COMPLETION TECHNICAL REPORT

06

A MULTIVARIATE STOCHASTIC FLOOD ANALYSIS USING ENTROPY

by

VIJAY P. SINGH

P. K. KRSTANOVIC

for

U. S. GEOLOGICAL SURVEY

Department of Interior

Reston, VA 22092

LOUISIANA WATER RESOURCES RESEARCH INSTITUTE

Louisiana State University

Baton Rouge, LA 70803

A MULTIVARIATE STOCHASTIC FLOOD ANALYSIS USING ENTROPY

By

V. P. Singh and P. F. Krstanovic
Department of Civil Engineering
Louisiana State University
Baton Rouge, LA 70803

Louisiana Water Resources Research Institute
Louisiana State University
Baton Rouge, LA 70803

October 1986

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

The principle of maximum entropy (POME) is used to perform a multivariate stochastic flood analysis. By specifying appropriate constraints in terms of means, variances, covariances and cross-covariances, various multivariate exponential distributions are derived. From these distributions, univariate, bivariate and general multivariate stochastic models are then derived. Both continuous and discrete cases are examined. Special emphasis is given to the structure of the matrix of Lagrange multipliers in the model.

The bivariate stochastic model for flood peaks and volumes is investigated for two cases: (1) The peaks and volumes are independent and occur the same number of times. (2) The number of peaks is more than that of volumes in the same time interval. Testing on two Louisiana rivers shows that case (1) is an approximation of case (2). Marginal frequency distributions of peaks, volumes and durations are obtained, first with no restrictions imposed, and then with assumptions of independent occurrences and a high threshold value. Conditional distributions of flood volumes, given peaks and of flood durations, given volumes and peaks are then presented.

ACKNOWLEDGEMENTS

This study was supported in part by funds provided by the Louisiana Water Resources Research Institute through the U.S. Department of Interior, Geological Survey, under the project, "A Multivariate Stochastic Analysis of Flood Magnitude, Duration and Volume." Dr. Helmut Schneider, Department of Quantitative Business Analysis, read the manuscript and made some valuable suggestions. His assistance is gratefully acknowledged.

1. INTRODUCTION

Stochastic approaches to flood analysis have been developed along two major lines: (1) analysis of flood peaks using the theory of probability, and (2) streamflow synthesis using time series analysis.

In the first category are considered the floods exceeding some threshold value X_0 (partial duration series approach) in a given time interval, or their maxima (complete duration series). Complete duration series primarily treats flood extremes and the time interval is usually a year, but can be longer. Partial duration series usually considers smaller time intervals. Both partial and complete duration series treat one stochastic process of flood characteristics and appear as realizations above the threshold value in a given time interval. The relationship among realizations of the process is very often neglected. Mostly presented are various univariate and multivariate frequency distributions.

The relationship among all realizations of the stochastic process is treated in time series analysis. Flood characteristics are observed either at specific instants of time (discrete time series) or sometimes continuously (continuous time series). Observations are usually in terms of discharges or water level elevations of streamflow. All other characteristics are calculated using these observations. A simple model for streamflow in this category is represented by a single probability function $f(x,\theta)$ with θ as row vector of the parameters $(\theta_1, \theta_2, \ldots, \theta_n)$ valid for the positions in time $t_1, t_2, \ldots, t_n$. θ can, for example, be streamflow discharge while each component θ_i $(1 = i, \ldots, n)$ is particular realization of that discharge at specific time t_i. This realization is just one of infinite possible realizations and is therefore random. θ is random vector and its components constitute the random sequence. In time series, this sequence is often represented by second moments such as variances or covariances (autocovariance for one realization and cross covariances for more realizations). Furthermore, one assumes second order stationarity and studies only the realizations of stationary random sequences.

The above two categories have been developed almost independently, but how they are related to each other is not clear except that they treat different aspects of the same subject. Recently, entropy theory has yielded encouraging results in both categories. In this study, we provide a brief historical review of this theory and extend it to flood analysis.

Specific objectives of this study are:

(1) To derive a multivariate stochastic model for flood analysis using the principle of maximum entropy (POME).

(2) To show that the model attempts to bridge the two categories based on theory of probability and time series analysis.

(3) To examine the mathematical form of the Lagrange multipliers.

(4) To investigate simplified forms of the model.

(5) To test the model using real world data.

2. HISTORICAL PERSPECTIVE

Both types of approaches to flood analysis have been widely treated in recent years. In the first category, the models of complete duration series are primarily empirical, based on the criteria of best fit to the data by means of Gumbel or Log-Pearson type III distribution. These have been used extensively prior to the 1970's because of ease in application. They consider only the biggest flood in the time interval (say, 1 year) and neglect others that may have larger peaks and volumes than the maximum smaller floods in other years. The partial duration series models correct that disadvantage by considering all flood peaks exceeding a given threshold value. The partial duration series models are more versatile than the complete duration series models.

Borgman (1963), and Shane and Lynn (1964) found the distribution function of the flood peaks in a certain time interval (0,t) assuming their number to be a time-homogeneous Poisson process. The flood magnitudes were assumed to follow an exponential distribution. Todorovic and Zelenhasic (1970) considered the case where the number of exceedances follows a nonhomogeneous Poisson process with the sequence of flood discharges kept exponentially independently identically distributed (iid). Todorovic and Rousselle (1971) relaxed independency assumption while dividing the water year into different seasons and keeping the iid assumption only within the season being considered. Todorovic and Woolhiser (1972), using the same assumptions, derived the distribution function for the time of the largest peak occurrence. Gupta, et al. (1976) found the joint distribution of the largest flood peaks and their associated times.

Todorovic (1978) generalized the previous results, and derived conditional distribution of the flood magnitude, given the time of occurrence. In spite of these generalizations and some further contributions (North, 1980, Ashkar and Rousselle, 1982), Todorovic retained the iid assumption, and consequently these models are valid only for a high threshold value. This assumption neglects seasonal variations, and the choice of a high threshold value may render a model invalid for lower threshold values.

A theoretically more general model was developed by Kavvas and co-workers (Kavvas and Delleur, 1975; Kavvas, 1982; Kavvas, et al., 1983) treating flooding as the clustering phenomenon and its mechanisms as centers of clusters of flood peaks. Although the disadvantages of the models developed by Todorovic and co-workers were removed, this model is not widely used because of its mathematical complexity.

In the second category are various models of time series analysis. The simplest ones are ARMA models (Thomas and Fiering, 1962; Yevjevich, 1963; Box and Jenkins, 1976). These models have been widely used for modeling long-term dependent hydrologic sequences assuming that flood mechanisms can be represented by either autoregressive (AR) or moving average (MA) components of the model. They do not simulate short-term properties satisfactorily.

Entropy, the subject matter of this report, was first defined by Bolzmann (1872) who used it in thermodynamics. A mathematical theory of entropy was developed by Shannon (1948a, 1948b) who used it in communication theory. A valuable contribution was made by Jaynes (1957) who expressed the principle of maximum entropy (POME). This principle has been applied to virtually all technical fields, for example, geophysics, radio-astronomy, communication theory, structural reliability, linguistics, economics, statistical physics, and hydrology. Wide ranging applications have been reported in workshops on entropy held regularly since 1981. Their proceedings serve as a guide to both mathematics and technology of entropy theory (Smith and Grandy, 1985).

Entropy has only recently been applied to hydrologic analyses. Sonuga (1972, 1976) used POME successfully in frequency analysis and rainfall-runoff modeling. Amorocho and Espildora (1973) used it for assessment of uncertainty in hydrologic frequency analysis. Jowitt (1979) used POME in parameter estimation. The most comprehensive derivation, using POME, of many frequency distributions used in hydrology was presented by Singh, et al. (1985, 1986). This work showed that by using POME, any probability distribution function (pdf) of exponential type can be obtained. Singh and Krstanovic (1985, 1986) applied POME to sediment yield, phosphorus loading and rainfall networks design (1986). Some new perspectives on applications of entropy in water resources have been discussed by Rajagopal, et al. (1986). They summarized advantages as well as disadvantages of the theory and its potential in hydrology, e.g., maximum entropy histograms originally derived by Collins and Wragg (1977).

In time series analysis, entropy was applied first in the astronomy (Ables, 1974). A landmark contribution was made by Burg (1975) that changed the theory of spectral analysis. This work has found useful applications in reconstruction of pictures, signals, waves, etc. These new perspectives are best summarized by Jaynes (1982) who also suggested guidelines for further research.

3. MATHEMATICAL PRELIMINARIES

3.1 Entropy and its Properties

Entropy is defined as expectation of information or measure of uncertainty. Let S be a system of events, E_1, E_2, ..., E_n, and $p(E_k)$ be the probability of k-th event to occur. The entropy of the system S is:

$$H(S) = - k \int_{\Omega} p(x) \ln p(x) \, pX \qquad (1)$$

where k is an arbitrary positive constant, Ω probability space, and p(x) is pdf associated with the random variable X from the probability space Ω. From now onwards, k is absorbed into the base of the logarithm, and (1) is written as:

$$H(S) = -\int_{\Omega} p(x) \ln p(x)\, dx \qquad (2)$$

Equations (1) and (2) represent the entropy of a continuous distribution with one random variable. For m-dimensional distribution $p(X_1, X_2, \ldots, X_m)$, the entropy is defined as:

$$H(S) = -\int \ldots \int p(x_1, \ldots, x_m) \ln p(x_1, \ldots, x_m)\, dx_1 \ldots dx_n \qquad (3)$$

For two random variables X_1 and X_2, the joint and conditional entropies of (X_1,X_2) can be defined as

$$H(X_1,X_2) = -\iint p(x_1,x_2) \ln[p(x_1,x_2)]\, dx_1\, dx_2 \qquad (4a)$$

$$H(X_2|X_1) = -\iint p(x_1,x_2) \ln[\frac{p(x_1,x_2)}{p(x_1)}]\, dx_1\, dx_2 \qquad (4b)$$

$$H(X_1|X_2) = -\iint p(x_1,x_2) \ln[\frac{p(x_1,x_2)}{p(x_2)}]\, dx_1\, dx_2 \qquad (4c)$$

The properties of entropy for continuous variables are:

(1) If X is confined to a certain volume v in its probability space Ω then H(X) has its maximum log v when p(X) is constant (= 1/v) in that volume. Entropy maximum corresponds to the uniform pdf in probability space Ω. We say that nothing is known about X and the uncertainty is the highest.

(2) If one of the events (E_k) is absolutely certain, the entropy achieves its minimum (H(X) = 0) and the uncertainty is the lowest.

(3) Joint entropy of two random variables (rv), X,Y is smaller or equal to the sum of individual entropies. Equality is achieved only if two rv's are independent.

(4) Entropy decreases with increasing knowledge about the variables of the system. Consequently, $H(Y) > H(Y|X)$ and $H(X) > H(X|Y)$.

(5) When assignment of the probabilities to various events is made by maximizing entropy subject to given information, then that assignment is minimally prejudiced as it excludes all assumptions and makes maximum use of the given information (POME). This then defines the principle of maximum entropy. This procedure prescribes adequate information and uses that information as constraints in order to describe the probability distribution. Usually these constraints are mean values of different statistics A_k. From now on we write A_k for the quantity and A_k' for its expected value:

$$A_k' = \int dx\ p(x)\ A_k(x) \tag{5}$$

Equation (5) is valid for both one-dimensional and multi-dimensional cases. Equation (2) or (3) is then maximized subject to (5), usually by the method of Lagrange multipliers. The resulting distribution p(x) will be most consistent with the information given by (5).

(6) Entropy concentration theorem (Jaynes, 1979) gives distribution of other entropies around the maximum entropy. It will not be used in this report.

For practical purposes, we normally deal with discrete entropy. To that end, mathematical equivalents of (1) - (4) are:

$$H(S) = -k \sum_{i=1}^{n} p(x_i) \ln[p(x_i)] \tag{6a}$$

$$H(S) = - \sum_{i=1}^{n} p(x_i) \ln[p(x_i)] \tag{6b}$$

$$H(S) = - \sum_{i=1}^{n_1} \ldots \sum_{i=1}^{n_n} p(x_1, \ldots, x_n) \ln p(x_1, \ldots, x_n) \tag{6c}$$

$$H(X_1,X_2) = - \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} p(x_{1i},x_{2j}) \ln p(x_{2i},x_{2j}) \tag{6d}$$

$$H(X_2|X_1) = -\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} p(x_{1i},x_{2j})\ \ln[\frac{p(x_{1i},x_{2j})}{p(x_{1i})}] \quad (6e)$$

$$H(X_1|X_2) = -\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} p(x_{1i},x_{2j})\ \ln[\frac{p(x_{1i},x_{2j})}{p(x_{2j})}] \quad (6f)$$

Here n is the number of possible events for one-dimensional random variable X; n_1 and n_2 are number of events for rv's X_1 and X_2, respectively; n_1, ..., n_n, are number of events for rv's in multi-dimensional space $\Omega = \{X_1, X_2, \ldots, X_m\}$.

Discrete entropy preserves all the properties defined for its continuous counterpart. However, there are two important differences: (1) Continuous entropy changes under coordinate transformation. There is no such problem for the discrete counterpart. (2) Discrete entropy uses the frequencies of possible occurrences or probability of long sequences. Continuous entropy uses probability density for a long series of samples. In this study, we treat both continuous and discrete entropies.

