WRC RESEARCH REPORT NO. 26

STOCHASTIC ANALYSIS OF HYDROLOGIC SYSTEMS

Ven Te Chow
Principal Investigator

FINAL REPORT

Project No. A-029-ILL

The work upon which this publication is based was supported by funds provided by the U.S. Department of the Interior as authorized under the Water Resources Research Act of 1964, P.L. 88-379
Agreement No. 14-01-0001-1632

UNIVERSITY OF ILLINOIS
WATER RESOURCES CENTER
3220 Civil Engineering Building
Urbana, Illinois 61801

December, 1969

ABSTRACT

STOCHASTIC ANALYSIS OF HYDROLOGIC SYSTEMS

Hydrologic phenomena are in reality stochastic in nature; that is, their behavior changes with the time in accordance with the law of probability as well as with the sequential relationship between the occurrences of the phenomenon. In order to analyze the hydrologic phenomenon, a mathematic model of the stochastic hydrologic system to simulate the phenomenon must be formulated. In this study, a watershed is treated as the stochastic hydrologic system whose components of precipitation, runoff, storage and evapotranspiration are simulated as stochastic processes by time series models to be determined by correlograms and spectral analysis. The hydrologic system model is then formulated on the basis of the principle of conservation of mass and composed of the component stochastic processes. To demonstrate the practical application of the method of analysis so developed, the upper Sangamon River basin above Monticello in east central Illinois is used as the sample watershed. The watershed system model so formulated can be employed to generate stochastic streamflows for practical use in the analysis of water resources systems. This is of particular value in the economic planning of water supply and irrigation projects which is concerned with the long-range water yield of the watershed.

Chow, Ven Te
STOCHASTIC ANALYSIS OF HYDROLOGIC SYSTEMS
Research Report No. 26 , Water Resources Center, University of Illinois at Urbana-Champaign, December 1969, 34 pp.
KEYWORDS--systems analysis/stochastic processes/synthetic hydrology/water resources development/watershed studies/precipitation/streamflow/evapotranspiration/storage/water yield/hydrologic models/hydrology

CONTENTS

I. INTRODUCTION

It is generally noted that the natural hydrological system and hydrologic process are truly "stochastic"; that is, the behavior of the system or the process varies with a sequential time function of the probability of occurrence [1,2].* In other words, the hydrologic phenomenon changes with the time in accordance with the law of probability as well as with the sequential relationship between its occurrences. For example, the occurrence of a flood is considered to follow the law of probability and also the relationship with the antecedant flood condition.

Most conventional methods for hydrologic designs are "deterministic," that is, the behavior of the hydrologic system or process is assumed independent of time variations. For example, a unit hydrograph derived for a given river basin for flood-control project design is based on historical flood records. Once derived, the unit hydrograph is used for analysis of future design floods. Thus, it is automatically assumed unchanged with time (from the past to the future) and therefore is deterministic.

Some conventional methods employ the concept of probability to the extent that no sequential relationship is involved in the probability. For example, the flood record is analyzed and fitted with a certain probability distribution to determine the recurrence intervals of the flood or the flood frequencies. Such methods are "probabilistic" but not in the true sense "stochastic."

---

* Numbers in parentheses refer to references listed at the end of the report.

The stochastic method, that is to employ the concept of probability as well as its sequential relationship, has not been well introduced in the practical design and planning of hydrologic projects, because such methods have not been fully developed. While the natural hydrologic phenomenon is stochastic, it is important to develop the stochastic method of hydrologic analysis for hydrologic system design. Conventional methods, deterministic and probabilistic, which do not conform more closely to the natural phenomenon, will produce results that depart from the true behavior of the hydrologic phenomenon and hence have the possibility to either overdesign or underdesign the hydrologic project [3].

The objective of this study is to formulate the mathematical model of a stochastic hydrologic system and the mathematical models of the hydrologic processes in the system, using the watershed as an example of the hydrologic system. In this study, in other words, the framework of a method was developed to utilize mathematical models to simulate the stochastic behavior of a watershed as the hydrologic system. The mathematical models so developed should have a practical application to the analysis of hydrologic systems in the water resources planning and development.

The initial step of the study involved a comprehensive review of the application of the theory of stochastic process in hydrology. The results of this initial step of investigation are reported separately as "Water Resources Systems Analysis - Annotated Bibliography on Stochastic Processes" [4] and "Water Resources Systems Analysis - Review of Stochastic Processes" [5].

II. FORMULATION OF THE HYDROLOGIC SYSTEM MODEL

In the formulation of the hydrologic system model, a watershed is used as the hydrologic system although the mathematical approach would be equally applicable to other kinds of hydrologic systems with some modifications depending on the nature of the system. The watershed is treated as a hydrologic system which has an input, mainly rainfall, and an output, mainly runoff and evapotranspiration. The input and output are to be treated as time series or stochastic processes which describe the stochastic behavior of the input and output processes. The amount of water stored in the watershed is also treated as a time series or stochastic process which describes the stochastic nature of infiltration, subsurface runoff and the soil moisture and groundwater storages.

To formulate a mathematical model for the watershed hydrologic system, the runoff is considered as the integral product of three component stochastic processes; namely, (1) a "conceptual watershed storage" at the end of the t-th time interval representing the storage of water on the ground surface, such as lakes, ponds, swamps and streams, as well as below the ground surface, such as soil moisture and groundwater reservoirs, (2) the total rainfall input during the t-th time interval, and (3) the total losses, mainly evapotranspiration, during the t-th time interval. These three component stochastic processes can be mathematically represented respectively by time series functions $[S(t);\ t \in T]$, $[X(t);\ t \in T]$ and $[E(t);\ t \in T]$ where T is the time range under consideration or the length of the hydrologic record. These stochastic processes can be simply denoted by S_t, X_t and E_t, respectively. They are not considered as independent but as a stochastic vector $[S(t), X(t), E(t);\ t \subset T]$. The theory of time series can therefore be used to formulate the stochastic model of this vector. A rigorous mathematical analysis of this vector would require the use of the theory of multiple time series analysis [6]. In view of the accuracy of the natural hydrologic data and for the purpose of practical application without resorting to excessive mathematical involvement, the stochastic vector is to be analyzed by the single time series analysis techniques of correlogram and spectrum in combination with the cross-spectrum theory which provides a powerful tool in the analysis of multiple time series.

By the basic concept of system continuity, the runoff, which is a stochastic process of total runoff output during the t-th time interval as denoted by $[Y(t);\ t \in T]$ or simply Y_t, can be related to the other three component stochastic processes of the hydrologic system as follows:

$$S_t = S_{t-1} + X_t - Y_t - E_t \tag{1}$$

where S_{t-1} is the conceptual watershed storage at the beginning of t-th time interval.

III. MATHEMATICAL TECHNIQUES

A. Mathematical Models for Time Series

In this study three models of time series which have been used in hydrologic study were reviewed. These models or their combinations would be employed to simulate the hydrologic stochastic processes. The hydrologic time series is denoted by $[u_t;\ t \in T]$ where u_t is the hydrologic variable attributed to the t-th time interval and T is the length of the hydrologic record.

1. Moving-Average Model. This model may be expressed as

$$u_t = a_1\varepsilon_t + a_2\varepsilon_{t-1} + \ldots + a_m\varepsilon_{t-m+1} \tag{2}$$

where ε is a random variable; $a_1, a_2, \ldots, a_m$ are the weights; and m is the extent of the moving average. This equation may be taken as the model representing the relation between, say, annual runoff u and, say, annual effective precipitation ε, where m is the extent of the carryover due to the water-retardation characteristics of the watershed. For such a model, the weights $a_1, a_2, \ldots, a_m$ must be all positive and sum to unity. By virtue of the moving average on the ε's, the simulated time series u is not random but stochastic.

2. Sum-of-Harmonics Model. This model may be expressed as

$$u_t = \sum_j^N \left(A_j \cos \frac{2\pi j t}{T} + B_j \sin \frac{2\pi j t}{T}\right) + \varepsilon_t \tag{3}$$

where A_j and B_j are the amplitudes; $2\pi jt/T$ is the period of cyclicity with $j = 1,2,\ldots,$ and N being the number of record intervals in months, years or other units used in the analysis; and ε_t is a random variable. This equation may be taken as a model representing a regular or oscillatory form of variations, such as diurnal, seasonal and secular changes that exist frequently in hydrologic phenomena. Such variations are of nearly constant period and they may be assumed sinusoidal as simulated in the model.

