Generation Models for Synthetic Annual and Monthly Flows for Some Indiana Watersheds

J. W. Delleur and I. T. Kisisel[1]
School of Civil Engineering
Purdue University, Lafayette, Indiana 47907

Abstract

A first order autoregressive process can be used to generate annual flows. The mean annual flow and the rank one serial correlation coefficient were related to watershed and climatological properties. A first order Markov process of the Fiering type adequately simulates the sequence of normalized logarithms of monthly flows. The means and standard deviations of the logarithms of the monthly flows and the autoregressive constant were related to climatological and basin characteristics.

Introduction

Among many possible methods of modeling watershed behavior for prediction and forecasting purposes, stochastic methods have the advantage of taking into account the chance dependent nature of hydrologic events. Using the methodology of time series analysis, two models for generating synthetic sequences of yearly and monthly flows were developed. For full details the reader is directed to reference (4).

Data Collection Procedures

The watersheds examined have areas ranging from 100 to 3,750 square miles, and had continuous runoff records varying from 21 to 44 years; all are free of regulation by reservoirs during the period of record.

The runoff data for the watersheds were annual, monthly and daily average discharges. Daily average discharges were obtained from a magnetic tape prepared by the National Weather Records Center. Monthly and annual series were constructed from the daily average discharge information.

Annual Average Discharges Model

The annual runoff data sequences obeyed a Gaussian probability distribution. An example is shown in Figure 1. A first-order autoregressive scheme was found to model satisfactorily the annual average flow sequences of the selected Indiana watersheds. The model was formulated as follows:

$$A_{i+1} = \overline{A} + \alpha(A_i - \overline{A}) + r_{i+1}\, S(1 - \alpha^2)^{1/2} \qquad [1]$$

where A_i = annual average discharge
$\overline{A}$ = mean of A_i sequence
α_j = autoregressive constant (rank one serial correlation coefficient of A_1 sequence)
r_i = random variable, zero mean and unit variance
S = standard deviation of A_i sequence

[1] Current Address: Dept. of Civil Eng., Middle East Tech. Univ., Ankara, Turkey.

If $\overline{A}$, S and α are known *a priori*, Equation 1 can be used to generate annual average flow sequences. The values of $\overline{A}$, S and α can be estimated by using regression equations that relate these properties to some other available information.

A regression equation relating $\overline{A}$ to some physical watershed characteristics was developed by Marie and Swisshelm (5), using all the available information for Indiana watersheds; and it is a reliable estimator of $\overline{A}$. The equations for S and α were developed by using the historic records of 11 of the watersheds used in this study. These three equations are of the form,

$$y = a_o \cdot x_1^{a_1} \cdot x_2^{a_2} \cdot x_3^{a_3} \cdots x_n^{a_n} \qquad [2]$$

and are given in tabular form in Table 1 in which basin characteristics used as dependent variables are the following: (see reference (5) for full definitions)

Ar, Drainage area in square miles.

Le, Main-channel length in miles.

Sl, Main-channel slope in feet per mile, as described by Benson (1, 2).

Fr, Forest cover expressed as the percent of the drainage area.

Pr, Mean annual precipitation in inches.

St, Area of lakes and ponds, expressed as percentage of the drainage area plus 1%.

I24,2, The maximum 24-hour rainfall having a recurrence interval of 2 years expressed in inches.

Tj, The average minimum daily January temperature expressed in degrees Fahrenheit.

Sn, The mean annual snowfall expressed in inches.

Gi, A geologic index expressed as a dimensionless number.

Per, A soil permeability index expressed as a dimensionless number.

The regeneration performance of the model suggested (Equation 1) was checked by computing the probability distribution, the autocorrelation function and the power spectrum for the generated series for all the watersheds and by comparing the results with those of the historic records. The comparison indicated that there were no substantial differences between these pairs of functions for each watershed. An example of the comparison of the probability distributions is given in Figure 1.

Monthly Runoff Sequences Model

The monthly runoff data possess a strong cyclic component which is exhibited by the sinusoidal appearance of the correlation function and by the strong peak in the spectrum at the annual frequency.

A logarithmic transformation was used to bring the probability distribution closer to normal. The removal of the annual cyclic com-

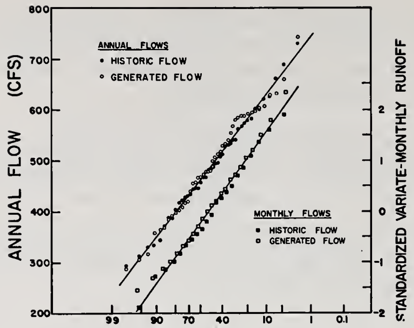

FIGURE 1. *Probability distributions of historic and generated annual flow sequences for the Kankakee River at Davis, Indiana, and Probability Distribution of Transformed, Standardized historic and generated Monthly flow sequences for Young's Creek near Edinburg, Indiana.*

ponent from the log-transformed monthly flow sequences was performed by introducing the standardized variate:

$$Z_j = \frac{\log Q_i - M_j}{S_j} \qquad [3]$$

where Q_i = monthly flow

where M_j = mean of log-transformed monthly flows for the month of j

S_j = standard deviation of log-transformed monthly flows for the month of j

i = index running from 1, N (N Total number of months in the series)

j = index for months, running 1 to 12

The standardized runoff sequences obeyed approximately a Gaussian probability distribution. An example is shown in Figure 1.