3.2 Entropy and Time Series

The best application of entropy to time series analysis can be found in Jaynes (1982) and Shilling and Gull (1984). Let X be a random variable with realizations in time interval (-T/2,T/2). From now onwards, we call X as a random process if it has continuous realizations in (-T/2,T/2) and write $X(t) = \{x(t), -T/2 \leq T \leq T/2\}$. Consequently, we use continuous definition of entropy. X is called the random sequence if it has only finite number of realizations in (-T/2,T/2), and we write $X(t) = \{x_{-T/2}, \ldots, x_0, \ldots, x_{T/2}\}$, and consequently discrete definition of entropy is used. This discussion encompasses only one time series where we measure certain flood properties (e.g., discharges, peaks, volumes, etc.). Quite often we need to monitor two or more flood properties in the same time interval. Thus, we measure two or more time series. In that case, we define X as multi-dimensional random sequence $X = \{X_1, X_2, \ldots, X_m\}$, where m is the number of properties measured and $X_i(t)$ is one random variable defined either as continuous or as discrete.

3.3 Entropy and Floods

Entropy can measure objectively the information content of a hydrological process. From its very definition, it is unbiased as it exclude all our assumptions and yields results consistent with respect to the information provided. These results may include pdf obtained by applying POME to any space-time function or pdf consistent with description in Section 3.2. Specifically, entropy can be used to analyze a set of simultaneous hydrological sequences: peaks, volumes, durations, etc. Each sequence carries its own information content on one time scale, e.g., realizations of flood peaks in one hydrological year. The various sequences interact among themselves and may not be independent. For example, in a given time interval (e.g., one month) flood peaks, durations and volumes will be closely related; floods of bigger volumes can be expected to have longer durations and higher peaks. This dependence among sequences can be specified as a constraint in the model using entropy. Furthermore, we may examine how the information of one sequence (e.g., peaks) influences the realization of another sequence (e.g., durations). This transfer of information might help improve predictions inside the given sequence (durations). Many scenarios are possible for building models, with the freedom to examine both space and time interactions uncharacteristic of other models.

4. MULTIVARIATE STOCHASTIC ANALYSIS

4.1 Continuous Description

4.1.1 Univariate Analysis

Let X be a one-dimensional random process of any flood characteristic (e.g., peak, volume or duration) and T the length of historical record from which the measurements are available. Then

$$X(t) = \{X_1(t)\} = \{x_1(t), -T/2 \leq t \leq T/2\} \quad (7)$$

The elements in the first bracket represent realizations of the flood characteristic. For a stationary process, the dependency among the elements is measured only by lag, and not by the position on the time-axis itself. Our objective is to derive a univariate pdf associated with the maximum of entropy H(X) as defined in (2). Sufficient information about the process is given in terms of variance and autocovariances of the process:

$$c(k) = \frac{1}{T} \int_{-T/2}^{T/2} [x(t) - \bar{x}]\ [x(t+k) - \bar{x}]\ dt \quad (8)$$

where $\bar{x}$ is the mean defined as

$$\bar{x} = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{+T/2} x_1(t)\ dt \quad (9)$$

and $-m_1 \leq k \leq m_1$, where $-T/8 \leq m_1 \leq T/8$. Since c(k) = c(-k), the maximum number of different constraints is $m_1 + 1$. The constraints for POME are conveniently defined as:

$$A(k) = T\ c(k) \quad (10)$$

Taking limit of T to ∞, and applying L'Hopital's rule

$$A(k) = \int_{-\infty}^{+\infty} [x(t) - \bar{x}]\ [x(t+k) - \bar{x}]\ dt \quad (11)$$

From data we have numerical values of A'_0, A'_1, ..., A'_{m1}. Maximizing (2) subject to (11) and

$$\int_{-\infty}^{+\infty} p(x)\ dx = 1 \tag{12}$$

we obtain pdf consistent with maximum entropy as:

$$p(X) = \frac{1}{Z(\Lambda)} \exp[-\ \Lambda A] \tag{13a}$$

where

$$\Lambda = [\lambda_{-m_1},\ \ldots,\ \lambda_0,\ \ldots,\ \lambda_{m_1}] \tag{13b}$$

$Z(\Lambda)$ is the partition function usually determined from (12) and A is the vector of the constraints A(k):

$$A = \begin{bmatrix} A(-m_1) \\ \cdot \\ \cdot \\ \cdot \\ A(0) \\ \cdot \\ \cdot \\ \cdot \\ A(m_1) \end{bmatrix} \tag{14}$$

The pdf of X can be rewritten as:

$$p(X) = \frac{1}{Z(\Lambda)} \exp\{-\sum_{k=-m_1}^{+m_1} \lambda_k \int_{-\infty}^{+\infty} [x(t) - \bar{x}]\ [x(t+k) - \bar{x}]\ dt\} \tag{15}$$

At 0-th lag, (15) simplifies to:

$$p(X) = \frac{1}{Z(\lambda_0)} \exp\{-\ \lambda_0\ T\ \sigma_x^2\} \tag{16}$$

where $Z(\lambda_0)$ is determined from:

$$-\frac{\partial}{\partial \lambda_0} \ln Z(\lambda_0) = A_0 \tag{17}$$

$$Z(\lambda_0) = \int_{-\infty}^{+\infty} \exp\{-\ \lambda_0\ T\ \sigma_x^2\}\ dx \tag{18}$$

4.1.2 Bivariate Analysis

Let X be two-dimensional random process of two flood properties X_1 and X_2 (e.g., peaks and volumes) and T defined as before. Then

$$X(t) = \{X_1(t),\ X_2(t)\} = \{[x_1(t),\ -T/2 \le t \le T/2],\ [x_2(t),\ -T/2 \le t \le T/2]\} \quad (19)$$

The elements $x_1(t)$ and $x_2(t)$ represent realizations of peaks and volumes respectively. Usually we create the pair (X_1,X_2) such that its elements are associated: single peak hydrograph combines X_1 and X_2 uniquely in time interval; and multiple peak hydrographs combine unique X_2 and one of the peaks X_1. For a stationary process, the dependence among elements of $X_1(t)$ and $X_2(t)$ is measured only by the time lag.

Our objective is to derive bivariate pdf associated with joint entropy $H(X_1,X_2)$ as defined by (4a). The information about the process $X(t)$ is expressed by autocovariance and cross covariance matrices (c_{ii} and c_{ij}) combined into information data matrix as:

$$C_{X_1,X_2} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \quad (20)$$

This matrix is called a partitioned matrix in multivariate analysis. Its elements are matrices that have the same number of rows and columns. c_{ii} and c_{ij} are of the dimension $(m_1+1) \times (m_1+1)$ where $-T/8 \le m_1 \le T/8$. In autocovariance matrices, c_{11} and c_{22}, m_1 is the number of lags used, while in cross covariance matrices c_{12} and c_{21}, m_1 is the number of lags by which X_1 is behind or in front of X_2. Since flood peaks X_1 will always be in the front, and never behind X_2, c_{12} and c_{21} will be lower and upper triangular matrices, respectively.

These matrices contain the most information about the process:

$$c_{12} = \begin{bmatrix} c(0) & & & \\ c(-1) & & & \\ \vdots & & & \\ c(-k) & & & \\ \vdots & & & \\ c(-m_1) & c(-k) & c(-1) & c(0) \end{bmatrix} \quad (-m_1 \leq k \leq 0)$$

$$c_{21} = \begin{bmatrix} c(0) & c(1) \; \dots & c(k) \; \dots & c(m_1) \\ & & & \vdots \\ & & & c(k) \\ & & & \vdots \\ & & & c(1) \\ & & & c(0) \end{bmatrix} \quad (0 \leq k \leq m_1) \qquad (21)$$

Each element c(k) of these matrices is defined as:

$$c(k) = \frac{1}{T} \int_{-T/2}^{+T/2} [x_1(t) - \bar{x}_1]\,[x_2(t+k) - \bar{x}_2]\,dt \qquad (22)$$

where $\bar{x}_1$ and $\bar{x}_2$ represent means of $x_1(t)$ and $x_2(t)$ defined as in (9). The information constraints A(k) in entropy formalism can be defined with any coefficients proportional to c(k). However, the scalar product λA_k and the final conclusion will be independent of our choice. Taking $A(k) = T\,c(k)$ and letting limit of T to ∞,

$$A(k) = \int_{-\infty}^{+\infty} [x_1(t) - \bar{x}_1]\,[x_2(t+k) - \bar{x}_2]\,dt \qquad (23)$$

The matrix elements in (21) are now replaced by A(k) of (23), and the partitioned matrix in (20) contains matrices A_{ii} and A_{ij} and A_{ij} $(i \neq 1)$. Information from (23) is available as numerical values A'_k defined in (5). We have m_1+1 of these values or:

$A'_0, A'_1, \dots, A'_m$

Maximizing (4a) subject to (12) and

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} p(x_1,x_2)\, dx_1\, dx_2 = 1 \tag{24}$$

we obtain:

$$p(X_1,X_2) = \frac{1}{Z(\Lambda)} \exp[- \sum_{k=0}^{m_1} \lambda_k \int_{-\infty}^{+\infty} [x_1(t) - \bar{x}_1]\, [x_2(t+k) - \bar{x}_2]\, dt] \tag{25a}$$

Considering all possible interactions among any two flood properties, one can generalize (25) as:

$$p(X_1,X_2) = \frac{1}{Z(\Lambda)} \exp[- \sum_{i=1}^{2} \sum_{j=1}^{i} \sum_{k=-m}^{+m} \lambda_k \int_{-\infty}^{+\infty} [x_i(t) - \bar{x}_i][x_j(t+k) - \bar{x}_j]\, dt] \tag{25b}$$

In (25a) and (25b) $Z(\Lambda)$ is the partition function usually determined from (24) and Λ is a row vector or

$$\Lambda = [\Lambda_1,\Lambda_2] \tag{26}$$

and

$$\Lambda_i = [\lambda_0, \lambda_1, \ldots, \lambda_{m_1}]$$

Equation (25) can be written as:

$$p(X_1,X_2) = \frac{1}{Z(\Lambda)} \exp\{\Lambda\} \begin{bmatrix} \int_{-\infty}^{+\infty} [x_1(t) - \bar{x}_1][x_2(t) - \bar{x}_2]dt \\ \cdot \\ \cdot \\ \cdot \\ \int_{-\infty}^{+\infty} [x_1(t) - \bar{x}_1][x_2(t+m_1) - \bar{x}_2]dt \end{bmatrix} \tag{27}$$

We note that λ_0 is strongly associated with covariance, λ_1 with cross-correlation on lag apart, etc. (27) can be written as:

$$p(X_1,X_2) = \frac{1}{Z(\Lambda_v)} \exp\{- \Lambda_v^T A_v\} \tag{28}$$

where subscript v denotes vector and superscript T denotes its transpose.

For the cases with strong correlation at associated time intervals and weak correlations for $k > 0$, the covariance term with λ_0 will be dominant. Then (25) and (28) will simplify to:

$$p(X_1,X_2) = \frac{1}{Z(\lambda_0)} \exp[-\lambda_0 T \operatorname{cov}(X_1,X_2)] \qquad (29)$$

for any associated pair (X_1,X_2). This assumption introduces independency among different occurrences (x_1,x_2), with $x_1 \varepsilon X_1(t)$ and $x_2 \varepsilon X_2(t)$.

From the bivariate process, one can derive the univariate case when two random variables X_1 and X_2 are equal. For that case, data matrix (2) (dimension 2 x 2) becomes $C_{x1} = [c_{11}]$ (dimension 1 x 1), where c_{11} is the autocovariance matrix with autocovariance elements defined in (8). Similarly, the information constraint A(k), dependent on the crosscovariance c(k) of the lag k, becomes dependent on the autocovariance of the same lag. Since autocovariances are symmetric with respect to zero, Λ of (26) increases its element by m_1 and becomes Λ defined in (13b). Consequently, the pdf of the bivariate process becomes the univariate pdf.

4.1.3 Multivariate Analysis

We extend the univariate and bivariate analyses to multi-dimensional random process X of a number of flood properties, e.g., peaks, volumes, durations, interarrival times, etc. With T defined as before,

$$X(t) = \{X_1(t), \ldots, X_m(t)\} \qquad (30)$$

where m is the number of measured properties. All properties are defined in the same time interval creating the total of m simultaneous time series.

