TIME SERIES MODELLING OF ANNUAL MAXIMUM FLOW OF RIVER INDUS AT SUKKUR

M. Shakeel Ahmad, M. Idrees Ahmad, H.M. Naeem & M. Sarwar
University of Agriculture, Faisalabad

A study was conducted to construct a forecasting model for the annual maximum flood data of river Indus at Sukkur barrage. The Box-Jenkins ARIMA (autoregressive integrated moving average) methodology was used for forecasting. The diagnostic checking showed that ARIMA (2, 1, 1) was appropriate. Using ARIMA (2, 1, 1), 11 years ahead forecasts and 95% confidence interval were made.

INTRODUCTION

Time series modelling of annual maximum flow data has several uses in water resources. These are used to forecast future series of floods. Generation of synthetic series are generally needed for reservoir sizing, for determining the risk of failure (or reliability) of water supplies for irrigation systems, for planning studies of future reservoir operation, for planning capacity expansion of water supply systems and for other similar applications.

Srikanthan *et al.* (1983) used time series models to analyse annual flow of Australian streams. Auto-correlation and partial auto-correlation functions were applied to determine the appropriate form of Box-Jenkins time series models. O'Connel (1977) studied the Auto-regressive moving average (ARMA) models to generate synthetic flow series and found that ARMA (1, 1) was better to preserve the characteristics of stream flow series. The application of time series models in water resources research has been neglected so far. The present study was attempted to construct a forecasting model for the annual flood data of river Indus at Sukkur barrage.

MATERIALS AND METHODS

The study was directed to model the annual flood data recorded at Sukkur barrage for the years 1901 to 1985. The data for 85 years were taken from Bhutto *et al.* (1989). A choice was to be made as to which type of the model should be developed. We made a choice from a class of linear time series models introduced by Box and Jenkins (1970). According to them, the ARIMA model is denoted by ARIMA (p, d, q), where "p" is the order of the autoregressive process; "d" is the order of homogeneity i.e. the number of differences to make the series stationary and "q" is the order of the moving average process. The general form of the ARIMA (p, d, q) given by Pindyck and Rubinfeld (1976) using the backward shift operator "B" would be as follows:

$$\Phi (B) \; {}^{d}y_t = \delta + \Theta (B) \epsilon_t \; \ldots\ldots\ldots\ldots \quad (2.1)$$

with

$$\Phi (B) = 1 - \Phi_1 B - \Phi_2 B^2 - \ldots\ldots\ldots\ldots - \Phi_p B^p$$

and

$$\Theta (B) = 1 - \Theta_1 B - \Theta_2 B^2 \ldots\ldots\ldots\ldots \Theta_q B^q$$

Φ (B) is called auto-regressive operator,
Θ (B) is called moving average operator
and

$$\blacktriangle yt = y_t - y_{t-1}$$
$$\blacktriangledown^2 yt = \blacktriangledown y_t - \blacktriangledown y_{t-1}$$

and so on, y_{t-1} .... y_{t-p}. The observations yt are generated by weighted average of past observations going back P-periods together with random disturbance in the current period in an auto-regressive process of order P and is given by:

$$y_t = \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \ldots\ldots + \Phi p y_{t-p} + \delta + \epsilon_t \qquad (2.2)$$

δ is constant term which relates to the mean of stochastic process, while in moving average process of order q observations yt will be generated by the equation:

$$yt = \mu + \epsilon t - \Theta_1 \epsilon_{t-1} - \Theta_1 \epsilon_{t-2} \ldots\ldots \Theta_p \epsilon_{t-q} \qquad (2.3)$$

RESULTS AND DISCUSSION

The annual maximum flood data for 85 years (1901-85) at Sukkur barrage were used for modelling purposes. The data were transformed logarithmically so as to make it normal. The modelling of the time series involved the steps of model specification, model estimation, diagnostic checking and forecasts. The model specification involved

Table 1. Goodness of fit statistics

ARIMA model	Modified Box-Pierce Chi-square statistic	Akaike's statistic
(1, 1, 1)	13.8 (DF = 10)	22.50
(2, 1, 1)	9.4 (DF = 9)	20.72
(1, 1, 2)	13.4 (DF = 8)	26.34
(2, 1, 2)	15.7 (DF = 9)	22.43

Table 2. Forecasts from period 1985-86 to 1994-95 (0000 cusecs)

		95% limits	
Period	Forecast	Lower	Upper
1985-86	2.76623	2.52870	3.00376
1986-87	2.72882	2.49112	2.96652
1987-88	2.77142	2.52540	3.01743
1988-89	2.76293	2.51673	3.00910
1989-90	2.77310	2.52599	3.02020
1990-91	2.77118	2.52392	3.01843
1991-92	2.77360	2.52610	3.02111
1992-93	2.77317	2.52553	3.02081
1993-94	2.77375	2.52595	3.02155
1994-95	2.77365	2.52572	3.02159

the plot of the differenced series. The plot of the 2nd differenced series showed that the parameter d (number of differences to make the series stationary) was one. Auto-correlation function indicates the order of auto-regressive components q of model, while the partial correlation function gives an indication for the parameter P. The plots of differenced series for auto-correlation function and partial auto-correlation function suggested that four specifications could be candidates for ARIMA model i.e. (1, 1, 1), (2, 1, 1), (1, 1, 2) and (2, 1, 2). These specified models were estimated by MINITAB Computer Package.

The modified Box-Pier statistic calculated for lag 12 alongwith Akaike's information criteria are given in Table 1 for the models under consideration. This table reveals that the Chi-square values are non-significant for all the models indicating that all these models have been adequate for these flood series. However, the Akaike's information criteria in Table 1 have given a clue that ARIMA (2, 1, 1) could be able to best preserve the characteristics of Indus river because it has the least value for Akaike's information as compared to other candidates. Thus, ARIMA (2, 1, 1) was taken for the synthetic generation of 11 years ahead flow which are given in Table 2 alongwith the 95% confidence interval values given in this table are the logarithm of actual values.

REFERENCES

Bhutto, H.B.G. and N.M. Shaikh. 1989. Flood frequency analysis of the Indus river at Sukkur barrage. J. Engg. & App. Sci. 8: 7-16.

Box, G.E.P. and G.M. Jenkins. 1970. Time Series Analysis: Forecasting and Control. Holden-Day, San Francisco, California, USA.

O'Connel, P.E. 1977. ARIMA model in synthetic hydrology: Mathematical model for surface hydrology. John Wiley, NY, USA.

Pindyck, R.S. and D.L. Rubinfeld. 1976. Econometric Models and Economic Forecasts. McGraw Hill Book Co. Inc., NY, USA.

Srikanthan, R., T.A. McMohan and J.L. Irish. 1983. Time series analysis of annual flows of Australian streams. J. Hydrology.