The autocorrelation functions of the log-transformed and normalized monthly runoff sequences exhibited positive values for lags less than about 20 months and the power spectra gradually decayed from low to high frequencies as shown in Figure 2. These results suggest that an autoregressive scheme can be used to model the runoff sequences. A Fiering (3) type first order autoregressive model was chosen to represent the sequences of the logarithm transformed and standardized monthly flows. Recalling Equation 3 for the normal variate Z_i, the model can be written as follows:

$$Z_{i+1} = a\, Z_i + r_{i+1}\, (1 - a^2)^{1/2} \qquad [4]$$

In this model the seasonal variation of the autoregressive constant a is omitted and one value for a watershed is assumed to be representative.

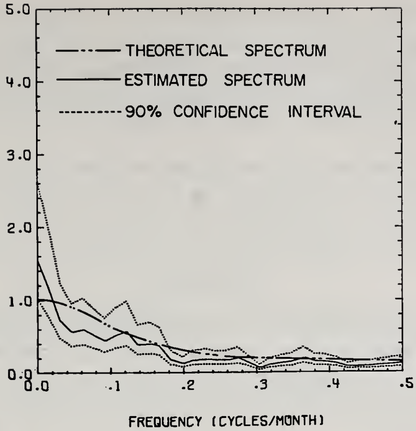

FIGURE 2. *Power spectra of transformed standardized monthly runoffs and of the fitted model for the White Water River, near Alpine, Indiana.*

The autoregressive constant α was related to the basin characteristics by a regression equation. For each watershed, the 12 mean logarithms of the monthly flows (M_j) and the 12 standard deviations of the logarithms of the monthly flows (S_j) were also related to the basin characteristics by regression equations. These 25 equations are similar to those in Table 1 and are presented in Ref. 4. Therefore, to generate a sequence of monthly flows the following steps can be taken:

1) Using basin characteristics and regression equation for α, determine α,

2) Using equation 4 generate sequences of Z_i,

3) Using $Z_i = (\log Q_i - M_j)/S_j$ and regression equations for M_j and S_j generate sequences of Q_i.

TABLE 1. *Regression equations for $\overline{A}$, S, and α. (Annual Flows).*

		A	S	α
	a_0	0.0033	0.1120	0.1330
a_1	(Ar)	0.993	1.33	—
a_2	(Le)	—	—	—
a_3	(Sl)	—	0.489	0.579
a_4	(Fr)	0.034	—	—
a_5	(Pr)	1.130	—	6.837
a_6	(St)	—	—	—
a_7	(I24,2)	—	—	2.818
a_8	(Tj)	0.496	—	9.720
a_9	(Sn)	—	—	—
a_{10}	(Gi)	0.018	—	—
a_{11}	(Per)	—	—	0.552
R^2		*	0.975	0.745

*After Reference (5), standard error of estimate 8%.

The regeneration performance of the suggested model was tested by comparing the probability distributions and the power spectra of the historic and generated records. An example of the probability distributions of the variate Z_i, for monthly historic flow sequences, and for the generated sequences is shown in Figure 1.

Another application of time series analysis distinct of generation is that of forecasting. Models for forecasting monthly flows have been developed by McKerchar and Delleur (6, 7) in an extension of this work.

Conclusions

Since only a small number of watersheds are considered in this preliminary study, definitive statements should be avoided as far as the particular details are concerned. The following general points can be used as guidelines for more detailed research based on more comprehensive chronological data.

1) The annual flow sequences for Indiana watersheds can be modeled by means of a first order autoregressive scheme. No particular transformation of the data sequences is needed. The model parameters were correlated to measurable climatological and geomorphological quantities.
2) The logarithm transformation for the monthly average flows is suggested in order to improve the normality of these sequences for the watersheds considered in this study.
3) The removal of the annual cycle from the sequences of monthly runoff yields residual sequences for which simple stochastic models can be formulated. Once the residual series is synthetically generated, means and standard deviations can be used to construct synthetic sequences with the same statistical properties (to the second order) as the historical sequences.
4) A first order autoregressive model was found satisfactory to generate the monthly runoff sequences. The model parameters were correlated to measurable climatological and geomorphological quantities.

Acknowledgments

The authors are grateful to Drs. R. Prasad, T. P. Chang, L. A. Schaal, J. Travis, and Dan Wiersma for assistance with various aspects of the project.

This work was supported by the Office of Water Resources Research, U.S.D.I., through Purdue University Water Resources Research Center under project number OWRR-A-020-IND.

Literature Cited

1. BENSON, M. A. 1962. Factors influencing the occurrence of floods in a humid region of diverse terrain. U.S.G.S. Water Supply Pap. 1580-B. B1-62 p.

2. __________. 1964. Factors affecting the occurrence of floods in the southwest. U.S.G.S. Water Supply Paper 1580-D. D1-D69 p.

3. FIERING, M. B. 1967. Streamflow synthesis. Harvard Univ. Press, Cambridge, Mass. 135 p.

4. KISISEL, I. T., and J. W. DELLEUR. 1971. An analysis of hydrologic time series and generation models of synthetic flows for some Indiana watersheds. Purdue Univ. Water Resources Center, Tech. Rep. No. 19, 60 p.

5. MARIE, J. R., and R. V. SWISSHELM, Jr. 1970. Evaluation of and recommendations for the surface-water data program in Indiana. U.S.G.S. Water Resources Div., Open File Rep., Indianapolis, Ind. 47 p.

6. MCKERCHAR, A. L., and J. W. DELLEUR. 1972. Stochastic analysis of monthly flow data. Application to Lower Ohio River Tributaries. Purdue Univ. Water Resources Res. Center, Tech. Rep. No. 26, 92 p.

7. __________. 1972. Evaluation of seasonal time series models: application to Midwest river flow data, Int. Symp. on Uncertainties in Hydrologic and Water Resources Systems, Univ. Ariz. Proc. 1:31-47.