Our objective is to derive multivariate pdf associated with multivariate entropy $H(X_1, X_2, \ldots, X_m)$ as defined in (3). The information about the process can be expressed in matrix form as

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1m} \\ c_{21} & & & \vdots \\ \vdots & & & c_{m-1,m} \\ c_{m1} & \cdots & c_{m,m-1} & c_{mm} \end{bmatrix} \tag{31}$$

We note that this matrix expresses dependencies across the space dimension, while each element is itself a matrix responsible for time dependencies: c_{ii} $(i = 1, \ldots, m)$ are autocorrelation matrices of the dimension $(m_1+1) \times (m_1+1)$. m_1's will be the same if with each flood we associate only one unique property (e.g., highest peak, one volume, one duration, time to the highest peak, etc.). For multiple peak hydrographs, autocorrelation matrix associated with the peak will be of the highest dimension. c_{ij} $(i,j = 1, \ldots, m;\ i \neq j)$ are cross-correlation matrices of $(m_1+1) \times (m_1+1)$ dimension. Both parts of the matrices (below and above the main diagonal) may be dominant so we are not necessarily restricted to the lower or upper triangular matrix. A matrix is written as

$$c_{ij} = \begin{bmatrix} c(0) & c(1) & \cdots & c(m_1) \\ c(-1) & & & \vdots \\ \vdots & & & c(1) \\ c(-m_1) & \cdots & c(-1) & c(0) \end{bmatrix}_{i,j} \tag{32}$$

where elements are cross-covariances defined as

$$c(k) = \begin{cases} \frac{1}{T} \int_{-T/2}^{+T/2} [x_i(t) - \bar{x}_i]\, [x_j(t+k) - \bar{x}_j]\, dt, \text{ for } k \geq 0 \\ \frac{1}{T} \int_{-T/2}^{+T/2} [x_j(t) - \bar{x}_j]\, [x_i(t-k) - \bar{x}_i]\, dt, \text{ for } k \leq 0 \end{cases} \tag{33}$$

where $\bar{x}_i$ and $\bar{x}_j$ are the means associated with $X_i(t)$ and $X_j(t)$ in $(-T/2, T/2)$. We define information constraint as the matrix:

$$A = T\, C \tag{34}$$

with each element $A_{ij} = T\, c_{ij}$ $(i,j = 1, \ldots, m)$.

The elements of A_{ij} are, in turn, $A(k) = T\, c(k)$. In total we may have $(2\, m_1+1)\cdot m$ possible constraints. By maximizing (3) subject to (34) and

$$\int_{-\infty}^{+\infty} \ldots \int_{+\infty}^{+\infty} p(x_1, x_2, \ldots, x_m)\, dx_1\, dx_2 \ldots dx_m = 1 \tag{35}$$

we obtain:

$$p(X_1, X_2, \ldots, X_m) = \frac{1}{Z(\Lambda)} \exp[-\, \Lambda A] \tag{36}$$

where Λ is the row vector or

$$\Lambda = [\Lambda_1\ \Lambda_2 \ldots \Lambda_m] \tag{37}$$

and $\Lambda_i = [\lambda_{m_1}, \ldots \lambda_0, \ldots \lambda_{-m_1}]$

$Z(\Lambda)$ is the partition function and is usually determined from (35) or

$$-\frac{\partial}{\partial \lambda_k} \ln Z(\lambda_k) = A_k \tag{38}$$

Assuming all possible interactions across space and time, (36) can be written as:

$$p(X_1, X_2, \ldots, X_m) = \frac{1}{Z(\Lambda)} \exp\{-\sum_{i=1}^{m} \sum_{j=1}^{i} \sum_{k=-m_1}^{+m_1} \lambda_k \int_{-\infty}^{+\infty} [x_i(t) - \bar{x}_i]\, [x_j(t-k) - \bar{x}_j]\, dt\} \tag{39}$$

For the special case when we have strong correlation at lag 0 and weak correlation for all other lags λ_0 is dominant in each Λ_i matrix so that

$\Lambda = [\lambda_{0(1)}\ \lambda_{0(2)} \cdots \lambda_{0(m)}]$

and

$$p(X_1,X_2,\ldots,X_m) = \frac{1}{Z(\Lambda)} \exp\{- \sum_{i=1}^{m} \sum_{j=1}^{i} \lambda_0\ T\ \mathrm{cov}(X_i,X_j)\} \quad (40)$$

for any associated m-dimensional random process.

Multivariate random process simplifies to the bivariate and univariate cases for m = 2 and m = 1, respectively: simplification of (31) to (20) and to matrix c_{11} given by the elements of (8); simplification of the multivariate information constaint A(k) to the bivariate and univariate counterparts: simplification of (39) to (25b) and (15).

A major problem in continuous description is determination of the Lagrange multipliers since the first partition function $Z(\Lambda_v)$ for bivariate and univariate cases, and $Z(\Lambda)$ for multivariate case must be explicitly determined. This involves precise solution to the continuous definition of either autocovariances or cross-covariances, that is still not tractable. One way out is to resort to the discrete case taking X(t) to be random sequence with finite number of observations.

4.2 Discrete Description

4.2.1 Univariate Analysis

Let X be a stationary random sequence of some selected flood property (e.g., flood peaks)

$X(t) = \{x_0, x_1, \ldots, x_T\}$

The constraints can then be written as autocorrelation matrix C with m_2+1 distinct values c(k) or

$$c(k) = \frac{1}{T+1} \sum_{j=0}^{T-k} (x_j - \bar{x})\ (x_{j+k} - \bar{x}) \quad (41)$$

where $\bar{x}$ is the sequence mean and k is the lag ($0 \leq k \leq m_2$, and $m_2 \leq T/4$). Let information constraint be defined as:

$$A(k) = (T+1)\ c_k \tag{42}$$

which is available as numerical values A'_k from the data. The pdf that yields maximum entropy is:

$$p(X) = \frac{1}{Z} \exp[-\sum_{k=0}^{m_2} \lambda_k\ A(k)] \tag{43}$$

or $$p(X) = \frac{1}{Z(\Lambda)} \exp[-X^* \Lambda X] \tag{44}$$

where $X^* = (x_1-\bar{x},\ \ldots,\ x_T-\bar{x})$, and

$$Z(\Lambda) = \frac{1}{|\Lambda|}$$

where

$$\Lambda = \begin{cases} \lambda_{i-j} & \ldots\ |j-1| \leq m_2 \\ 0 & \ldots \text{ otherwise} \end{cases}$$

We emphasize that (44) is of multivariate Gaussian form. On comparing with multivariate Gaussian distribution:

$$p(X) = \frac{1}{(2\pi)^{(T+1)/2}\ |\Sigma|^{1/2}} \exp[-\frac{1}{2}\ (X-\mu)^*\ \Sigma^{-1}(X-\mu)] \tag{45}$$

The additional constant 1/2 can be taken inside A(k), while Λ matrix is proportional to the inverse of the sample variance-covariance matrix Σ.

Thus, the POME distribution is equivalent to the Gaussian weighted by some constant value $(1/2\pi)^{(T+1)/2}$. Here we assume dependency among all terms in Λ matrix. It is shown in Appendix A that every new Lagrange multiplier in the model adds one new diagonal to the Λ matrix and is inversely proportional to the autocovariance or

$$\lambda_k = 1/c_k \tag{46}$$

and accounts for additional information being introduced with each new lag. When most of the significant information is already introduced with the 0-th lag, λ_0 is dominant and Λ becomes a diagonal matrix:

$$p(X) = \frac{1}{Z(\lambda_0)} \exp[-\lambda_0 \sum_{i=0}^{T} (x_i - \bar{x})^2] \tag{47}$$

and for any member of stationary sequence:

$$p(x_i) = \frac{1}{s_x^2} \exp[-\frac{(x_i - \bar{x})^2}{s_x^2}] \tag{48}$$

where s_x^2 is the sample variance. (48) is valid assuming independent occurrences of members of the stationary sequence X(t).

Very often information is not available in terms of variances, but only as the mean. Then equivalency with Gaussian distribution does not hold and we use pdf that gives maximum entropy, given the mean as the constraint (Singh, et al., 1985):

$$p(x_i) = \frac{1}{\bar{x}} \exp[-\frac{x_i}{\bar{x}}] \tag{49}$$

4.2.2 Bivariate Analysis

Let X be a two-dimensional random sequence of two flood properties X_1 and X_2. Then:

$$X(t) = \{X_1, X_2\} = \{[x_0, x_1, \ldots, x_T]_1, [x_0, x_1, \ldots, x_T]_2\} \tag{50}$$

The discussion about (X_1, X_2) properties for the continuous description in section 4.1.2 remains the same. The constraint that carries the most information from X_1 to X_2 and vice versa is cross-covariance available from the data as the sample cross-covariance:

$$c_{X_1X_2}(k) = \begin{cases} \frac{1}{T+1} \sum_{t=0}^{T-k} [x_1(t) - \bar{x}_1][x_2(t+k) - \bar{x}_2], \; k = 0,1,..,m_2 \\ \frac{1}{T+1} \sum_{t=0}^{T+k} [x_2(t) - \bar{x}_2][x_1(t+k) - \bar{x}_1], \; k = 0,-1,..,-m_2 \end{cases} \quad (51)$$

where $m_2 \leq T/4$.

For any two flood properties X_1 and X_2, the number of occurrences is either the same or one property follows the other. If X_1 is peak and X_2 volume, then the volume sequence is always behind the peak sequence, so (51) reduces to:

$$c_{X_1X_2}(k) = \frac{1}{T+1} \sum_{t=0}^{T+k} [x_2(t) - \bar{x}_2][x_1(t+k) - \bar{x}_1], \; k = 0,-1,..,-m_2 \quad (52)$$

We define suitable information constraint $A_{12}(k) = (T+1)\, c_{x_1,x_2}(k)$. Maximizing joint entropy $H(X_1,X_2)$ subject to m_2+1 available constaints A_{12} $(k = 0, \ldots, -m_2)$ and (24), the following pdf is obtained:

$$p(X_1,X_2) = \frac{1}{Z(\Lambda)} \exp\{- \Lambda_M A\} = \frac{1}{Z(\Lambda)} \exp\{- X_1^* \Lambda_M X_2\} \quad (53)$$

where Λ_M is lower diagonal matrix of Lagrange multipliers

$$\Lambda_M = \begin{bmatrix} \lambda_0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \lambda_{-1} & & & & & & & 0 \\ \cdot & & & & & & & 0 \\ \cdot & & & & & & & \\ \lambda_{-m_2} & & & & & & & 0 \\ \cdot & & & & & & & 0 \\ \cdot & & & & & & & \\ 0 & \ldots & \lambda_{-m_2} & \ldots & \lambda_{-1} & & \lambda_0 & 0 \end{bmatrix} \quad (54)$$

and

$$X_1^* = [x_1(0) - \bar{x}_1, \ldots, x_1(T) - \bar{x}_1]$$

$$X_2^* = [x_2(0) - \bar{x}_2, \ldots, x_2(T) - \bar{x}_2]$$

Considering all possible interactions among any two flood properties, one can generalize (53) as:

$$p(X_1,X_2) = \frac{1}{Z(\Lambda)} \exp[-\sum_{i=1}^{2} X_i^* \Lambda_i \sum_{j=1}^{i} X_j] \qquad (55)$$

Comparing (53) with the pdf of continuous case, we see that Λ_v is substituted by Λ_M and the term A is factorized into two vector products X_1^* and X_2. Equation (54) is the matrix of the Toeplitz form. Some properties of the Toeplitz matrices help us in solving the partition function $Z(\Lambda)$ and determining Lagrange multipliers λ. That approach was first suggested by Jaynes (1982). When $T >> m_2$, from Toeplitz theory the eigenvalues of Λ matrix become:

$$g_j = g(z_j)$$

where z_j are the roots of $z^{T+1} = 1$ on the unit circle or

$$z_j = \exp[2\pi i j/(T+1)] \qquad (56)$$

where $0 \leq j \leq T$. Since the partition function is given as

$$\ln Z = - \sum_{j=0}^{T} \ln g_j + \text{constant}$$

$$= - \ln[g_0 \ldots g_T] \qquad (57)$$

From multivariate analysis in statistics, we know that the product of the eigenvalues of the matrix is equal to the matrix determinant. Thus,

$$\ln Z = - \ln \det \Lambda = - \ln|\Lambda|$$

and

$$Z = \frac{1}{|\Lambda|} \qquad (58)$$

From (56), we define

$$g(z_j) = \sum_{k=0}^{m_2} \lambda_k z^k \qquad (59)$$

For long time base ($T \to \infty$),

$$\ln Z \to - \sum_{j=0}^{m_2} \ln[g_j(\exp \frac{2\pi ij}{T+1})] \tag{60a}$$

The summation becomes integration on the unit circle and roots z_j will be close to one another:

$$\frac{1}{T+1} \ln Z(\lambda_k) = - \frac{1}{2\pi} \int_0^{2\pi} \ln[g(\exp(i\theta))] \, d\theta$$

or

$$\ln Z(\lambda_k) = - \frac{T+1}{2\pi} \int_0^{2\pi} \ln[\sum_{k=0}^{m_2} \lambda_k \exp(ik\theta)] \, d\theta \tag{60b}$$

Finally, the Lagrange multipliers are determined from:

$$(T+1) \; c_{X_1,X_2}(k) = - (\frac{\partial}{\partial \lambda_k}) \ln Z(\lambda_k) \tag{61a}$$

$$c_{X_1,X_2}(k) = \frac{1}{2\pi} \int_0^{2\pi} \frac{\exp(ik\theta)}{\sum_{k=0}^{-m_2} \lambda_k \exp(ik\theta)} \tag{61b}$$

A couple of simple examples are shown in Appendix A.

