STOCHASTIC GENERATION OF INFLOW SCENARIOS TO BE USED BY OPTIMAL RESERVOIR OPERATION MODELS

Alcigeimes B. CELESTE[1], Koichi SUZUKI[2], Akihiro KADOTA[3] and Camilo A. S. de FARIAS[4]

[1]Member of JSCE, Ph.D. Student, Dept. of Civil and Environmental Engineering, Ehime University (3 Bunkyo-cho, Matsuyama, Ehime 790-8577, Japan)
[2]Member of JSCE, Dr. of Eng., Professor, Dept. of Civil and Environmental Engineering, Ehime University (3 Bunkyo-cho, Matsuyama, Ehime 790-8577, Japan)
[3]Member of JSCE, Dr. of Eng., Assistant Professor, Dept. of Civil and Environmental Engineering, Ehime University (3 Bunkyo-cho, Matsuyama, Ehime 790-8577, Japan)
[4]Undergraduate Student, Dept. of Civil Engineering, Federal University of Campina Grande (Av. Aprígio Veloso, 882, Campina Grande 58.109-970, Paraíba, Brazil)

The Thomas-Fiering stochastic model for synthetic streamflow generation is used to determine monthly inflow scenarios for the watershed of the reservoir that supplies the city of Matsuyama, Ehime Prefecture. The scenarios are going to be used by a stochastic programming model which is being developed for the optimal operation of the reservoir. The Thomas-Fiering model allows for the non-stationarity of seasonal data. Twenty years of historical data are used for calibrating the model parameters and a new 20-year synthetic series is generated. The comparison between the statistics of historical and synthetic discharges shows that the model can preserve the characteristics of the historical series and effectively incorporate them into the generated data.

Key Words: *Inflow generation, Thomas-Fiering model, stochastic forecast*

1. INTRODUCTION

The sustained economic development of any region is directly related to reliable and well-managed water resources. In areas with scarcity of water and increasing population, the principal means of avoiding water shortage is by an efficient management of the existing systems.

The city of Matsuyama, capital of Ehime Prefecture, suffers periodically from problems originated from the lack of water. There is a great necessity of a better development and management of the water resources in the region. Several studies have been carried out in order to apply mathematical *optimization methods* to the operation of the reservoir that supplies the city (Ishitegawa Dam)[1]. The uncertainties of the inflows into the reservoir play a very important role in the determination of the optimal operating policy and thus should be taken into account by optimization models. One way of doing so is by using *stochastic programming* techniques, which implicitly incorporate the uncertainties of the inflows into the optimization model.

The authors are currently implementing an optimization model based on the so-called *stochastic programming with recourse*[2] for the optimal operation of Ishitegawa Dam. The model aims to determine the best allocation of water for city supply and irrigation that meet their respective demands to the greatest extent possible. The general structure of the stochastic programming model is as follows:

$$\min \sum_{t=1}^{N} \left\{ \begin{array}{l} \alpha_1 \left(\dfrac{Q_{rel}^t - T_{dem}^t}{T_{dem}^t} \right)^2 + \alpha_2 \left(\dfrac{Q_{irr}^t - T_{irr}^t}{T_{irr}^t} \right)^2 \\ + \alpha_3 \sum_{l=1}^{L} p^l \left(\dfrac{E_{stor}^{t,l} + D_{stor}^{t,l}}{V_{stor}^{\max}} \right)^2 \end{array} \right\} \quad (1)$$

subject to

$$Q_{rel}^t + Q_{irr}^t + E_{stor}^{t,l} + D_{stor}^{t-1,l} - D_{stor}^{t,l} = \tilde{V}_{inf}^{t,l}; \forall t; \forall l \quad (2)$$

$$0 \le Q_{rel}^t \le \min[T_{dem}^t, Q_{rel}^{\max}]\,;\forall\, t \quad (3)$$

$$0 \le Q_{irr}^t \le \min[T_{irr}^t, Q_{irr}^{\max}]\,;\forall\, t \quad (4)$$

$$0 \le E_{stor}^{t,l} \le V_{spill}^{\max}\,;\forall\, t \quad (5)$$

$$0 \le D_{stor}^{t,l} \le V_{stor}^{\max} - V_{stor}^{\text{dead}}\,;\forall\, t\,;\forall\, l \quad (6)$$

