>
<th>$\hat{\sigma}^2$</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>(6,2,3)</td>
<td>0.3179x10^{-2}</td>
<td>-2042.6</td>
</tr>
<tr>
<td>(6,1,3)</td>
<td>0.3202x10^{-2}</td>
<td>-2042.1</td>
</tr>
<tr>
<td>(5,2,3)</td>
<td>0.3207x10^{-2}</td>
<td>-2041.5</td>
</tr>
<tr>
<td>(6,1,4)</td>
<td>0.3191x10^{-2}</td>
<td>-2041.3</td>
</tr>
<tr>
<td>(6,3,3)</td>
<td>0.3176x10^{-2}</td>
<td>-2041.0</td>
</tr>
<tr>
<td>(7,2,3)</td>
<td>0.3176x10^{-2}</td>
<td>-2041.0</td>
</tr>
<tr>
<td>(7,3,3)</td>
<td>0.3173x10^{-2}</td>
<td>-2038.8</td>
</tr>
</tbody>
</table>

Among the models listed in Table 1, the following model (6,2,3) gives the minimum value of AIC.

$$y_t = 0.05882 + z_{t-1} + 0.3567z_{t-1}^2 - 0.1804z_{t-1}^3 + \epsilon_t,$$

$$z_{t-1} = 0.7458y_{t-1} - 0.03255y_{t-2} - 0.03313y_{t-4} + 0.04519y_{t-4}$$

$$- 0.01054y_{t-5} + 0.04930y_{t-6} + 0.01183x_{t-1} - 0.001133x_{t-2},$$

$$\hat{\sigma}^2 = 0.3179x10^{-2}. \quad (4.1)$$

The validity of the model for riverflow prediction and the identification procedure are checked as follows. We estimate the parameters of the model order (6,2,3) using the first 280 length data, and we predict one-step ahead the latter 86 data using the estimated model. Results obtained are shown in Fig. 8, where the continuous line shows the observations and the mark $\Diamond$ for $1 \leq t \leq 280$ and the mark $*$ for $281 \leq t \leq 366$ show the predicted values. It seems that the results show good agreement between

Fig. 8 should be here.
the predicted riverflow values with the observed ones for $281 \leq t \leq 366$, when we take account into the fact that the estimated model has not used any information of the rainfall time series $x_t$ and riverflow time series $y_t$ for $281 \leq t \leq 366$. The estimated "state variable" $z_t$ of the model varies between 0.2060 and 2.2512 for $7 \leq t \leq 365$. The non-linear function $y_t = f(z_{t-1})$ of the model (4.1) for $0.2060 \leq t \leq 2.2512$ is shown in Fig. 9.

Fig. 9 should be inserted here.

The fitting procedure is also applied to the daily riverflow data $Q_t$ ($1 \leq t \leq 2192$) (See Fig. 10) of Bird Creek from 1st Oct. 1955 to 30th Sept. 1961 given by Sugawara (1974). Fig. 10 shows the daily rainfall $P_t$ of drainage basin of the Bird Creek. Since $Q_t$ takes zero value in some period, we $\log_{10}$-transform the positive $Q_t$, and $\log_{10} 0$ is replaced for convenience by -3 which is slightly smaller than $\log_{10}$ of the smallest positive value of $Q_t$. For this transformed data $y_t$, the model (6,6,6) is adopted among several models of different orders by the MAICE procedure. The estimated model is given by

$$y_t = -0.3056 \cdot 10^{-1} + z - 0.3337 \cdot 10^{-1} z_{t-1}^2 - 0.4095 \cdot 10^{-2} z_{t-1}^3$$

$$- 0.7682 \cdot 10^{-3} z_{t-1}^4 - 0.6001 \cdot 10^{-3} z_{t-1}^5 + 0.9488 \cdot 10^{-4} z_{t-1}^6 + \epsilon_t$$

$$z_{t-1} = 1.0279 y_{t-1} - 0.2710 y_{t-2} + 0.1777 y_{t-3} + 0.01171 y_{t-4}$$

$$+ 0.04447 y_{t-5} - 0.02065 y_{t-6} + 0.3203 \cdot 10^{-1} x_{t-1} + 0.6718 \cdot 10^{-2} x_{t-2}$$

$$- 0.8059 \cdot 10^{-3} x_{t-3} - 0.3529 \cdot 10^{-2} x_{t-4} - 0.1944 \cdot 10^{-2} x_{t-5}$$

$$- 0.2445 \cdot 10^{-2} x_{t-6}$$
\[ \hat{\sigma}^2 = 0.4886 \times 10^{-1} \]

The one-step ahead predictions \( \hat{y}_t (1993 \leq t \leq 2192) \) by the model whose parameters are estimated from \( y_t (1 \leq t \leq 1992) \) and \( x_t (1 \leq t \leq 1992) \) are shown in Fig. 12. Continuous line in the figure shows the observed riverflow, and the mark \( \hat{\Phi} \) for \( 1 \leq t \leq 1992 \) and the mark \( * \) for \( 1993 \leq t \leq 2192 \) show the predicted values. It seems that our non-linear dynamic system model fits very well and gives good prediction of the riverflow of Bird Creek when compared with Sugawara's results by tank models (See Sugawara (1974)). The estimated non-linear function \( y_t = f(z_{t-1}) \) of the model \( (6,6,6) \) for the whole data of Bird Creek is shown in Fig. 13. The figure of the function is more complicated and more non-linear than the one of Kanna River.

Fig. 13 should be inserted here.

5. CONCLUSION

It has been shown that the non-linear dynamic system model, which was introduced by taking account of the dynamic structure of rainfall and riverflow of Sugawara's tank model, is practically useful. One of the weak points of the Sugawara's model has been the difficulty of the identification of the model. Although Sugawara (1977) introduced a semi-automatic identification procedure, no explicit objective function for the goodness of fit of the model was given. Our method gives not only a mathematical form for Sugawara's model but also the objective function
for the evaluation of the goodness of fit of models.

Acknowledgement

The author is grateful to Dr M. Sugawara and Mrs Y. Katsuyama of National Research Center of Disaster Prevention, Japan, and to Mr N. Komura of Ministry of Construction, Japan, for providing him the data of Bird Creek and Kanna River respectively. The author is also grateful to Dr H. Akaike and Dr H. Tong for their discussions and encouragements.

This work was supported in part by the Science Research Council, U. K., at the University of Manchester Institute of Science and Technology and in part by a grant from Ministry of Education, Japan, at the Institute of Statistical Mathematics, Tokyo.
REFERENCES


Fig. 1. Tank model for Bikin River.
Fig. 2. Tank model for Bird Creek.
Fig. 3. A tank with three exits.

Fig. 4. A piece-wise linear function $y = f(z)$.

Fig. 5. A non-linear feedback system.
Fig. 6. Riverflow data of Kanna River (from 1st January 1956 to 31st December 1956)

Fig. 7. Rainfall data of the basin of Kanna River (from 1st January 1956 to 31st December 1956).
Fig. 8. One-step ahead prediction of the discharge of Kanna River by the model \((6,2,3)\).
Fig. 9. Non-linear function $y = f(z)$ of the model (6,2,3) for Kanna River.
Fig. 11. Rainfall data of the basin of Bird Creek (from 1st October, 1955 to 30th September, 1961)

Fig. 10. Riverflow data of Bird Creek (from 1st October, 1955 to 30th September, 1961)
Fig. 13. Non-linear function $y = f(z)$ of the model $(6,6,6)$ for Bird Creek.