It is of interest to study simplification of (53) for some special cases. For example, for a bivariate stationary sequence whose pairs are strongly correlated at the same time of occurrence (as is usual for peaks and volumes), while the rest of the pairs are weakly correlated:

$$[(x_{01},x_{02}), (x_{11},x_{12}), \ldots, (x_{T1},x_{T2})] \tag{62}$$

These pairs are strongly correlated. Note that first subscript denotes time of occurrence and the second one the element of the variable (X_1 or X_2).

For that case, the dominant constraint is covariance since

$$c_{x_1,x_2}(0) = \text{cov}(X_1,X_2)$$

and Λ matrix becomes diagonal with λ_0 significant. The pdf for maximum entropy is again (53), while λ_0 is shown to be equal (Appendix A).

$$\lambda_0 = \frac{1}{\cos(X_1,X_2)} \tag{63}$$

Occurrence of any particular pair of (62), using (53) and (63) is:

$$p(x_{i1},x_{j2}) = \frac{1}{\mathrm{cov}(X_1,X_2)} \exp[- \frac{(x_{i1} - \bar{x}_1)(x_{j2} - \bar{x}_2)}{\mathrm{cov}(X_1,X_2)}] \tag{64}$$

We emphasize that (64) is obtained assuming independent occurrence of any pair from (62).

Bivariate random sequence simplifies to the univariate case in the same manner as for the continuous case.

4.2.3 Multivariate Analysis

This analysis is a discrete analog of Section 4.1.3. We study the multi-dimensional random sequence X of a number of flood properties, each property known as the sequence of observations, all with the common time base. Constraints and assumptions are equivalent to (31) - (33), but each element or cross-covariance is defined as in (51) for any two properties i and j (i,j = 1, ..., m). The pdf that gives maximum entropy is of the form as in (36), but now we are able to factor A term into the products of two vectors. Specifically,

$$p(X_1,\ldots,X_m) = \frac{1}{Z(\Lambda_1,\ldots,\Lambda_m)} \exp[- \sum_{i=1}^{m} X_i^* \Lambda_i \sum_{j=1}^{i} X_j] \tag{65}$$

where * denotes the transpose of the vector, m is the number of considered flood properties, X_i^*,X_j are as defined in (55) for any i,j = 1, ..., m, and $Z(\Lambda)$ can be determined by using the same procedure as for the bivariate case. Consider for example three flood properties (m = 3). Then from (53)

$$p(X_1,X_2,X_3) = \frac{1}{Z(\Lambda_1,\Lambda_2,\Lambda_3)} \exp[- X_1^* \Lambda_1 X_1 - X_2^* \Lambda_2(X_1 + X_2) - X_3^* \Lambda_3(X_1 + X_2 + X_3)] \quad (66)$$

For particular triplet in the same time interval $(x_1,x_2,x_3) \varepsilon (X_1,X_2,X_3)$ and assuming $\Lambda_1,\Lambda_2,\Lambda_3$ to be diagonal matrices (λ_0 dominant), (66) can be simplified to:

$$p(x_1,x_2,x_3) = \frac{1}{Z(\lambda_{01},\lambda_{02},\lambda_{03}} \exp[- \sum_{i=1}^{3} \frac{(x_i - \bar{x}_i)^2}{S_{X_i}^2} - \frac{(x_1 - \bar{x}_1)(x_2 - \bar{x}_2)}{\mathrm{cov}(X_1,X_2)} - \frac{(x_1 - \bar{x}_1)(x_3 - \bar{x}_3)}{\mathrm{cov}(X_1,X_3)} - \frac{(x_2 - \bar{x}_2)(x_3 - \bar{x}_3)}{\mathrm{cov}(X_2,X_3)}] \quad (67)$$

This form is convenient for practical applications.

Multivariate random sequence simplifies to bivariate and univariate cases in the same way as for the continuous case.

4.3 Summary

All derived cases of discrete and continuous analysis are presented in Table 1. The table shows the similarities and differences for both cases. Discrete case is easier to use for practical applications since partition function can easily be derived from the Toeplitz theory. Both dependencies in space and time are accounted for by taking nonzero Lagrange multipliers for diagonals surrounding the main one in Λ matrix that connects two variables X_i and X_j. When only one variable X_i is considered, Λ matrix is limited only to the dependency in time. In examining strongly independent events, for example by taking a sufficiently high threshold value for the floods of partial duration series, only diagonal Λ matrices are left easy to use.

Table 1

Analysis	pdf	Information Constraint	Sample Information
Univariate-continuous Var: $X_i(t)$	$p(X_i) = \frac{1}{Z(\Lambda)} \exp\{-[\lambda_{-m_1},\dots,\lambda_{m_1}] \begin{bmatrix} T\,c(-m_1) \\ \vdots \\ T\,s^2 \\ \vdots \\ T\,c(m_1) \end{bmatrix}$ $Z(\lambda_{-m},\dots,\lambda_m) = \int_{-\infty}^{+\infty} d\,X_i \exp[-\Lambda_v A_v]$	$A_v = T \begin{bmatrix} c(-m_1) \\ \vdots \\ s_x^2 \\ \vdots \\ c(m_1) \end{bmatrix}$	autocovariance (continuous) $c(k) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{+T/2} [x(t)-\bar{x}]$ $[x(t+k)-\bar{x}]dt$
Univariate-discrete Var: $X_i(t)$	$p(X) = \frac{1}{Z(\Lambda)} \exp[-X^* \Lambda X]$ $\Lambda = \begin{bmatrix} \lambda_0 & \lambda_1 & \dots & \lambda_{m_1} & 0 \\ \lambda_1 & & & & \lambda_m \\ \vdots & & & & \vdots \\ \lambda_{m_1} & & & & \lambda_1 \\ 0 & \lambda_{-m_1} & & \lambda_1 & \lambda_0 \end{bmatrix}$ $X = \begin{bmatrix} x_0-\bar{x} \\ x_1-\bar{x} \\ \vdots \\ x_T-\bar{x} \end{bmatrix}$	$A_M = T \begin{bmatrix} c(0) & c(1) & \dots & c(m_1) \\ c(-1) & & & \vdots \\ \vdots & & & c(1) \\ c(-m_1) & \dots & c(-1) & c(0) \end{bmatrix}$	autocovariances (discrete) $c(k) = \frac{1}{T+1} \sum_{t=0}^{T-k} [x_t-\bar{x}]$ $[x_{t+k}-\bar{x}]$
Bivariate-continuous Var: $X(t) = (X_i(t), X_j(t))$	$p(X_i,X_j) = \frac{1}{Z(\Lambda)} \exp\{-[\lambda_0,\dots,\lambda_{m_2}] \begin{bmatrix} T\,c_{ij}(0) \\ \vdots \\ T\,c_{ij}(m_2) \end{bmatrix}$ $\Lambda = [\lambda_0,\dots,\lambda_{m_2}]$ $Z(\lambda_0,\dots,\lambda_{m_2}) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} dx_i\, dx_j \exp\{-\Lambda A_v\}$	$A_v = T \begin{bmatrix} c_{ij}(0) \\ \vdots \\ c_{ij}(m_2) \end{bmatrix}$	cross-covariance (continuous) $c_{ij}(k) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{+T/2} [x_i(t)-\bar{x}_i]$ $x_j(t+k)-\bar{x}_j]dt$

Table 1

Analysis	pdf	Information Constraint	Sample Information
Bivariate-discrete Var: $X(t) = (X_i(t), X_j(t))$	$p(X_i,X_j) = \frac{1}{Z(\Lambda)} \exp[- X_i^* \Lambda X_j]$ $\Lambda = \begin{bmatrix} \lambda_0 & & & 0 \\ \lambda_{-1} & & & \\ \vdots & & & \\ \lambda_{m_2} & & & \\ 0 & \lambda_{-m} & \lambda_1 & \lambda_0 \end{bmatrix}$ $X_i = \begin{bmatrix} x_i(0)-\bar{x}_i \\ \vdots \\ x_i(T)-\bar{x}_i \end{bmatrix}$ $X_j = \begin{bmatrix} x_j(0)-\bar{x}_j \\ \vdots \\ x_j(T)-\bar{x}_j \end{bmatrix}$	$A_v = T \begin{bmatrix} c(0) & \dots & c(m_2) \\ \vdots & & \\ c(-m_2) & \dots & c(0) \end{bmatrix}$	cross-covariance (discrete) $c_{ij}(k) = \frac{1}{T+1} \sum_{t=0}^{T+k} [X_j(t)-\bar{x}_j]$ $[x_i(t-k)-\bar{x}_i]$
Multivariate-continuous Var: $X(t) = (X_1(t),\dots,X_m(t))$	$p(X_1,X_2,\dots,X_m) = \frac{1}{Z(\Lambda)} \exp[- \Lambda A]$ $\Lambda = [\Lambda_1,\dots,\Lambda_m],\ \Lambda_i = [\lambda_{m_2},\dots,\lambda_0,\dots,\lambda_{-m_2}]$ $A = T \begin{bmatrix} c_{11} & \dots & c_{1m} \\ \vdots & & \\ c_{m1} & \dots & c_{mm} \end{bmatrix} = \begin{bmatrix} A_{11} & \dots & A_{1m} \\ \vdots & & \\ A_{m1} & \dots & A_{mm} \end{bmatrix}$	$A_{ij} = T \begin{bmatrix} c(0) & \dots & c(m_2) \\ \vdots & & \\ c(-m_2) & \dots & c(0) \end{bmatrix}$	cross-covariance (discrete) $1 \le i,j \le m \quad (i \ne j)$ autocovariance (continuous) $1 \le i,j \le m \quad (i = j)$
Multivariate-discrete Var: $X(t) = (X_1(t),\dots,X_m(t))$	$p(X_1,X_2,\dots,X_m) = \frac{1}{Z(\Lambda)} \exp[- \Lambda A]$ $= \frac{1}{Z(\Lambda)} \exp[- \sum_{i=1}^{m} X_i^* \Lambda_i \sum_{j=1}^{m} X_j]$ $X_i = \begin{bmatrix} x_i(0)-\bar{x}_i \\ \vdots \\ x_i(T)-\bar{X}_i \end{bmatrix}$; $X_j = \begin{bmatrix} x_j(0)-\bar{x}_j \\ \vdots \\ x_j(T)-\bar{X}_j \end{bmatrix}$	$A = \begin{bmatrix} X_1^*X_1 & X_2^*X_1 & \dots & X_m^*X_1 \\ \vdots & X_2^*X_2 & & \\ \vdots & & & \\ \vdots & & & X_m^*X_m \end{bmatrix}$	cross-covariance (discrete) $1 \le i,j \le m \quad (i \ne j)$ autocovariance (continuous) $1 \le i,j \le m \quad (i = j)$

In addition, three more appendices are added. Appendix B provides the Levinson-Burg algorithm that can be used for an easy alternative way to calculate Lagrange multipliers. In Appendix C, it is shown that the Gumbel distribution is a special consequence of univariate distribution, assuming annual maximum peaks to be iid. Similarly, Appendix D shows that the Poisson distribution of the number of flood events in an arbitrary interval is also the special case of our univariate distribution.

5. APPLICATION

5.1 Application to Real-World Data

The various components of the model have been tested on two data sets. Thirty years of daily mean discharges (in period 1956-1985) were obtained from the U.S. Geological Survey (USGS) for two Louisiana rivers: Spring Creek and Bayou Bayou. For both rivers low threshold value was chosen (100 cfs for Spring Creek and 80 cfs for Bayou Bayou) so that in every year at least one flood was available. In this way, the model performance can be evaluated more thoroughly. For Spring Creek, multiple peak hydrographs appeared often, but in total there were only 28 peaks more than volumes or approximately one more per hydrological year. Solutions for both multiple and single peak hydrographs were tested with only slight difference (λ_1 that appeared in Λ matrix still was much smaller in magnitude than λ_0). Further, we used only single peak hydrographs, taking only the maximum peak when multiple hydrographs did occur.