$$0 \le V_{spill}^t \le V_{spill}^{\max}\,;\forall\, t\,;\forall\, l \quad (7)$$

$$V_{stor}^t - E_{stor}^{t,l} + D_{stor}^{t,l} = V_{stor}^{\max}\,;\forall\, t\,;\forall\, l \quad (8)$$

in which t is the time index; N is the operating horizon; l is the scenario index; L is the number of scenarios; α_1, α_2, α_3 are coefficients that measure the relative importance given to each of the reservoir operation purposes; Q_{rel}^t is the allocation for city supply; T_{dem}^t is the demand for city supply; Q_{irr}^t is the allocation for irrigation; T_{irr}^t is the demand for irrigation; p^l is the probability of occurrence of scenario l; $E_{stor}^{t,l}$ and $D_{stor}^{t,l}$ are, respectively, surplus (excess) and shortage (deficit) variables defined as in Eq. (8); V_{stor}^t is the reservoir storage; $\widetilde{V}_{inf}^{t,l}$ is the random inflow to the reservoir; V_{spill}^t is the amount of water that might spill from the weir; $Q_{rel}^{\max}$ is the total capacity of the surface water treatment plants; $Q_{irr}^{\max}$ is the capacity of the irrigation system; V_{stor}^{dead} is the dead storage of the reservoir; $V_{stor}^{\max}$ is the capacity of the reservoir; and $V_{inf}^{\max}$ is the maximum spill.

As seen above, the stochastic programming approach requires a set of possible future inflow scenarios ($\widetilde{V}_{inf}^{t,l}$), each of them having a particular probability of occurrence (p^l). This paper deals with the determination of these scenarios.

Forecast of future discharges is usually performed by means of *stochastic streamflow generation models*[3], which produce synthetic data having the same statistical features of the historical streamflow series. In this study a stochastic streamflow generation model is applied to the watershed of Ishitegawa Dam. First, the parameters of the model are calibrated and used to produce a synthetic monthly inflow series, which is compared to the historical one. Then, a set of inflow scenarios are generated for past years and compared to the actual inflow records for those years.

2. DESCRIPTION OF THE MODEL

Basically, two approaches are available to generate synthetic monthly flows[3]. The first one removes the seasonalities and periodicities in the monthly flows and then the resulting stationary season is modeled[4]. The second approach uses the parametric autoregressive *Thomas-Fiering*[5),6)] model, which tries to preserve the various statistic parameters such as mean, standard deviation, correlation coefficient, etc. Either of the approaches has its own advantages and disadvantages. The Thomas-Fiering model was selected in this study due to its relative simplicity and for allowing the seasonal data series to be non-stationary.

The Thomas-Fiering model may be viewed as a non-stationary first order autoregressive (AR) model[3]. First order AR models, also called *Markov* models, are normally used to describe *stationary* series, e.g., annual runoff series. A series is said to be stationary when the probabilistic laws that rule the process do not alter with time. AR models adequately reflect the phenomenon known as *persistency*, or *memory* of the process, according to which the value of the discharge at any time is dependent in part on the value of the discharge in the previous time. This dependency can be measured by a regression analysis of the flow at a given time and its previous value. As show in Fig. 1, if the regression coefficient of season j on j-1 is b_{j-1}, then the regression line value of a seasonal flow, X_j, can be determined from the previous season's flow, X_{j-1}, by the following equation:

$$X_j = \mu_j + b_{j-1}(X_{j-1} - \mu_{j-1}) \quad (9)$$

in which μ_j and μ_{j-1} are the average flows in seasons j and j-1, respectively.

Any hydrologic series observed with a time interval of less than a year would have a non-stationary structure because of the cyclic component with a period of one year introduced by the astronomic cycle. The seasonal models must take this non-stationarity into account. The Thomas-Fiering model was developed to implicitly allow for the non-stationarity of seasonal data. In this

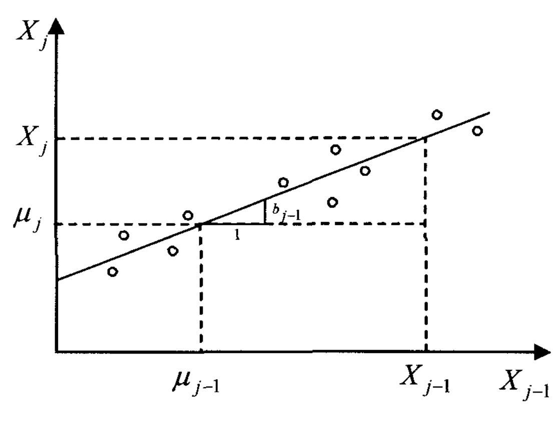

Fig. 1 Principle of the Thomas-Fiering model.

model, the parameter b_{j-1} is calculated as below:

$$b_{j-1} = \frac{\rho_j \sigma_j}{\sigma_{j-1}} \tag{10}$$

where ρ_j is the correlation coefficient between the flows of jth and (j-1)th seasons; σ_j and σ_{j-1} are the standard deviations of the flows in seasons j and j-1, respectively.