3. Autoregression Model. The general form of this model may be expressed as

$$u_t = f(u_{t-1}, u_{t-2}, \ldots, u_{t-k}) + \varepsilon_t \tag{4}$$

where f( ) is a mathematical function, k is an integer, and ε_t is a random variable. A special case of this model is the linear autoregressive model of the n-th order:

$$u_t = a_1 u_{t-1} + a_2 u_{t-2} + \ldots + a_n u_{t-n} + \varepsilon_t \tag{5}$$

where $a_1, a_2, \ldots, a_n$ are the regression coefficients. For n = 1, the above equation becomes the first-order Markov process:

$$u_t = a\, u_{t-1} + \varepsilon_t \tag{6}$$

where a is the Markov-process coefficient.

The autoregression model may be used as a model representing hydrologic sequences whose nonrandomness is due to storage in the hydrologic system, such as a watershed.

B. The Correlogram

The choice of an appropriate time series model for a given hydrologic process is not an easy task because the above-mentioned three models all exhibit oscillations resembling the fluctuations which one usually observes on most hydrologic data by visual inspection. A well-known analytical approach which can help one to select the best model is the analysis of the sample correlogram.

The correlogram is a graphical representation of the serial correlation coefficient r_k as a function of the lag k where the values r_k are plotted as ordinates against their respective values of k as abscissas. In order to reveal the features of the correlogram better, the plotted points are joined each to the next by a straight line. The serial correlation coefficient of lag k is computed by

$$r_k = \frac{\text{cov}(u_t, u_{t+k})}{[\text{var}(u_t)\ \text{var}(u_{t+k})]^{1/2}} \tag{7}$$

where $\text{cov}(u_t, u_{t+k})$ is the sample autocovariance and $\text{var}(u_t)$ and $\text{var}(u_{t+k})$ are the sample variance; or

$$\text{cov}(u_t, u_{t+k}) = \frac{1}{N-k} \sum_{t=1}^{N-k} u_t\, u_{t+k} - \frac{1}{(N-k)^2} \left(\sum_{t=1}^{N-k} u_t \right) \left(\sum_{t=1}^{N-k} u_{t+k} \right)^2 \tag{8}$$

$$\text{var}(u_t) = \frac{1}{N-k} \sum_{t=1}^{N-k} u_t^2 - \frac{1}{(N-k)^2} \left(\sum_{t=1}^{N-k} u_t \right)^2 \tag{9}$$

and

$$\mathrm{var}(u_{t+k}) = \frac{1}{N-k} \sum_{t=1}^{N-k} u_{t+k}^2 - \frac{1}{(N-k)^2} \left(\sum_{t=1}^{N-k} u_{t+k} \right)^2 \tag{10}$$

The correlogram provides a theoretical basis for distinguishing among the three types of oscillatory time series mentioned previously. It has been proved analytically that if the time series is simulated by a moving-average model for random elements of extent m, then the correlogram will show a decreasing linear relationship and vanishes for all values of $k > m$. For a sum-of-harmonics model, the correlogram itself is a harmonic with periods equal to those of the harmonic components of the model and it will therefore show the same oscillations. In the case of an autoregression model, the correlogram will show a damping oscillating curve. In the case of a first-order Markov process with a serial correlation coefficient r_1, it will oscillate with period unity above the abscissa with a decreasing but nonvanishing amplitude if r_1 is negative [7].

It may be noted that, when the time series is too short, the computed correlogram may exhibit substantial sampling variations and thus may conceal its actual form.

C. The Spectrum Analysis

This method is another diagnostic tool for the analysis of time series in the frequency domain, which can help develop an appropriate time series model for the hydrologic process.

All stationary stochastic processes can be represented in the form

$$u_t = \int_{-\pi}^{\pi} e^{it\omega} \, dz(\omega) \tag{11}$$

where $i = \sqrt{-1}$ and $z(\omega)$ is a complex, random function. Using this as a generating process, it can be shown that the autocovariance for a stationary process is [8]

$$\gamma_k = \int_{-\pi}^{\pi} e^{ik\omega}\, dF(\omega) \tag{12}$$

where $i = \sqrt{-1}$, k is the time lag, ω is the angular frequency, and $F(\omega)/\gamma_o$ is a distribution function monotonically increasing and bounded between $F(-\pi) = 0$ and $F(\pi) = \gamma_o = \sigma^2$ where σ is the standard deviation. The function $F(\omega)$ is called the "power spectral distribution function." For $k = 0$, Eq. (12) gives

$$\gamma_o = \sigma^2 = \int_{-\pi}^{\pi} dF(\omega) \tag{13}$$

which shows that $dF(\omega)$ represents the variance attributed to the frequency band $(\omega, \omega+d\omega)$. Thus, $dF(\omega) = f(\omega)d\omega$ where $f(\omega)$ is called the "power spectrum" of the process.

In the practical hydrologic application of the spectral theory the processes are real and the imaginary component is dropped off, thus Eq. (12) becomes

$$\gamma_k = 2 \int_{0}^{\pi} \cos k\omega\, f(\omega)\, d\omega \tag{14}$$

The mathematical inversion of the above equation gives the power spectrum as

$$f(\omega) = \frac{1}{2\pi} (\gamma_o + \sum_{k=1}^{\infty} \gamma_k \cos k\omega) \tag{15}$$

For a finite amount of data $[u_t;\ t\epsilon T]$ an estimate of the power spectrum is

$$\hat{f}(\omega) = \frac{1}{2\pi} (C_o + 2 \sum_{k=1}^{T-1} C_k \cos k\omega) \quad (16)$$

where C_k is the autocovariance for a time lag k.

The estimate of the power spectrum by Eq. (16) is called the "raw spectral estimate" because it does not give a smooth power spectral diagram. To adjust for the smoothness, it is common to use the "smoothed spectral estimate" in the form

$$\hat{f}(\omega) = \frac{1}{2\pi} [\lambda_o(\omega)C_o + 2 \sum_{k=1}^{m} \lambda_k(\omega)C_k \cos k\omega] \quad (17)$$

where $\lambda_k(\omega)$ are selected weighting factors and m is a number to be chosen much less than T. A commonly used weighting factor is the "Tukey-Hamming" weights [9]:

$$\lambda_k(\omega) = 0.54 + 0.46 \cos \frac{\pi k}{m} \quad (18)$$

where m is taken as less than T/10.

The significance of the spectrum is that it exhibits less sampling variations than the corresponding correlogram. Consequently, the estimated spectrum would provide a better evaluation of the various parameters involved in a model. If the generating process contains periodic terms, the frequencies of these terms will appear as high and sharp peaks in the estimated spectrum and the height of the peaks will give a rough estimate of the amplitude.

IV. ANALYSIS OF THE HYDROLOGIC SYSTEM

A. The Watershed under Study

The watershed chosen as the hydrologic system to be analyzed in this study is the upper Sangamon River basin of 550 sq. mi. in size, above Monticello, Illinois, and located in east central Illinois. The criteria for selecting this watershed are that the available hydrologic data such as the precipitation, streamflow and temperature records have a reasonably concurrent period and that additional data if needed can be relatively easily collected due to convenient access to its location and to its data collecting agencies. Figure 1 shows the map of the Sangamon River basin above Monticello, Illinois with the locations of the stream gaging station at Monticello and the precipitation gages where data were observed for use in the analysis.

B. The Hydrologic Data

1. Precipitation. The monthly precipitations in inches were used in the analysis as the historical hydrologic inputs to the watershed system. The data were taken from the "Climatic Summary of the United States" published by the U.S. Weather Bureau for Illinois. The period of records used in the analysis extends from October 1914 through September 1965 for stations at Urbana, Clinton, Bloomington and Roberts, from March 1940 through September 1965 for the station at Rantoul, and from June 1942 through September 1965 at Monticello. The average monthly precipitations over the watershed were computed by the Thiessen polygon method.