The frequency histograms of peaks, volumes and durations are presented in figures 1-3 for Spring Creek, and in figures 8-10 for Bayou Bayou. Flood peaks, volumes and durations may all be represented by either exponential or Gaussian distribution, taking either the mean

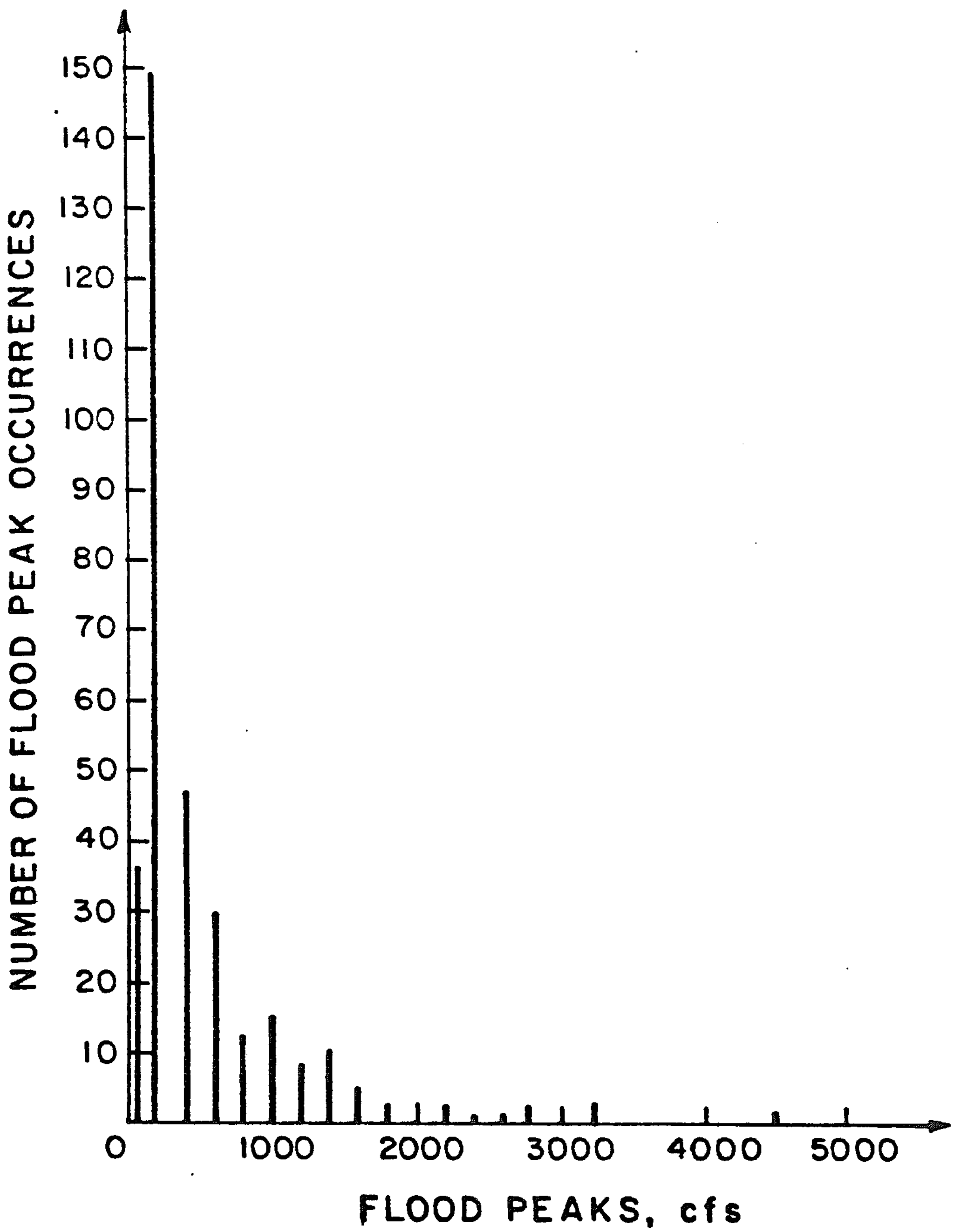

Figure 1. Histogram of flood peaks for Spring Creek (Louisiana) for the period 1956-1985.

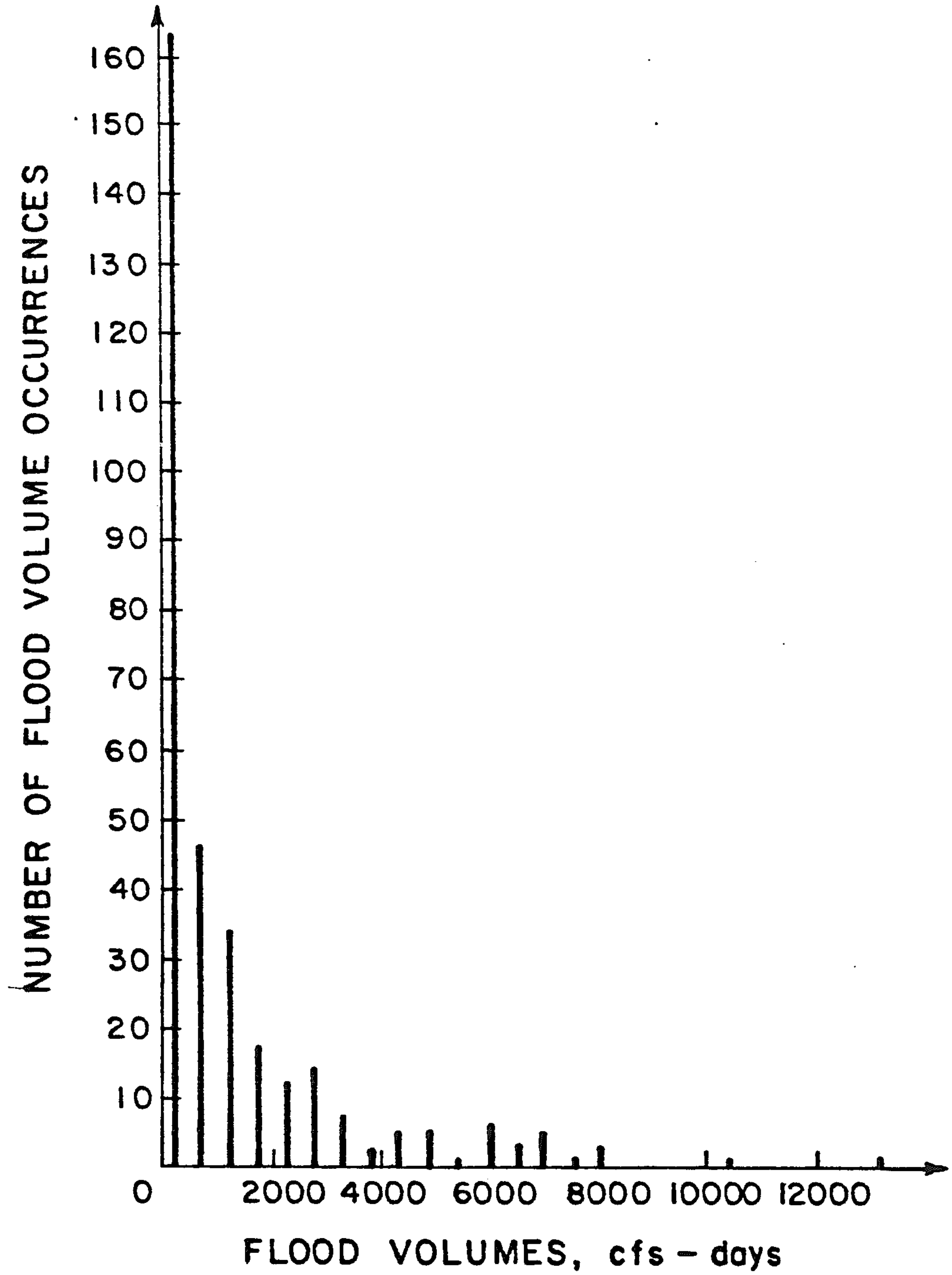

Figure 2. Histogram of flood volumes for Spring Creek (Louisiana) for the period 1956-1985.

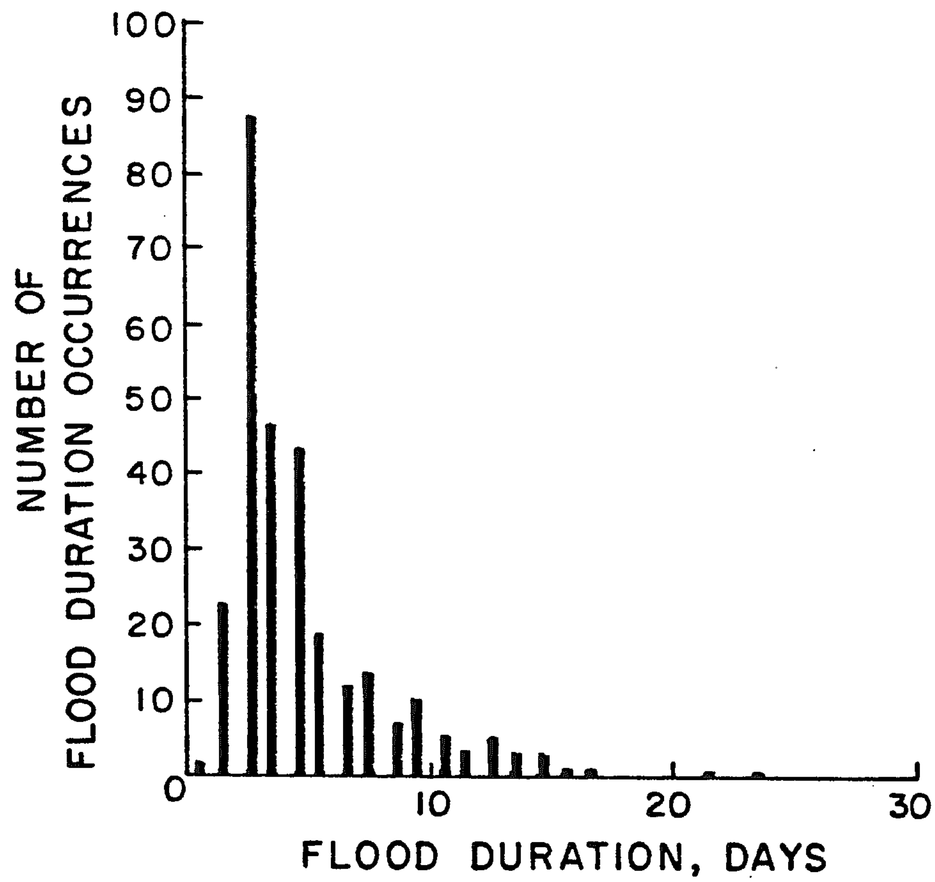

Figure 3. Histogram of flood durations for Spring Creek (Louisiana) for the period 1956-1985.

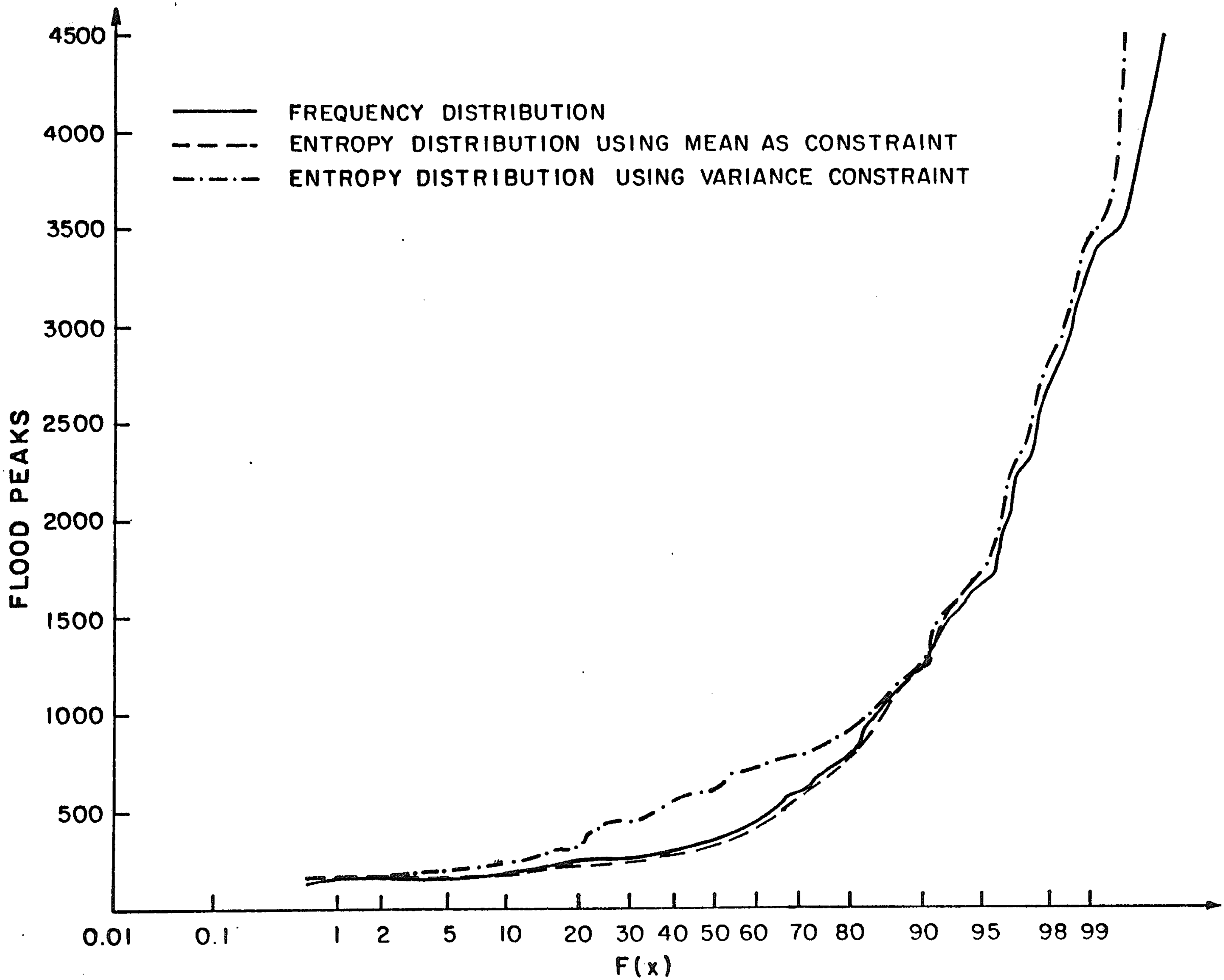

Figure 4. Cumulative distribution of flood peaks for Spring Creek (Louisiana).