The variance of the measured data about the regression line is taken into account by adding a further stochastic component given by $z\sigma_j\sqrt{(1-\rho_j^2)}$ in which z is a normal process with mean zero and variance unity $N(0,1)$. In practice, the seasonal model is used by replacing the population parameters by their respective sample estimates obtained from the historical record. Thus, the general Thomas-Fiering model in terms of the sample estimates is written as follows:

$$x_j = \bar{x}_j + \frac{r_j s_j}{s_{j-1}}[x_{j-1} - \bar{x}_{j-1}] + z s_j \sqrt{(1-r_j^2)} \tag{11}$$

in which x_j is flow in the jth season; $\bar{x}_j$ and s_j are the mean and standard deviation of the flows in jth season, respectively; r_j is the correlation coefficient between the flows of jth and (j-1)th seasons; and z is a normal process with mean zero and variance unity.

In the Thomas-Fiering model the flow in any season is thus given by a sum of three terms: the first term is the mean flow in that season; the second term is the regressed component on the flow in the previous season; and the third term is a random component to reflect the desired variance.

3. DETERMINATION OF THE PARAMETERS

The Thomas-Fiering model was implemented in MATLAB environment. Twenty years of historical monthly inflow data were used for calculating the statistical sample parameters of Eq. (11).

Analyzing the histogram of the 20-year inflow data (Fig. 2), it can be noted that they follow a skewed distribution. In order to determine the distribution that would best represent these data, the sample skewness coefficient was compared with the skewness coefficients of the Gamma and the Log-Normal distributions, as shown in Table 1. Since the value of the sample coefficient was closer to that of the Gamma distribution, this distribution was selected for the random component z.

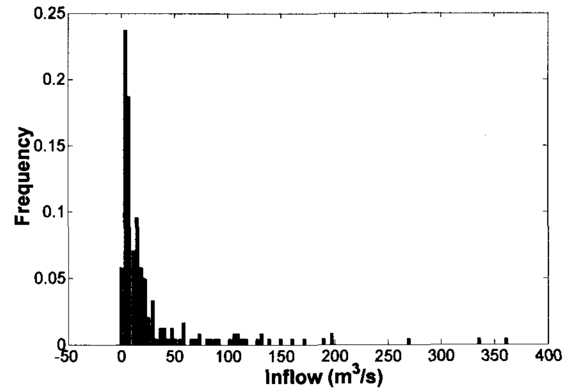

Fig. 2 Histogram of the historical series.

Table 1 Skewness coefficients

Sample	Gamma	Log-Normal
3,77	3,59	10,94

When the flows follow Gamma distribution, the random component z has to be replaced by ξ given below, so that the skewness verified in the sample is preserved in the data to be generated:

$$\xi = \frac{2}{\hat{g}_j}\left[1 + \frac{\hat{g}_j z}{6} - \frac{\hat{g}_j^2}{36}\right]^3 - \frac{2}{\hat{g}_j} \tag{12}$$

where z is from $N(0,1)$, as usual, and g_j is called the modified skewness coefficient given by

$$\hat{g}_j = \frac{g_j - r_j^3 g_{j-1}}{(1-r_j^2)^{1.5}} \tag{13}$$

in which g_j is the skewness coefficient of the observed inflow data.

Table 2 shows the parameters from expressions (11) and (13) calculated by using the historical inflow data (m^3/s).