2. Streamflow. The monthly streamflow records for the Sangamon River at Monticello, Illinois, were used as the historical

hydrologic outputs of the watershed system in the analysis. The U.S. Geological Survey, in its cooperative program with the Illinois State Water Survey and other state, local and federal agencies, collects long-term streamflow records to determine the performance of rivers and streams. The gaging station on the Sangamon River about one-half mile west of Monticello had published data available for the periods of February 1908 to December 1912 and June 1914 to September 1968. The monthly streamflows from September 1914 through September 1965 were used in the analysis.

3. Temperature. In the analysis, the average monthly temperatures from October 1914 through September 1965 were taken from the "Climatic Summary of the United States" published by the U.S. Weather Bureau for Illinois. The mean of the monthly average temperatures at the stations in Urbana and Bloomington was considered as the average monthly temperature of the watershed. The relative location of these two stations with respect to the watershed has suggested this choice.

4. Potential Evapotranspiration. Necessary to the analysis of the watershed hydrologic system is the estimation of the monthly potential evapotranspiration. There are several methods for the computation of the potential evapotranspiration. The method proposed by Hamon [10] was used because it has been tested in Illinois [11] with satisfactory results and the computation and the data requirement are rather simple.

The formula proposed by Hamon is

$$E_p = 0.0055\ D^2 P_t \tag{19}$$

where E_p is the daily potential evapotranspiration in inches, D is the possible hours of sunshine in units of 12 hours and P_t is the saturation vapor density (absolute humidity) in grams per cubic meter at the daily mean temperature. The value of D depends on the latitude of the watershed and the month of the year. The value of P_t depends on the temperature. Tables for evaluating the values of D and P_t are provided by Hamon [12]. The value of D is essentially the monthly daytime coefficient of the Hargreaves evapotranspiration formula [13]. The value of P_t can be found from the Smithsonian Meteorological Tables. For the watershed under consideration, its average latitude is 40° N. The values of D^2 for the twelve months are 0.64 (Jan.), 0.79 (Feb.), 0.99 (Mar.), 1.22 (Apr.), 1.44 (May), 1.56 (June), 1.51 (July), 1.31 (Aug.), 1.08 (Sept.), 0.86 (Oct.), 0.69 (Nov.), and 0.61 (Dec.).

The monthly potential evapotranspiration can then be computed by

$$E_{pm} = 0.0055\ nKD^2P_t \tag{20}$$

where n is the number of days for each month and K is a correction factor equal to 1.04 because P_t is estimated for the monthly mean temperature instead of the daily mean temperature.

C. Establishing the Records for Conceptual Watershed Storage and Actual Evapotranspiration

Rewriting Eq. (1) gives

$$E_t = X_t - Y_t - (S_t - S_{t-1}) \tag{21}$$

Since the values of monthly precipitation X_t and monthly runoff Y_t are known from the historical records, it is obvious from the above equation that if the record for the conceptual watershed storage S_t were known then the record for the actual monthly evapotranspiration E_t could be easily established. On the other hand, if the record of E_t were known and an initial value of S_t were assumed, then the record of S_t could also be established. Unfortunately neither S_t nor E_t can be computed in a direct manner.

It is known, however, that in late September and early October of each year in Illinois the amount of surface water on the watershed and the soil moisture are at a minimum. Especially in the case of very low amount of precipitation during the months of August, September and October, the watershed storage must be the lowest. This lowest amount of storage can be considered as the reference point of the conceptual watershed storage. In other words, the conceptual watershed storage is taken as zero at the beginning of the October of the year having very low precipitation during the months of August, September and October. In the present analysis, this happens to be the case for the year of 1914.

Once the initial stage of the conceptual watershed storage is established, the following procedure may be followed to establish the records of conceptual watershed storage and actual evapotranspiration.

If $S_{t-1} + X_t - Y_t \geqq E_{pt}$ where E_{pt} is the potential evapotranspiration for the t-th time interval, then the actual evapotranspiration $E_t = E_{pt}$. Thus, the initial storage S_t for the next time interval can be computed by Eq. (1).

If $S_{t-1} + X_t - Y_t < E_{pt}$, then $E_t = S_{t-1} + X_t - Y_t$ and Eq. (1) gives $S_t = 0$.

The mass curves of X_t, Y_t, E_t and $S_t - S_{t-1}$ are shown in Fig. 2. The difference between ΣX_t and ΣY_t is essentially equal to ΣE_t since $\Sigma(S_t - S_{t-1})$ is relatively small as plotted in an enlarged scale. The mass curve for $S_t - S_{t-1}$ represents the variation in conceptual watershed storage with a mean of 3.5 inches.

D. Analysis of the Hydrologic Processes

In this analysis, the stochastic processes of precipitation, conceptual watershed storage and evapotranspiration are not to be treated independently of each other but they are considered as a three-dimensional vector or a multiple-time series. Without introducing the theory of multiple-time series, which has yet to be further developed and refined, the following assumptions are to be made in the present analysis:

(a) Each stochastic process consists of two parts; namely, one deterministic and the other random and uncorrelated to the deterministic part and the parts of other processes.

(b) The deterministic part of each stochastic process consists also of two parts; one part depending only on time and the other part depending on the vector of the stochastic processes of precipitation, conceptual watershed storage and actual evapotranspiration at previous time intervals.

Based on the above assumptions, the first step is to determine the deterministic part of each process which depends on time. From the experience in hydrology and the exhibition of hydrologic data, the deterministic part appears to be a periodic function rather than a polynomial of time. Hence, the sample correlograms can be computed for each process to test the existence of harmonic components in the process.

The serial correlation coefficients r_k for time lag k for the processes of precipitation, conceptual watershed storage and the evapotranspiration were computed by Eqs. (7), (8), (9) and (10) for $t = 1,2,\ldots,T$.

In the present study, T is the length of the records equal to 612 months and k is from zero to n/10, say 60. The correlograms, or the plots of r_k versus k, for precipitation, conceptual watershed storage and evapotranspiration are shown in Figs. 3, 4 and 5 respectively. For all three processes these correlograms are oscillating without any indication of damping, thus revealing the presence of harmonic components in all the processes.

In order to determine the periods of the harmonic components which will be included in the model to simulate the hydrologic processes and the hydrologic system, the power spectrum for each of the processes should be computed.

From Eqs. (16) and (17), the raw and smoothed spectral estimates may be written respectively as

$$L(\omega_t) = \frac{1}{2\pi} \left(C_o + 2 \sum_{k=1}^{m-1} C_k \cos \frac{\pi k t}{m} + C_m \cos \pi t\right) \tag{22}$$

and

$$U(\omega_t) = \frac{1}{2\pi} \left(\lambda_o C_o + 2 \sum_{k=1}^{m-1} \lambda_k C_k \cos \frac{\pi k t}{m} + \lambda_m C_m \cos \pi t\right) \tag{23}$$

Substituting Eq. (18) for the Tukey-Hamming weights in Eq. (23) and simplifying,

$$U(\omega_t) = \frac{1}{2\pi} [0.54(C_o + 2 \sum_{k=1}^{m-1} C_k \cos \frac{\pi k t}{m} + C_m \cos \pi t)$$

$$+ 0.46(C_o + 2 \sum_{k=1}^{m-1} C_k \cos \frac{\pi k t}{m} \cos \frac{\pi k}{m} - C_m \cos \pi t)] \tag{24}$$

Since

$$\cos \frac{\pi kt}{m} \cos \frac{\pi k}{m} = \cos \frac{\pi k}{m}(t+1) + \cos \frac{\pi k}{m}(t-1) \tag{25}$$

and

$$\cos \pi t = -\frac{1}{2} [\cos \pi(t+1) + \cos \pi(t-1)] \tag{26}$$

Eq. (24) becomes

$$U(\omega_t) = 0.54\ L(\omega_t)$$

$$+ \frac{0.23}{2\pi} [C_o + 2 \sum_{k=1}^{m-1} C_k \cos \frac{\pi k}{m}(t+1) + C_m \cos \pi(t+1)]$$

$$+ \frac{0.23}{2\pi} [C_o + 2 \sum_{k=1}^{m-1} C_k \cos \frac{\pi k}{m}(t-1) + C_m \cos \pi(t-1)] \tag{27}$$

As the raw spectral estimates can be represented by Eq. (22), Eq. (27) may be written as

$$U(\omega_t) = 0.23\ L(\omega_{t-1}) + 0.54\ L(\omega_t) + 0.23\ L(\omega_{t+1}) \tag{28}$$

Computer programs were written to compute the autocovariance by Eq. (8) and the raw and smoothed spectral estimates by Eqs. (22) and (28). The smoothed spectra for precipitation, conceptual watershed storage and evapotranspiration are shown in Figs. 6, 7 and 8, respectively. The sharp peaks exhibited in these spectra indicate a significant amount of the variance with the periodicities of 12-month and 6-month which are appropriate for use in the model.