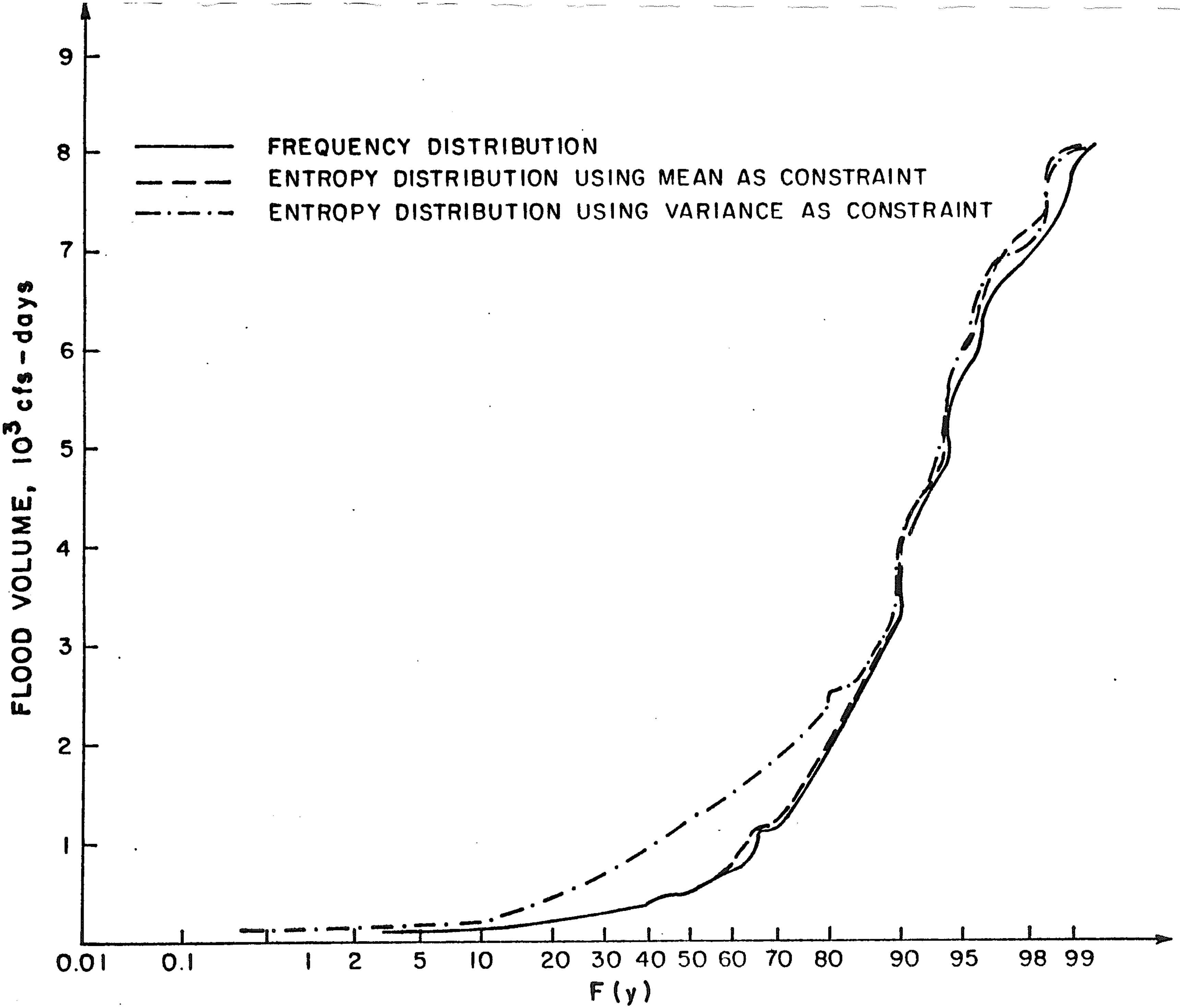

Figure 5. Cumulative distribution of flood volumes for Spring Creek (Louisiana).

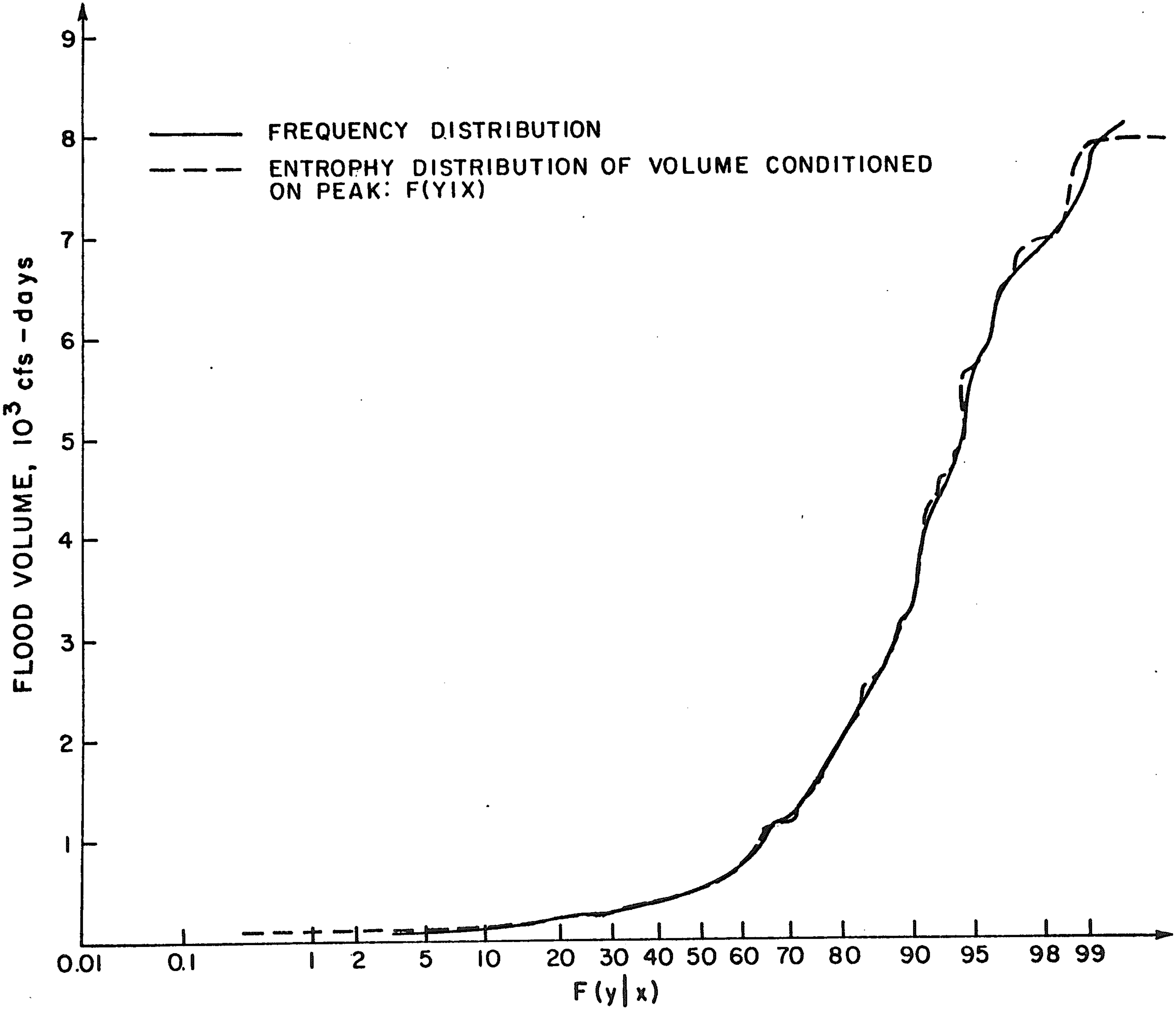

Figure 6. Cumulative distribution of flood volumes conditioned on peaks for Spring Creek (Louisiana).

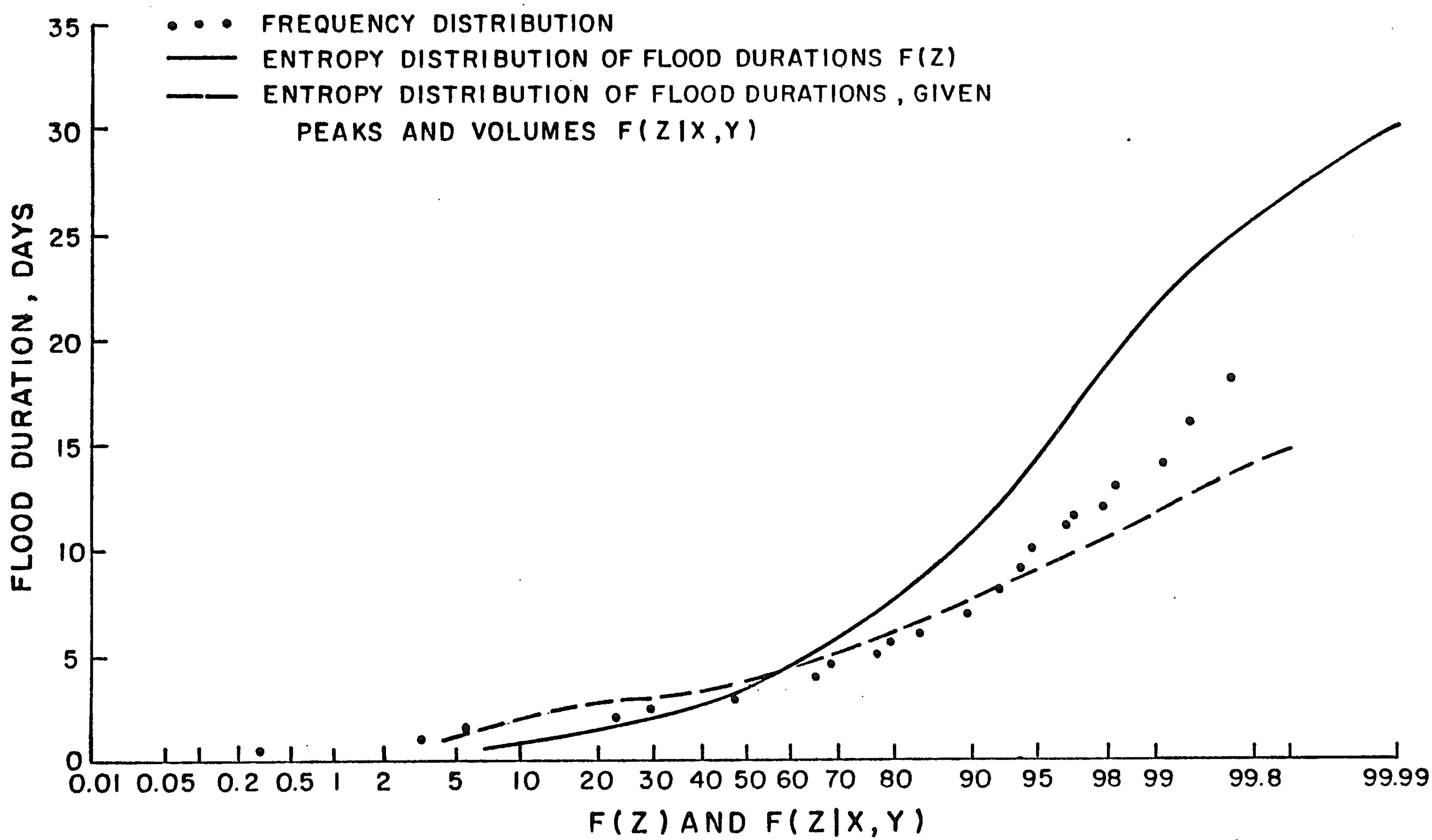

Figure 7. Cumulative distribution of flood durations for Spring Creek (Louisiana).