4. GENERATION OF THE SYNTHETIC SERIES

With the parameters from Table 2, a new 20-year synthetic series was generated. Statistics such as mean, standard deviation and skewness coefficient in each month of the synthetic data were compared with those from the historical series. Different numbers of independent processes z were used for each year of the synthetic series: one single process (and thus one single series generated for the given year), 10 processes, or 100 processes. Table 3 shows the

Table 2 Monthly sample estimates

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
$\bar{x}$	6,25	8,41	10,17	12,89	16,88	58,23	102,14	35,51	39,95	20,48	11,89	9,72
r	0,10	0,54	0,50	-0,02	0,37	-0,16	0,11	0,74	0,76	0,04	0,08	0,73
s	5,70	7,30	6,83	8,70	16,12	50,16	109,77	59,73	48,44	24,96	11,74	10,46
$\hat{g}$	1,34	0,68	1,54	0,86	1,58	0,68	1,14	4,57	1,26	2,41	1,45	2,63

Table 3 Comparison of monthly and annual statistics for the historical and synthetic series

	Historical Series			Synthetic Series (1 independent process)			Synthetic Series (10 independent processes)			Synthetic Series (100 independent processes)		
	Mean	Standard Deviation	Skewness Coefficient	Mean	Standard Deviation	Skewness Coefficient	Mean	Standard Deviation	Skewness Coefficient	Mean	Standard Deviation	Skewness Coefficient
Jan	6,2	5,7	1,7	8,4	5,9	1,0	6,6	5,9	1,2	6,3	5,7	1,6
Feb	8,4	7,3	0,8	9,9	6,8	0,6	9,8	7,1	0,8	8,4	6,9	0,9
Mar	10,2	6,8	1,4	12,5	6,4	1,1	10,9	7,0	0,8	10,2	7,0	1,3
Apr	12,9	8,7	1,1	12,1	9,8	1,5	12,9	9,3	1,1	13,0	8,8	0,9
May	16,9	16,1	1,6	16,3	15,3	1,1	16,9	15,5	1,7	16,5	14,9	1,3
Jun	58,2	50,2	0,8	55,5	51,1	0,7	59,7	50,4	1,0	60,2	48,3	1,0
Jul	102,1	109,8	1,4	122,7	104,3	1,0	106,9	113,3	1,8	107,7	103,9	1,3
Aug	35,5	59,7	2,4	58,3	48,8	0,9	36,9	48,4	2,0	41,4	56,3	2,5
Sep	39,9	48,4	1,5	49,5	42,8	0,7	36,4	38,1	1,2	43,9	45,9	1,6
Oct	20,5	25,0	3,0	14,8	20,4	1,8	19,1	23,9	3,4	20,8	25,4	2,5
Nov	11,9	11,7	1,8	11,2	8,9	0,9	13,1	11,4	1,2	11,7	11,0	1,6
Dec	9,7	10,5	1,8	9,9	12,7	2,2	10,6	9,0	1,2	10,0	10,3	1,8
Annual	27,7	49,7	3,8	31,8	50,8	3,2	28,3	49,8	4,6	29,2	49,7	3,8

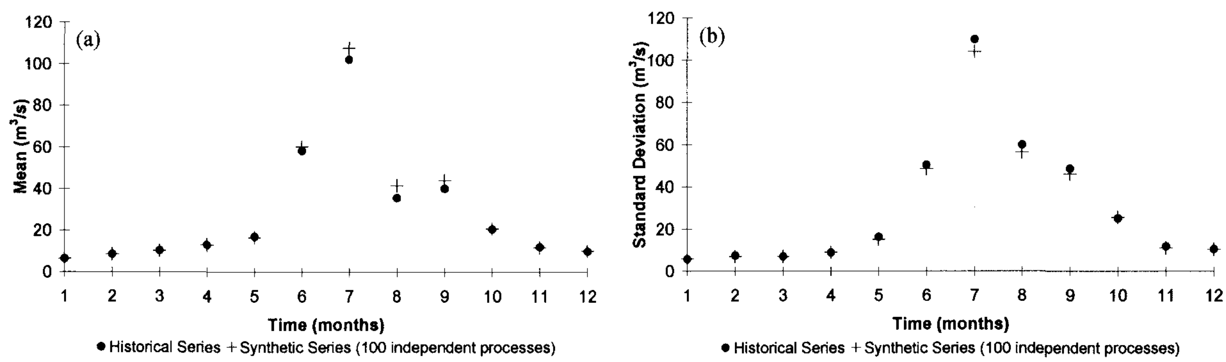

Fig. 3 Comparison between (a) mean and (b) standard deviations of the historical and synthetic series (100 independent processes).

results obtained.