E. Determination of the System Model

The proposed model for the hydrologic processes is a combination of the sum-of-harmonics and the autogression time series models. Since the results of the correlogram and spectral analyses indicate the presence of the 12-month and 6-month periodicities, the general model for the hydrologic stochastic processes under study may be written in the form

$$U_t = c_1 + c_2 \sin\frac{2\pi t}{12} + c_3 \cos\frac{2\pi t}{12}$$

$$+ c_4 \sin\frac{4\pi t}{12} + c_5 \cos\frac{4\pi t}{12} + u_t' \qquad (29)$$

where c_1, c_2, c_3, c_4 and c_5 are the coefficients to be estimated and u_t' is the residual stochastic process with zero mean. This model was therefore used to fit the hydrologic processes of precipitation, conceptual watershed storage, and evapotranspiration by the least-square method such as the one described by Brown [14]. The coefficients of the model determined for precipitation, conceptual watershed storage and evapotranspiration are as follows:

	c_1	c_2	c_3	c_4	c_5
X_t	3.0382	-0.9701	0.1365	0.3717	0.0647
S_t	3.5231	0.5786	-2.3821	0.5583	-0.1366
E_t	2.2346	-2.0511	0.7408	-0.1563	-0.4051

The first five terms in the time series model represented by Eq. (29) are a portion of the deterministic part of the simulated hydrologic stochastic processes. The first term is a constant while the second, third, fourth and fifth terms are deterministic harmonics as functions of time. The last term u_t' represents the residual stochastic process which may consist of a deterministic portion and the random part of the model.

This deterministic portion may be correlated with the vector of the processes of precipitation, conceptual watershed storage and evapotranspiration at previous time intervals, while the random part of the process may be simulated by a representative probability distribution. The determination of a suitable model for the residual stochastic process will require further investigation. In further investigation, it may be suggested that the deterministic portion of the residual stochastic processes be analyzed by the cross-spectrum theory [8]. Although the residual stochastic process is a significant component of the model, its magnitude is of relatively low order. As a first approximation the residual stochastic processes in the watershed system may be considered completely random with their means equal to zero. Thus, for the present study, $X'_t=E'_t=S'_t=0$ and their variances were found to be 2.754, 0.465 and 4.136 respectively. Their probability distributions may be roughly assumed as normal at present until better probability distribution models are to be found in future investigation.

With the hydrologic processes of precipitation, conceptual watershed storage and evapotranspiration being determined, the runoff process may be formulated from Eqs. (1) and (29) as

$$Y_t = X_t - E_t - (S_t - S_{t-1}) \tag{30}$$

or

$$\begin{aligned} Y_t = {} & 0.8036 + 0.5024 \sin \frac{\pi t}{6} + 1.7778 \cos \frac{\pi t}{6} - 0.0303 \sin \frac{\pi t}{3} \\ & + 0.6064 \cos \frac{\pi t}{3} + 0.5786 \sin \frac{\pi(t-1)}{6} - 2.3821 \cos \frac{\pi(t-1)}{6} \\ & + 0.5583 \sin \frac{\pi(t-1)}{3} - 0.1366 \cos \frac{\pi(t-1)}{3} + X'_t - E'_t - (S'_t - S'_{t-1}) \end{aligned} \tag{31}$$

This is the system model expressed for the runoff process of the upper Sangamon River basin above Monticello, Illinois. This model can be employed to generate stochastic monthly streamflow values for use in the analysis of water resources systems. It is of particular value in the economic planning of water supply and irrigation projects which is concerned with the long-range water yield of the watershed.

V. CONCLUSIONS

The ultimate objective of the research on the stochastic analysis of stochastic hydrologic systems is to formulate the mathematical model for a stochastic hydrologic system for which a watershed is considered. The upper Sangamon River basin above Monticello, Illinois, is taken as an example of the watershed. This study has demonstrated that such a model is feasible and its application to a practical problem is workable.

For this study the literature on stochastic processes and their application in hydrology were reviewed. It was found that the application of the theory of stochastic processes in hydrology has barely begun and the theory has applied mostly to single processes but not to composite hydrologic systems. The mathematical theory of stochastic processes is very extensive, but unfortunately most of it is written not for practicing engineers and hydrologists. Furthermore, a systematic theory for the formulation of a stochastic system model is unavailable because the formulation of the model requires the practical knowledge on the physical characteristics of the process and the system which is usually lacking on the part of the mathematician. This study therefore attempts to introduce the use of a theoretical model to the simulation of a practical hydrologic system.

Based on the principle of conservation of mass, the watershed system is represented by the mass balance equation in which the system components of precipitation, conceptual watershed storage, evapotranspiration and runoff are considered as stochastic processes. While the data of precipitation and runoff are given, a method was developed to establish the unknown records of conceptual watershed storage and evapotranspiration.

A deterministic portion of the system component process is analyzed by the theory of correlogram and spectrum. Computer subroutines were programmed for the computation of correlograms and spectra of a discrete time series of finite length. The expected values of the system components of precipitation, conceptual watershed storage and evapotranspiration were thus found to be best simulated by harmonics of 12-month and 6-month periodicities. This analysis constitutes an important step in the attempt of considering the nonstationarity of the processes involved in the hydrologic system because the expected values are taken as functions of time but not constants.

The hydrologic system model so formulated for the upper Sangamon River basin can be used to generate stochastic streamflows for the use in the planning of water supply and irrigation projects in the basin. The method developed in this study is therefore formed to be of practical value in the analysis of water resources systems.

VI. ACKNOWLEDGMENT

This report is the result of a research project on "Stochastic Analysis of Hydrologic Systems" sponsored by the U.S. Office of Water Resources Research, which began in July 1968 and was completed in June 1969. Under the direction of the Project Investigator, the hydrologic data used in this study were mainly collected by Mr. Gonzalo Cortes-Rivera, Research Assistant in Civil Engineering, and the mathematical analysis and computations were largely performed by Mr. Sotirios J. Kareliotis, Research Assistant in Civil Engineering.

VII. REFERENCES

1. Chow, V. T., Statistical and probability analysis of hydrologic data: Part 1. Frequency analysis, Section 8 in Handbook of Applied Hydrology, ed. by V. T. Chow, McGraw-Hill Book Co., New York, 1964, p. 8-9.

2. Chow, V. T., A general report on new ideas and scientific methods in hydrology (Simulation of the hydrologic behavior of watersheds), Proceedings, International Hydrology Symposium, Fort Collins, Colorado, 6-8 September 1967, pp. 50-65.

3. Chow, V. T., Hydrologic systems for water resources management, Conference Proceedings of Hydrology in Water Resources Management, Water Resources Research Institute Report No. 4, Clemson University, Clemson, South Carolina, March 1968, pp. 8-22.

4. Chow, V. T., and Meredith, D. D., Water resources systems analysis - Part I. Annotated bibliography on stochastic processes, Civil Engineering Studies, Hydraulic Engineering Series No. 19, University of Illinois, Urbana, Illinois, July 1969.

5. Chow, V. T., and Meredith, D. D., Water resources systems analysis - Part III. Review of stochastic processes, Civil Engineering Studies, Hydraulic Engineering Series No. 21, University of Illinois, Urbana, Illinois, July 1969.

6. Quenouille, M. H., The analysis of multiple time series, Hafner Publishing Co., New York, 1957.

7. Dawdy, D. R., and Matalas, N. C., Analysis of variance, covariance, and time series, Section 8-III, Part III in Handbook of Applied Hydrology, ed. by V. T. Chow, McGraw-Hill Book Co., New York, 1964, p. 8-85.