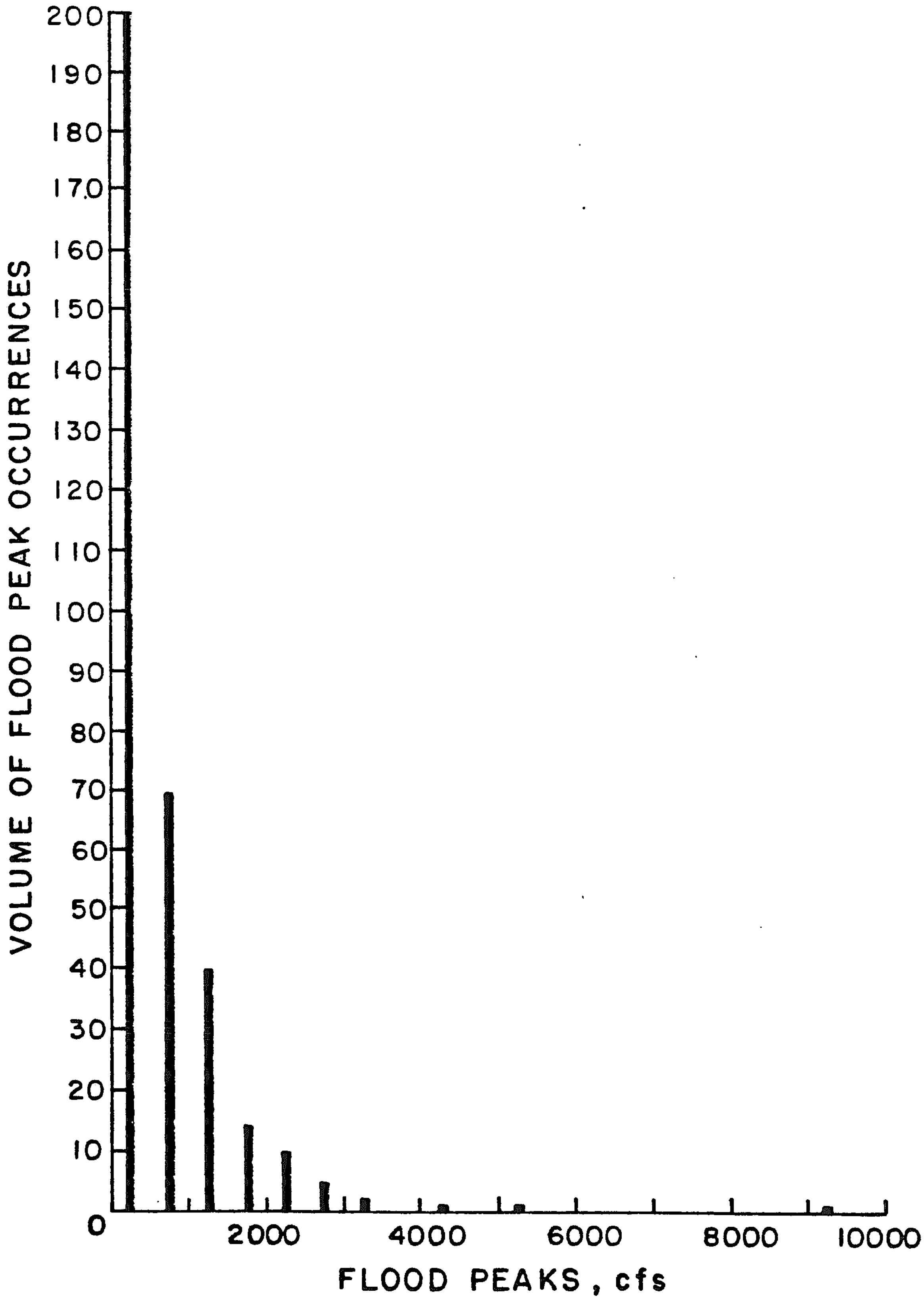

Figure 8. Histogram of flood peaks for Bayou Bayou (Louisiana) for the period 1956-1985.

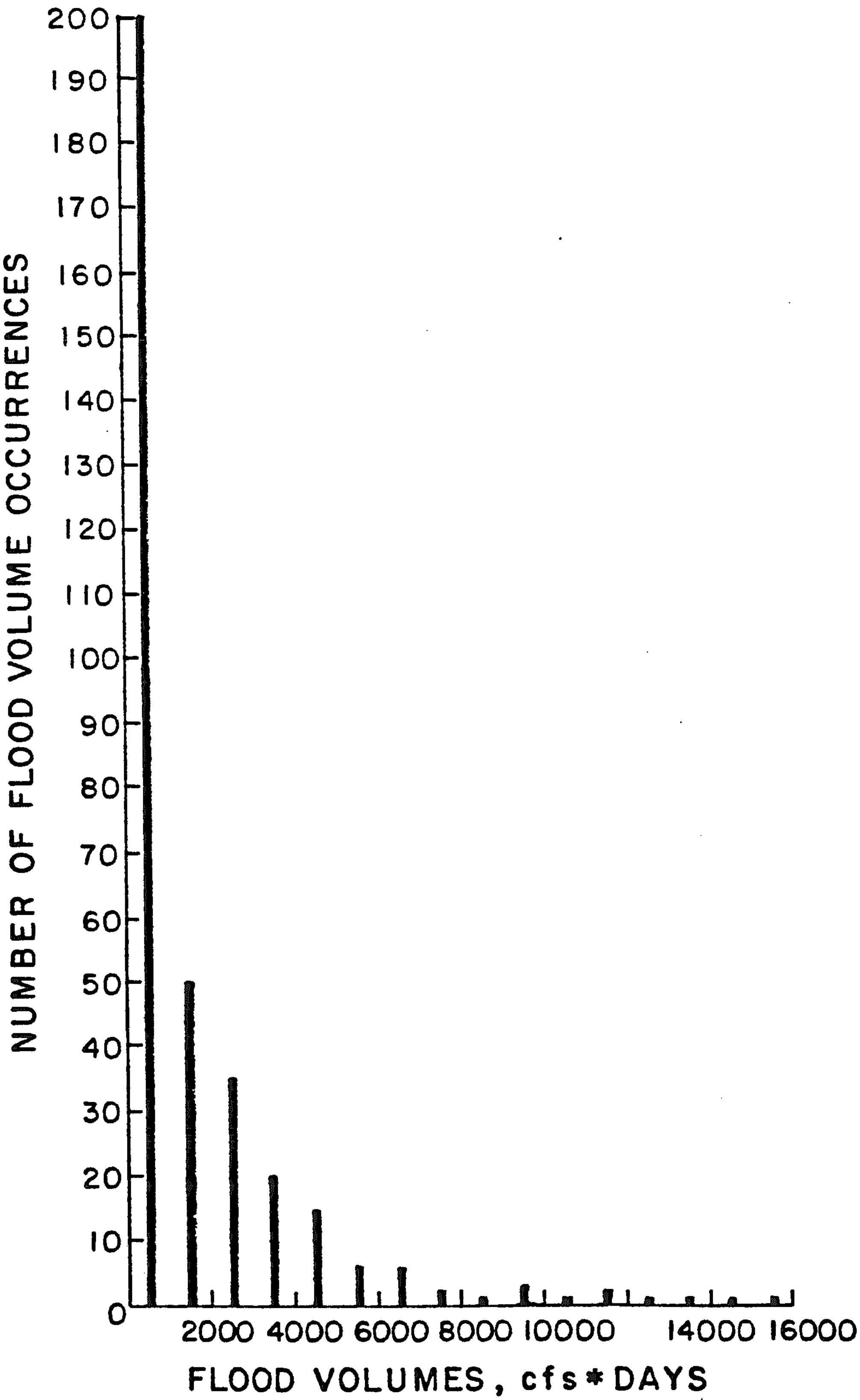

Figure 9. Histogram of flood volumes for Bayou Bayou (Louisiana) for the period 1956-1985.

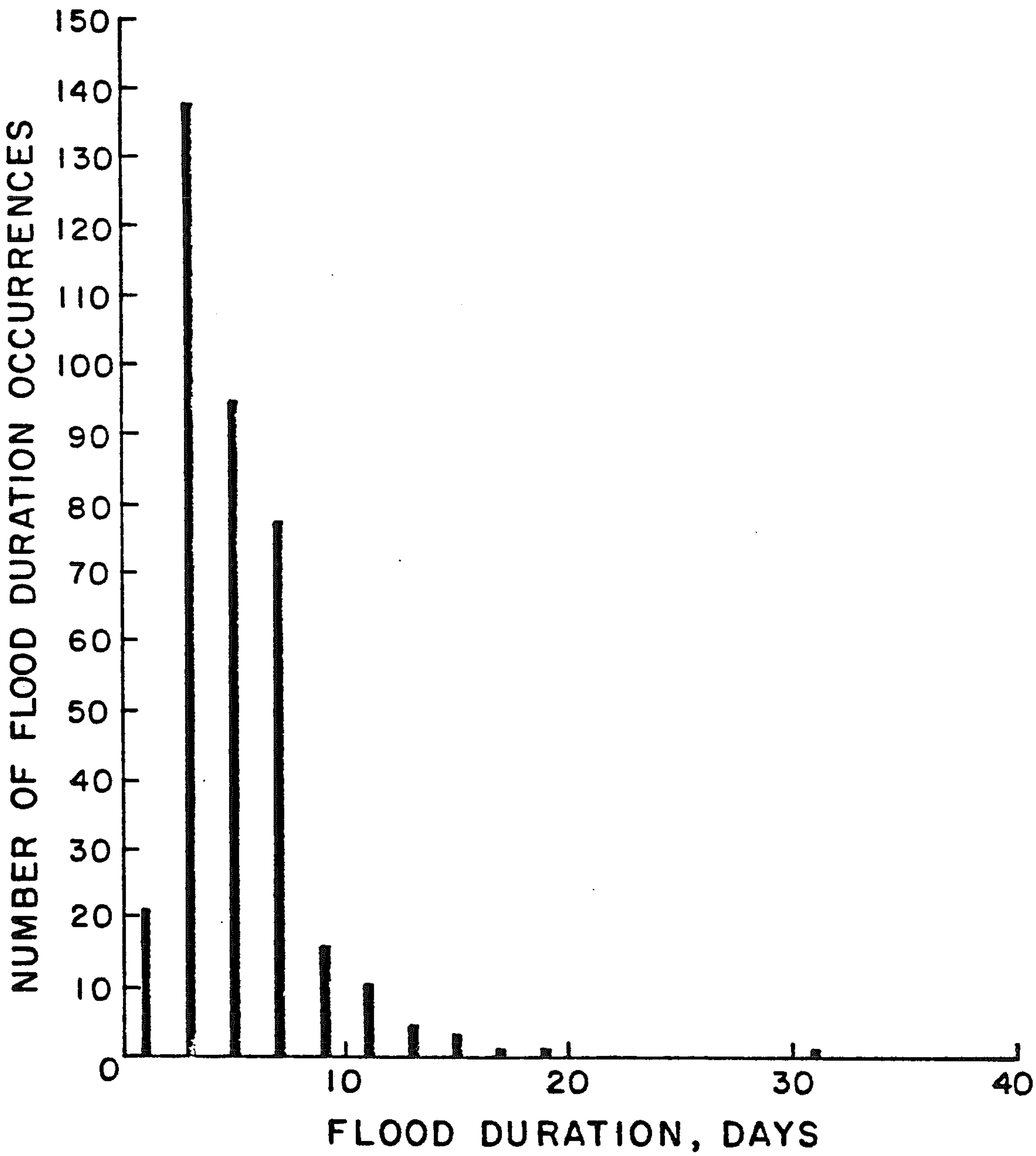

Figure 10. Histogram of flood durations for Bayou Bayou (Louisiana) for the period 1956-1985.

and/or the variance as the information available from data, and using simple univariate models of (48) or (49). For Spring Creek, both possibilities were tested for distributions of peaks (figure 4) and volumes (figure 5), while exponential distribution of (49) was used for flood durations (figure 7). Both tests showed that prediction was much better with mean than with variance. Exponential distribution using mean as the information was fitted to peaks, volumes and durations in figures 11-13 for Bayou Bayou river.

Further, we tested the bivariate distribution of (64) for flood peaks and volumes, and the trivariate distribution of (67) for peaks, volumes and durations. To that end, we improved prediction of volumes using information on peaks. Let X and Y be random variables of peak and volume, respectively. Then

$$p(Y|X) = \frac{p(X,Y)}{p(X)} \tag{68}$$

using (64) for p(X,Y) and (49) for p(X). Results are presented in figure 6 for Spring Creek and in figure 12 for Bayou Bayou. Similarly, let Z be a random variable of flood duration. Using Bayes' formula:

$$p(Z|X,Y) = \frac{p(X,Y,Z)}{p(X,Y)} \tag{69}$$

with (67) for trivariate pdf and (65) for the bivariate case, results are presented in figure 7 for Spring Creek and in figure 13 for Bayou Bayou.