Fig. 3 shows graphs comparing the mean and standard deviation of the historical and synthetic (with 100 independent processes) series, as illustrated in Table 3. The histograms of the synthetic series for all numbers of independent processes are presented in Fig. 4.

5. GENERATION OF THE INFLOW SCENARIOS

The stochastic programming model (1)-(8) is going to be tested with data from past years (1979-1996). For each year, the optimization model will use monthly inflow scenarios for the current year as well as the next year, e.g., for the operation in 1996 the model needs the scenarios for the period 1996-1997, and so on. Consequently, the Thomas-Fiering model was used to generate monthly inflow scenarios for each pair of years within the period 1979-1996. As illustration, Fig. 5 shows data for a 5-scenario

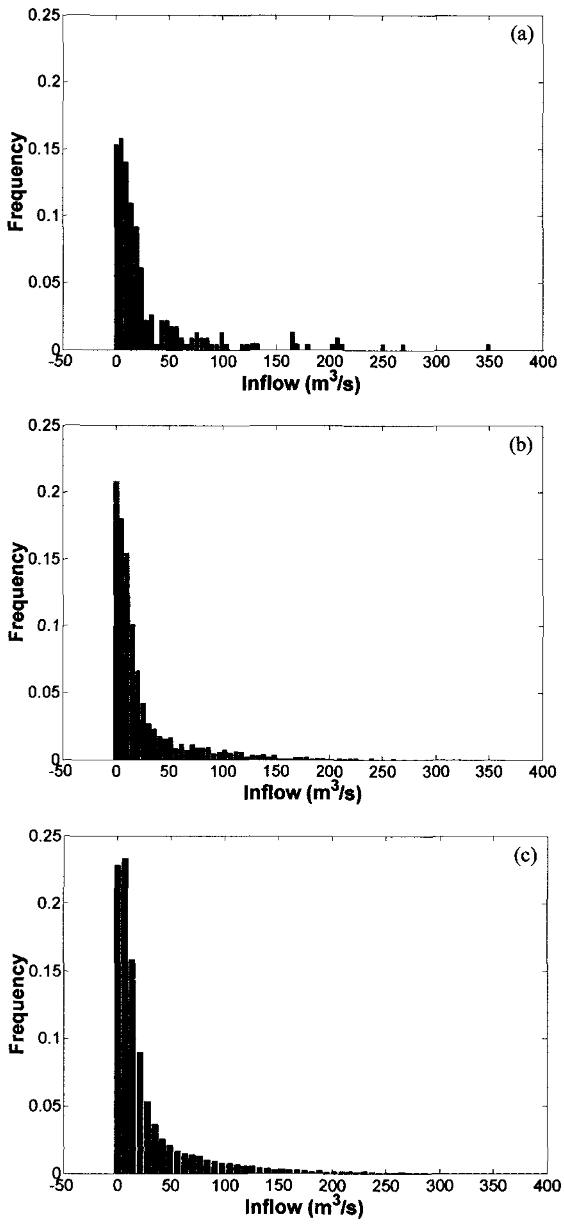

Fig. 4 Histogram of the synthetic series according to the number of independent processes: (a) 1 process, (b) 10 processes, (c) 100 processes.