8. Granger, C. W. J., and Hatanaka, M., Spectral analysis of economic time series, Princeton University Press, Princeton, New Jersey, 1964.

9. Blackman, R. B., and Tukey, J. W., The measurement of power spectra, Dover Publications, Inc., New York, 1959.

10. Hamon, W. R., Estimating potential evapotranspiration, Proceedings, American Society of Civil Engineers, Journal of Hydraulics Division, Vol. 87, No. HY3, pp. 107-120, May 1961.

11. Jones, D. M. A., Variability of evapotranspiration in Illinois, Illinois State Water Survey Circular 89, 1966.

12. Hamon, W. R., Estimating potential evapotranspiration, Massachusetts Institute of Technology Department of Civil and Sanitary Engineering, unpublished M.S. thesis, 1960.

13. Veihmeyer, F. J., Evapotranspiration, Section 11 in Handbook of Applied Hydrology, ed. by V. T. Chow, McGraw-Hill Book Co., p. 11-30, 1964.

14. Brown, R. C., Smoothing, forecasting and prediction of discrete time series, Prentice Hall, Inc., Englewood Cliffs, N.Y., 1962.

VIII. FIGURES

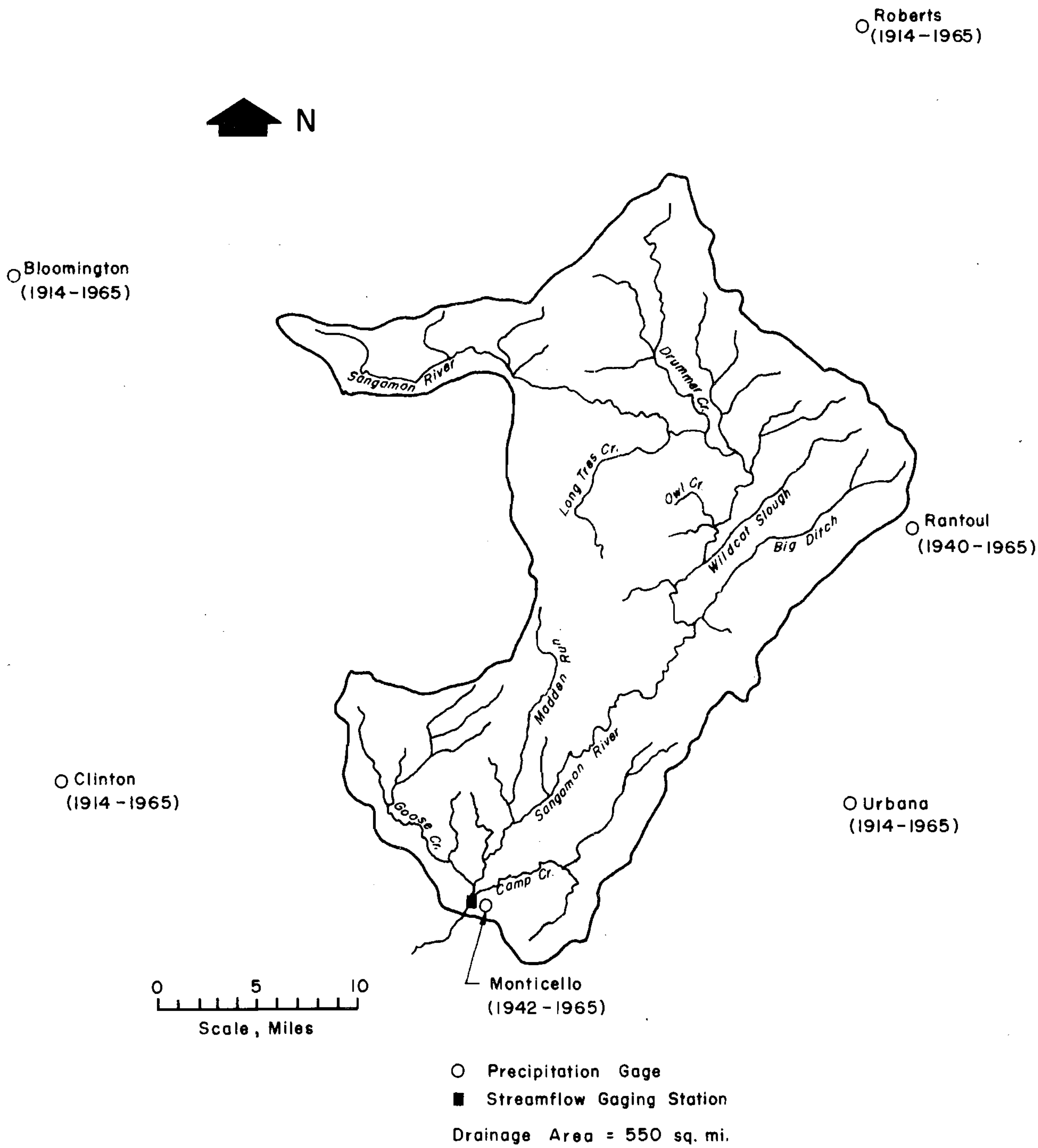

FIG. I SANGAMON RIVER BASIN ABOVE MONTICELLO, ILLINOIS

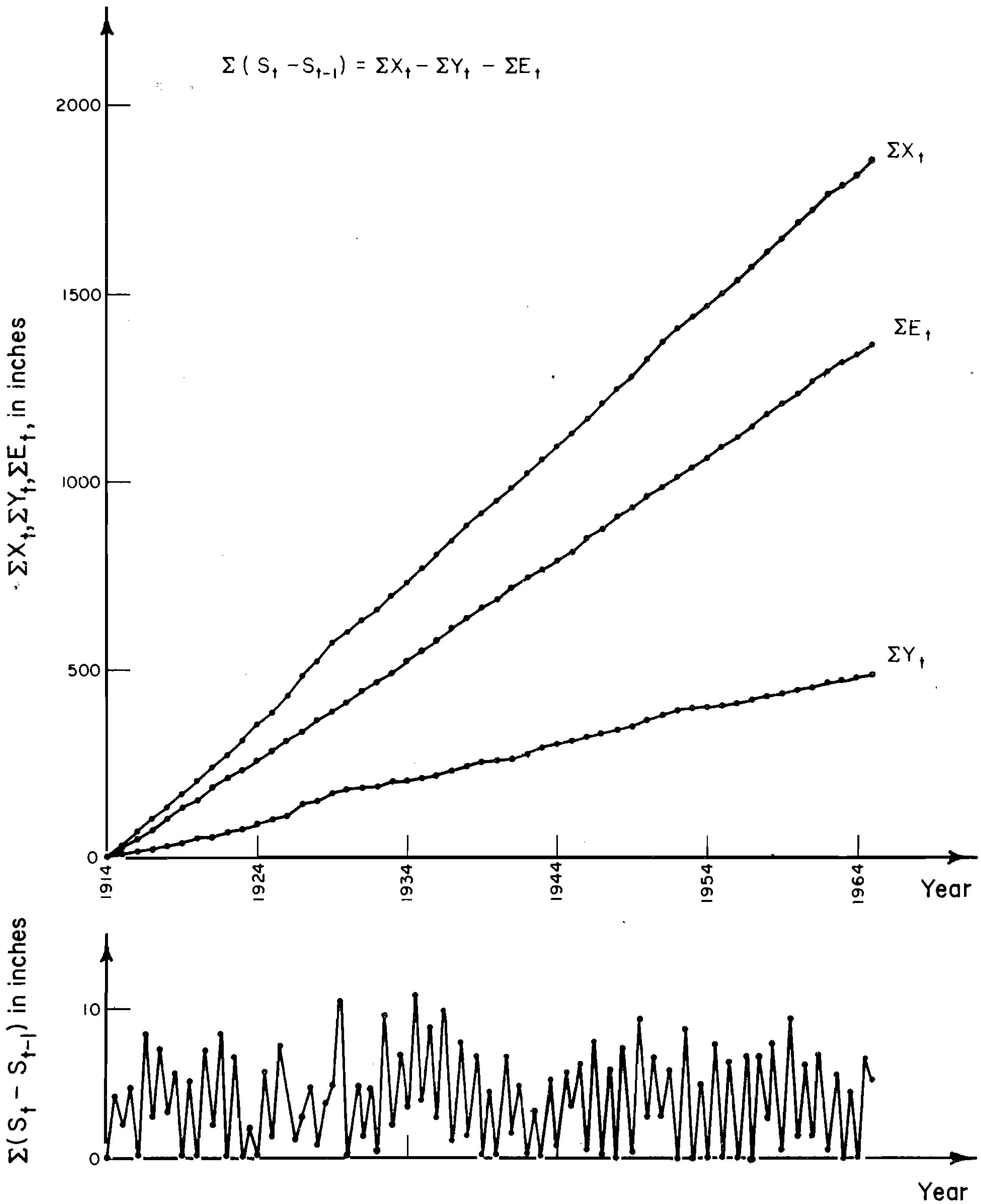

FIG. 2 MASS CURVES OF PRECIPITATION, EVAPOTRANSPIRATION, RUNOFF AND CONCEPTUAL WATERSHED STORAGE

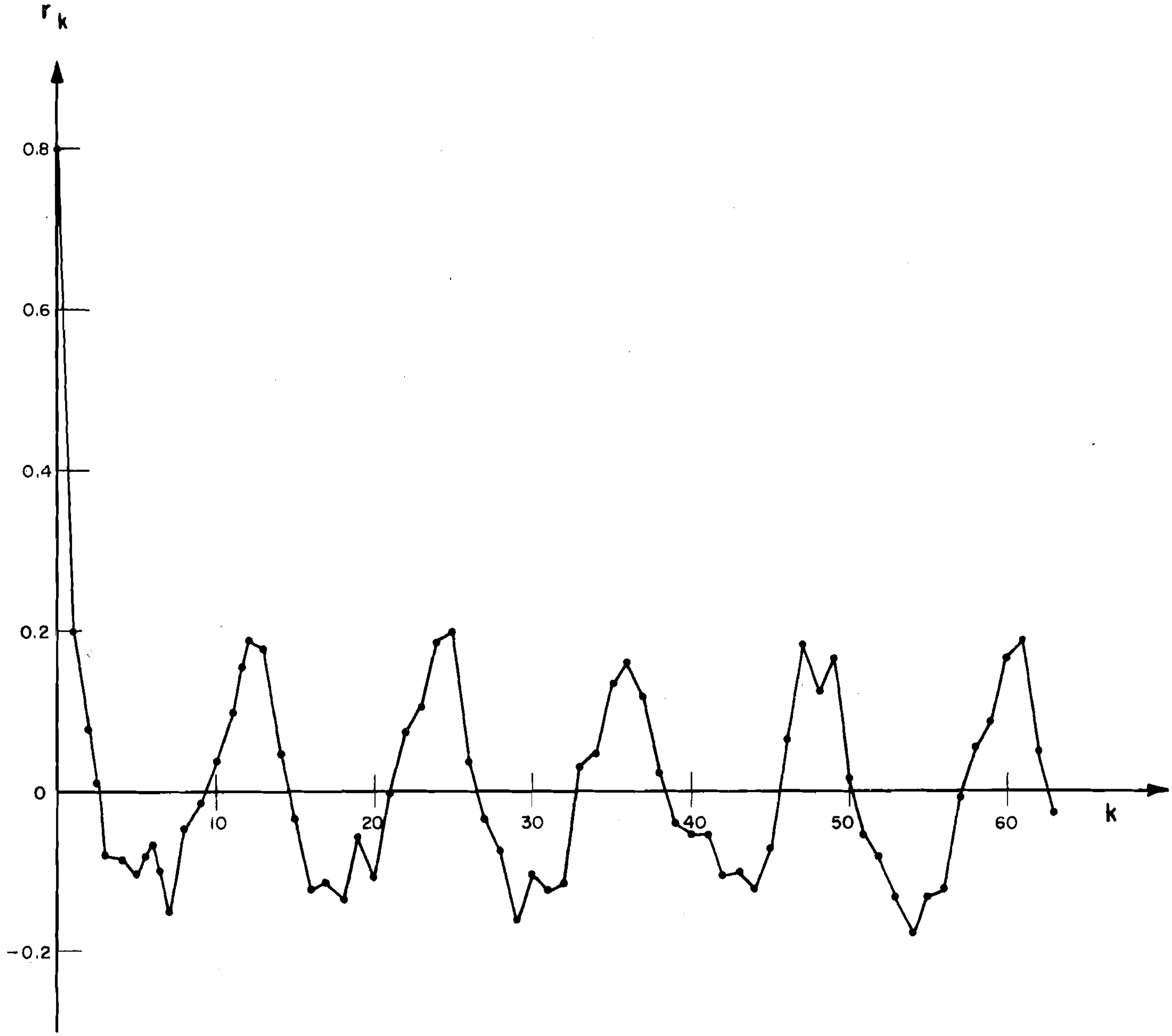

FIG. 3 CORRELOGRAM FOR PRECIPITATION