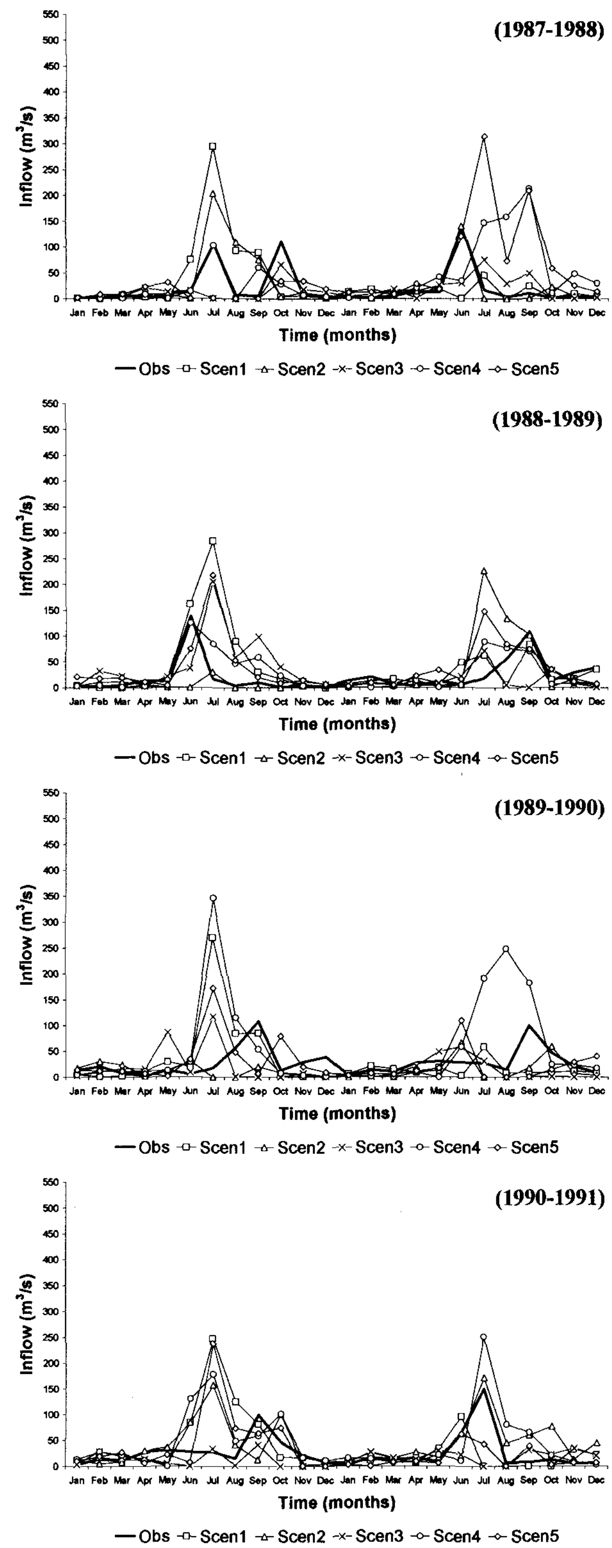

Fig. 5 Generated scenarios (1987-1989 until 1990-1991).

generation (i.e., using five number of random processes z) for the last ten years. The observed inflow data are also shown for comparison. Any number of scenarios can be generated by selecting different independent processes. Each scenario represents one possibility of occurrence of the inflow in the given year.

6. CONCLUSIONS

In this study, the stochastic streamflow generation model of Thomas-Fiering was applied to synthetically generate monthly inflow scenarios for the reservoir that supplies Matsuyama City, in Ehime Prefecture. The scenarios are going to be used by an optimization model based on stochastic programming which is currently being developed for the operation of Ishitegawa Dam.

The main characteristic of the Thomas-Fiering model is to implicitly allow for the non-stationary data of monthly inflows.

From the results obtained with the use of a 20-year historical series for computing the model parameters,

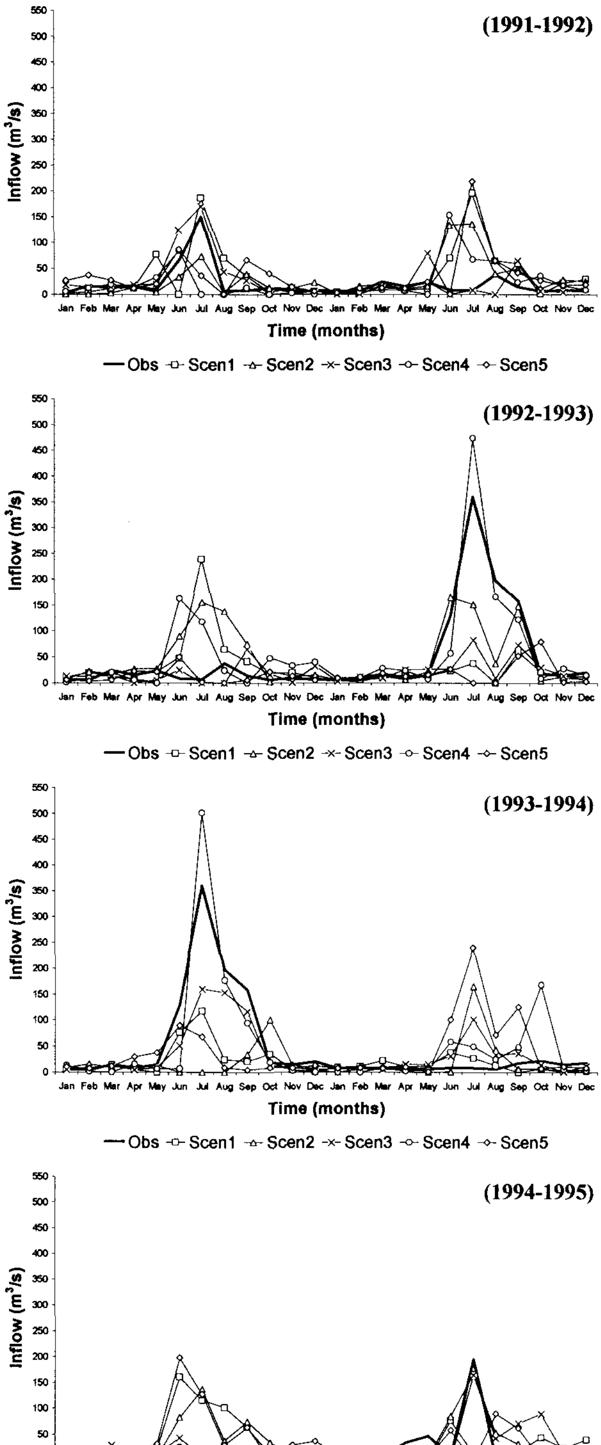

Fig. 5 (cont.) Generated scenarios (1991-1992 until 1994-1995).

it was noticed that the statistics (mean, standard deviation and skewness coefficient) of the generated data were very close to their historical correspondents. Besides, the histogram of the historical series was similar to those from the

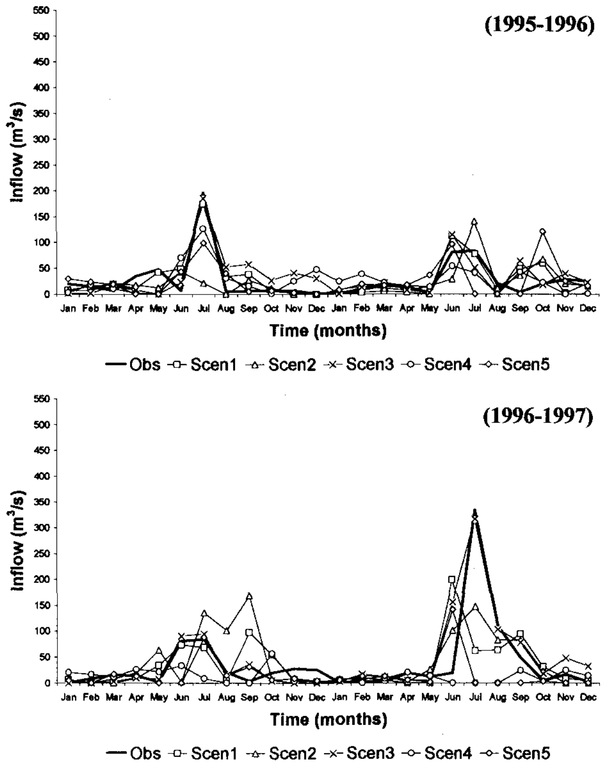

Fig. 5 (cont.) Generated scenarios (1995-1996 and 1996-1997).

synthetic series. This signifies that the Thomas-Fiering model could successfully incorporate the statistical features of the historical data into the generated values. Thus, it can be said that the model is suitable to be used for producing the inflow scenarios needed by the stochastic optimization model.

REFERENCES

1) Celeste, A. B., Ishikawa, T., Kadota, A. and Suzuki, K. Operation of Ishitegawa Dam by Mathematical Programming. *9th JSCE Congress of Civil Engineering - Shikoku Division*, Takamatsu, Japan, pp. 111-112, 2003.
2) Birge, J. R., and Louveaux, F. *Introduction to stochastic programming*. Springer-Verlag, New York, 1997.
3) Reddy, P. J. *Stochastic hydrology*. Laxmi Publications, Delhi, 1987.
4) Yevjevich, V. *Stochastic processes in hydrology*. Water Resources Publication, Fort Collins, Colorado, 1972.
5) Thomas, H. A., & Fiering, M. B. Mathematical synthesis of stream-flow sequences for the analysis of river basins by simulation. in: A. Maass et al. (ed) *Design of Water Resource Systems*. Harvard Univ. Press, Cambridge, Massachusetts, pp. 459–493, 1962.
6) Fiering, M. B. *Streamflow synthesis*. Harvard Univ. Press, Cambridge, Massachusetts, 1967.

(Received September 30, 2003)