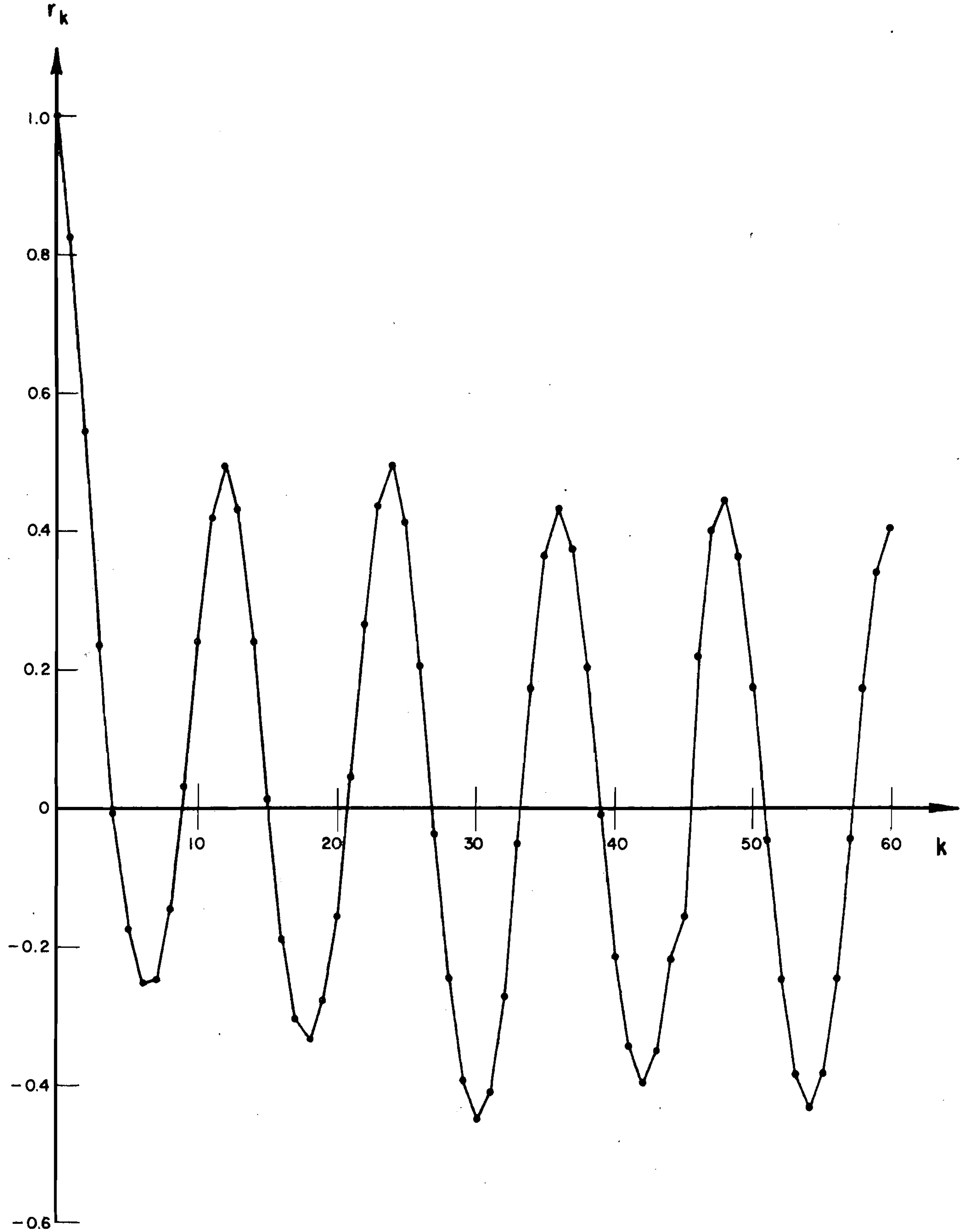

FIG. 4 CORRELOGRAM FOR CONCEPTUAL WATERSHED STORAGE

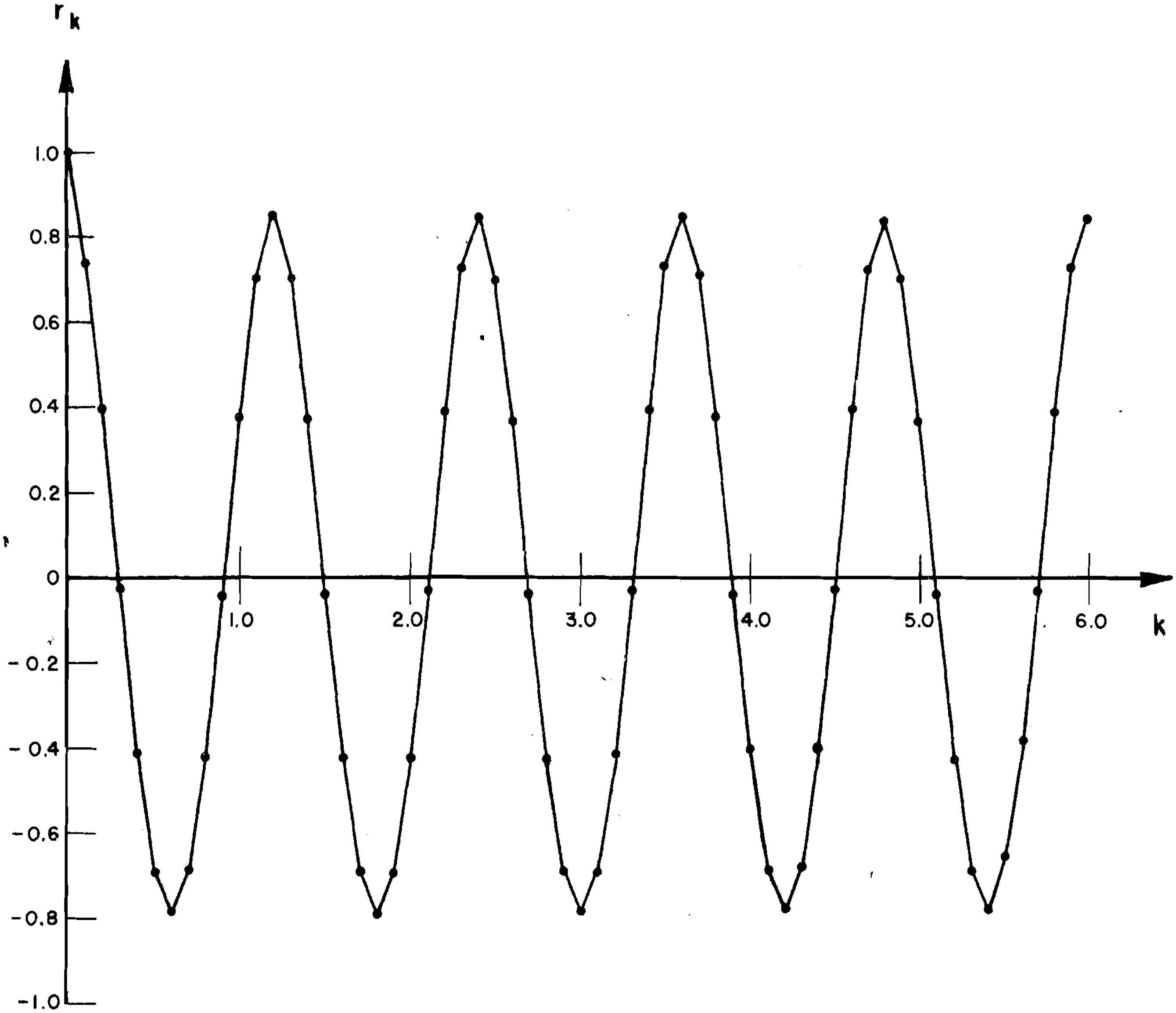

FIG. 5 CORRELOGRAM FOR EVAPOTRANSPIRATION

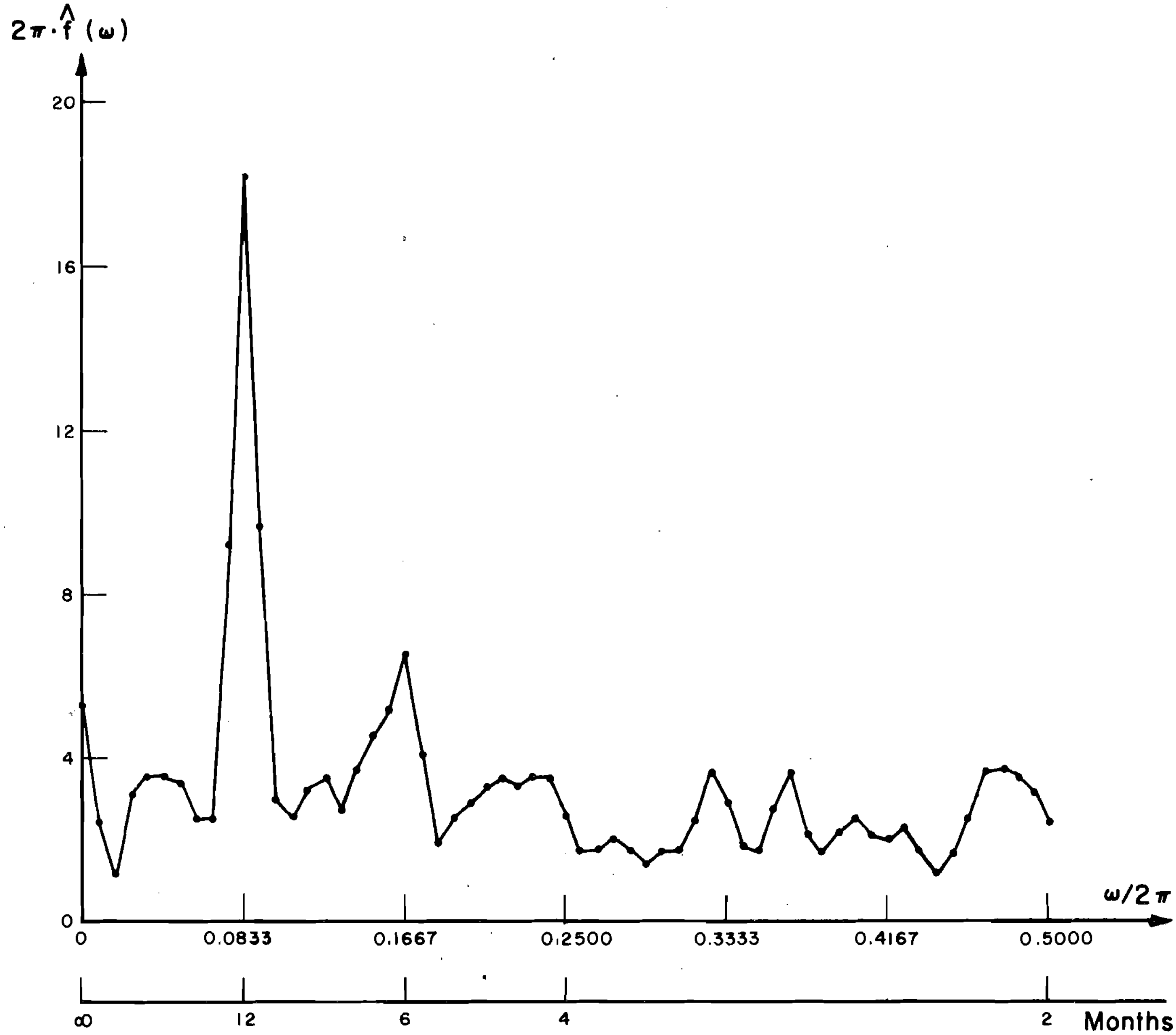

FIG. 6 SPECTRUM OF PRECIPITATION

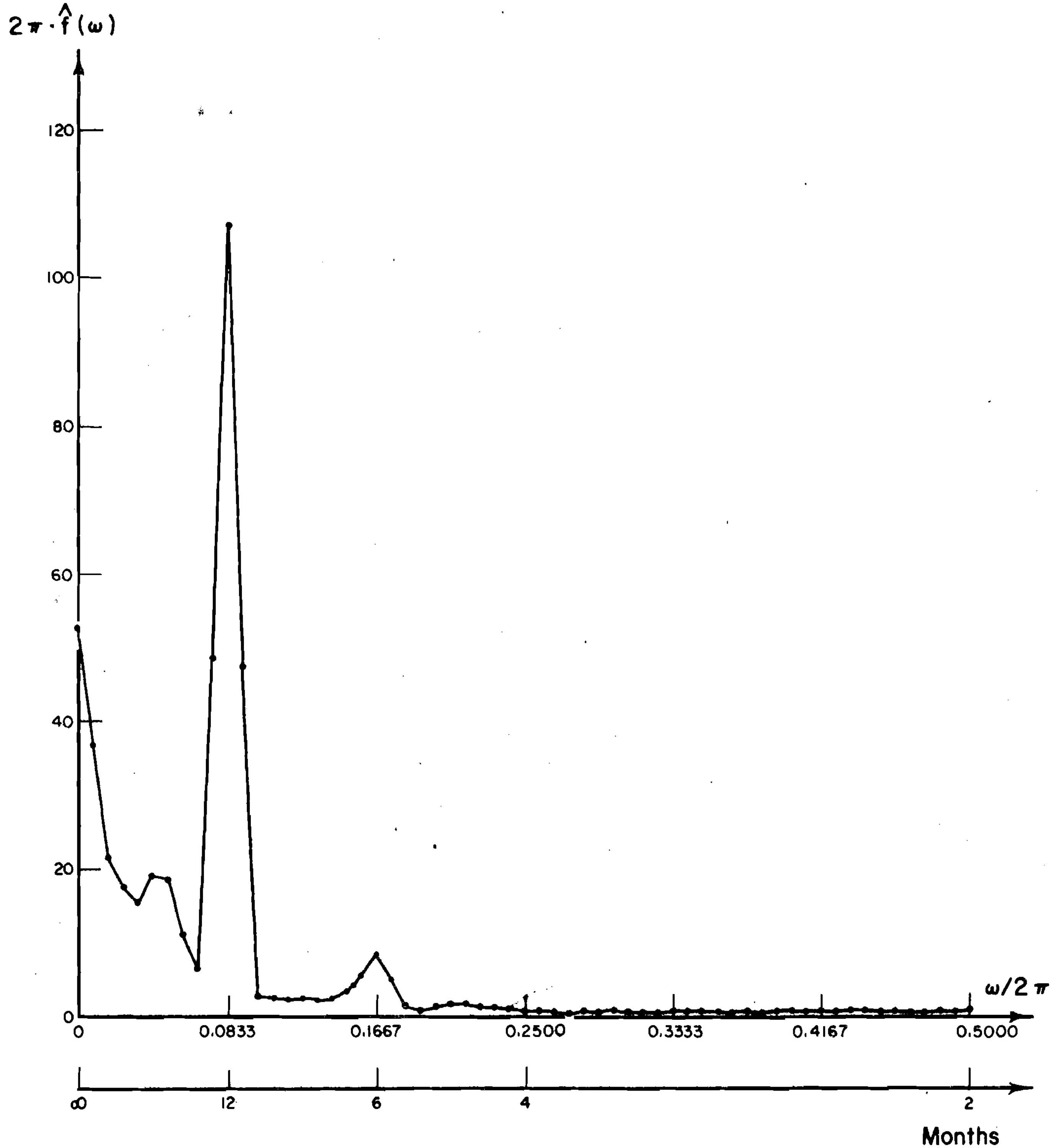

FIG. 7 SPECTRUM OF CONCEPTUAL WATERSHED STORAGE

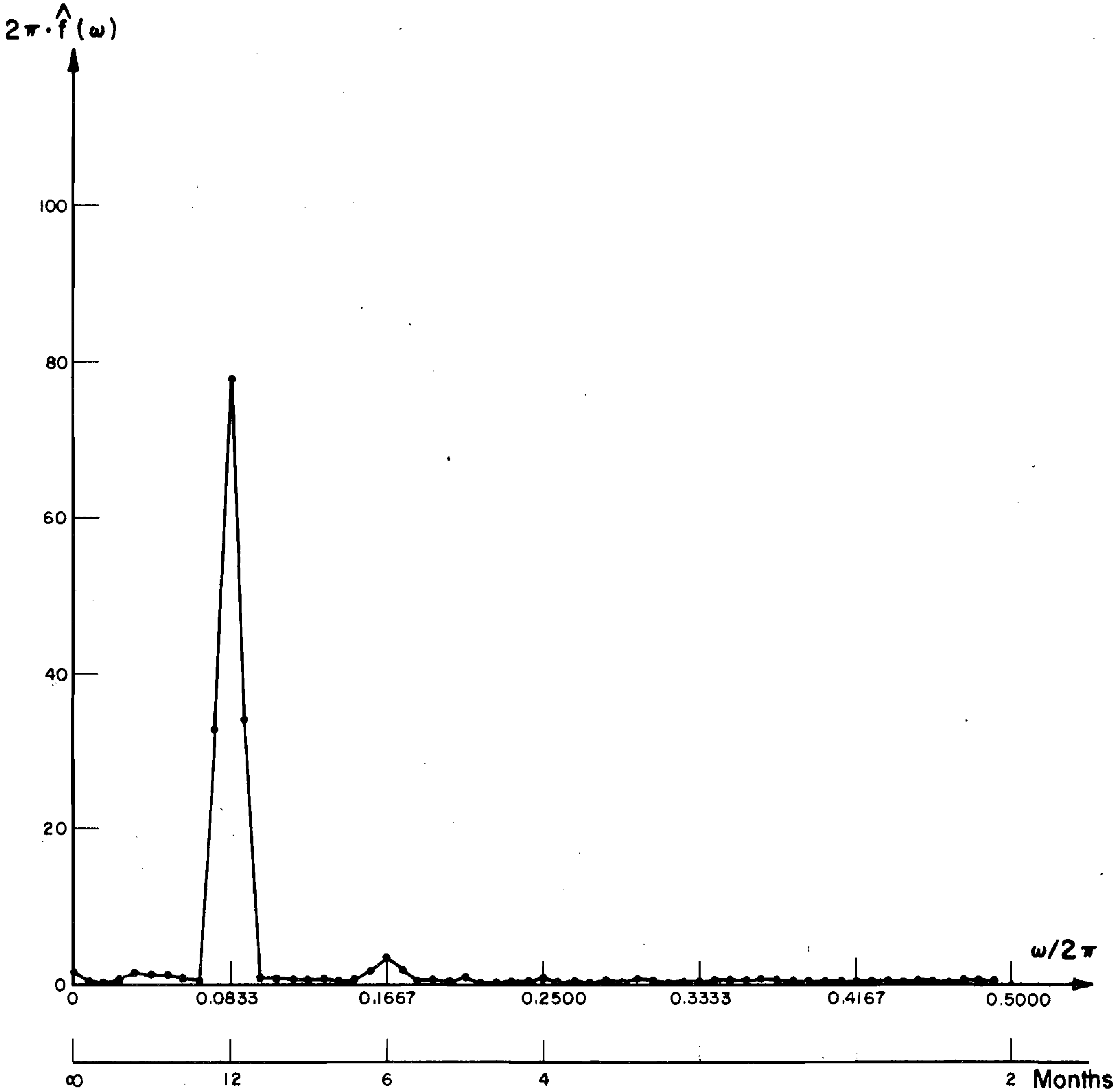

FIG. 8 SPECTRUM OF EVAPOTRANSPIRATION