In these applications, we deal with only one Lagrange multiplier, taking independency of flood occurrences. This assumption was not made in advance. Using the Levinson-Burg algorithm, λ's were obtained for all lags and λ_0 was 10 times higher in magnitude than the nearest one (λ_1). Since λ in the model is responsible for the information intro-

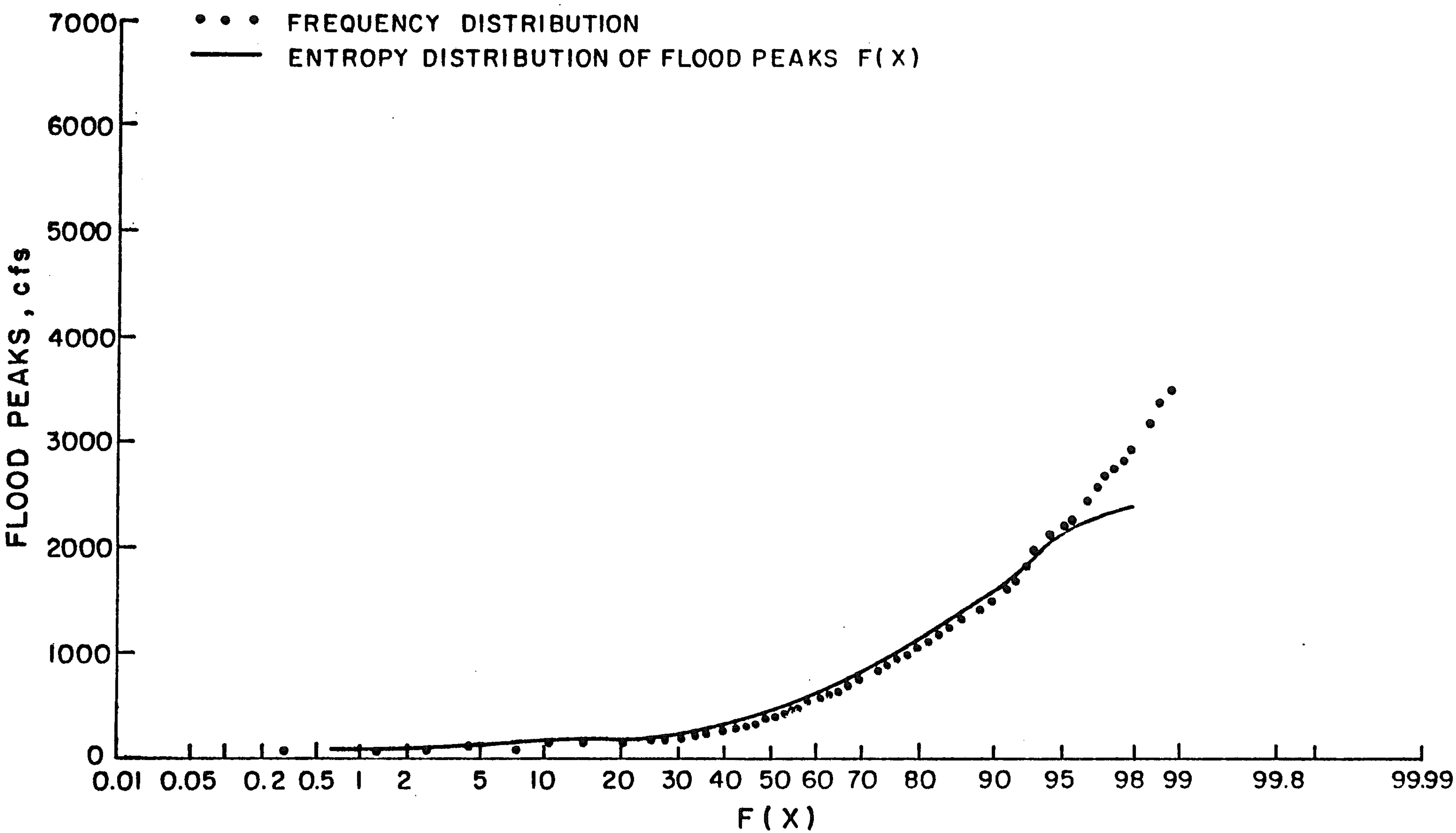

Figure 11. Cumulative distribution of flood peaks for Bayou Bayou (Louisiana).

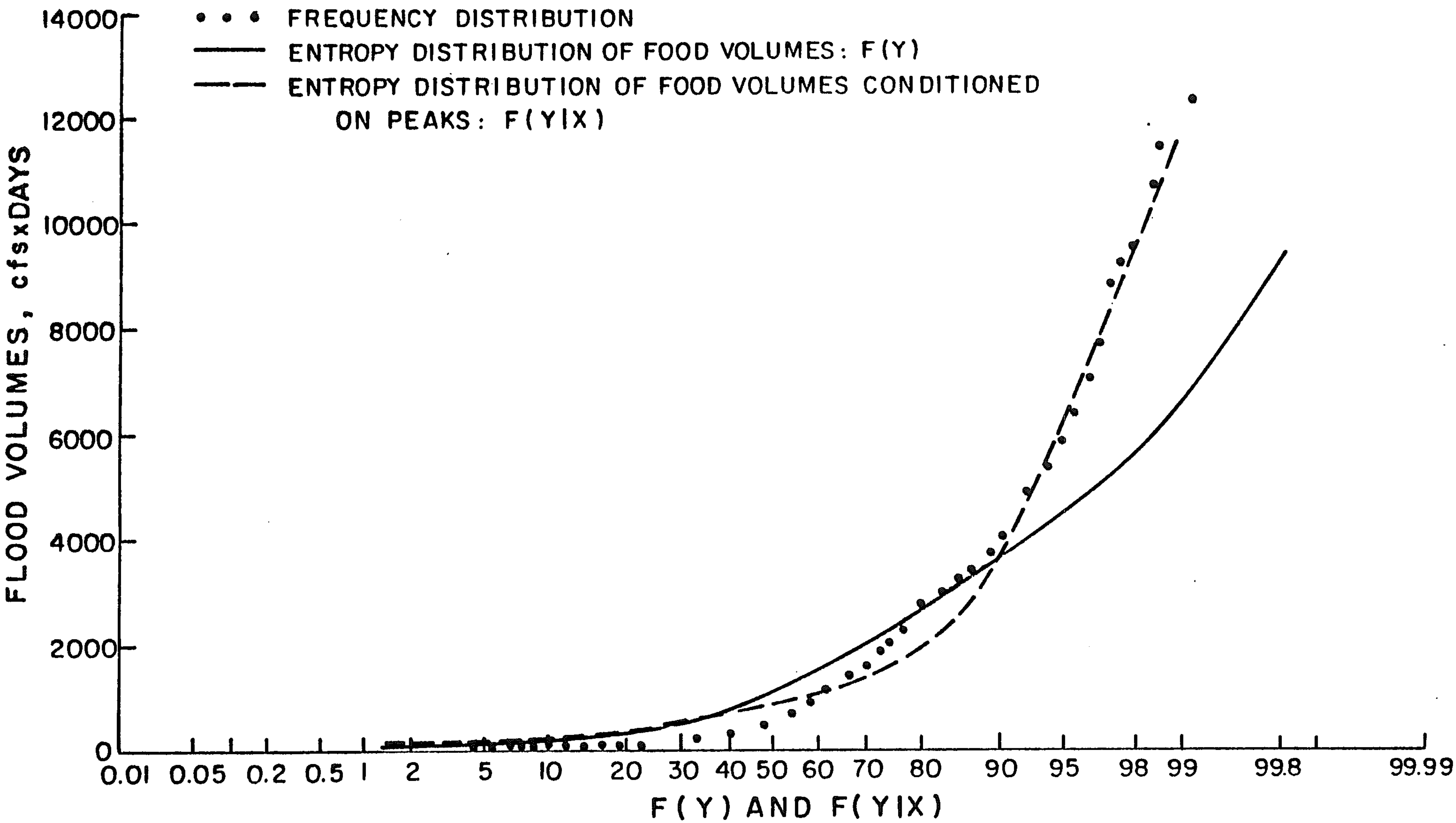

Figure 12. Cumulative distribution of flood volumes for Bayou Bayou (Louisiana).

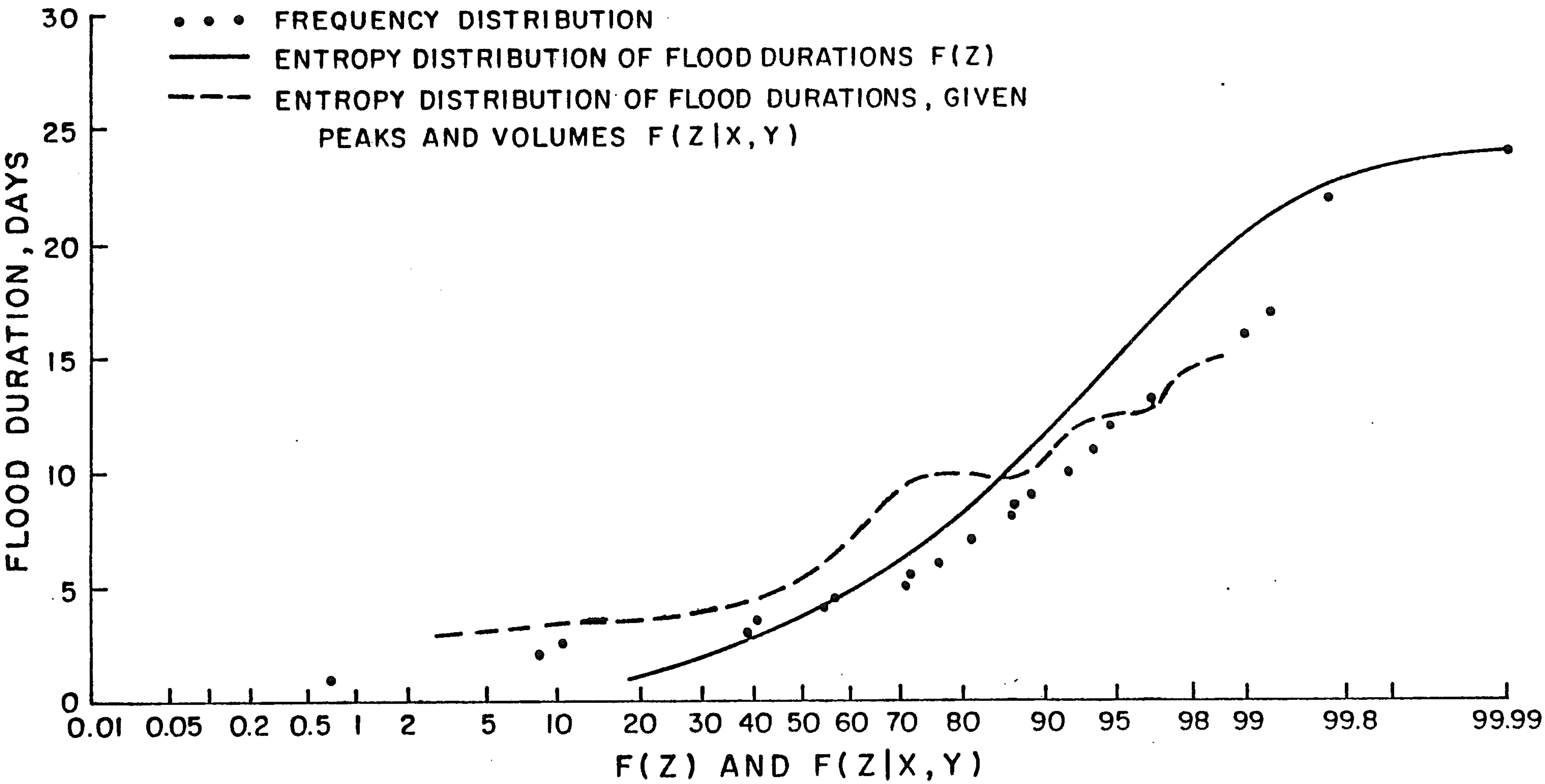

Figure 13. Cumulative distribution of flood durations for Bayou Bayou (Louisiana).

duced, it is concluded that the most significant information is introduced at lag 0. (The next significant lag was only 18 showing seasonality of flood occurrences.) For example for Spring Creek:

λ_0 (peaks) = 1.21 >> λ_{18} (peaks) = 0.19

λ_0 (volumes) = 1.38 >> λ_{18} (volumes) = 0.12

5.2 Discussion

Results of figures 6, 7, 12 and 13 showed that both bivariate and trivariate distributions considerably improved prediction of one univariate flood property (e.g., volumes, durations, etc.). We did not have to assume any arbitrary shape of the flood hydrograph. Although Bayes' formula as used, entropy served primarily to transfer the knowledge from one flood property to another. In this transfer, the relationship with another variable both through space and time was accounted for.

6. CONCLUSIONS

The most important features of the model are: (1) The most general entropy model is drived for flood prediction. Given certain information, it specializes in most of the existing models: exponential distribution for flood peaks of partial duration series, Gumbel distribution for annual series, etc. (2) Both continuous and discrete analyses of the model are performed. Discrete analysis is shown to be easier for the application purposes. (3) The model accounts for the transfer of information through space and time. (4) The model is not confined to iid assumption; both dependent and independent flood occurrences can be treated. (5) Multivariate model can be used to improve prediction of a single flood property through the use of Bayes' formula. (6) Structure of Lagrange multipliers: Λ_V (vector) in the continuous analysis and Λ_M (matrix) in the discrete analysis prove to be important factors in studying space and time dependency. Each new λ-diagonal is responsible for new information introduced with a new lag. (7) The model shows sufficient flexibility, since under certain assumptions, it reduces to simpler forms easily applicable in practice.

7. REFERENCES

Ables, J. G., 1974. Maximum entropy spectral analysis. Journal of Astronomy and Astrophysics Suppl. 15, pp. 383-393.

Amorocho, J. and Espildora, B., 1973. Entropy in the assessment of uncertainty in hydrologic systems and models. Water Resources Research Vol. 9, No. 6, pp. 1511-1522.

Ashkar, F. and Rousselle, J., 1982. A multivariate statistical analysis of flood magnitude, duration and volume. in: V. P. Singh (editor), Statistical Analysis of Rainfall and Runoff, published by Water Resources Publications, Littleton, Colorado.

Boltzmann, L., 1872. Weitere studien uber das warmegleichgewich unter gasmolekulen. K. Acad. Wiss. (Wien) Sitzb., II Abt., Vol. 66, pp. 275.

Borgman, L. E., 1963. Risk criteria. Journal of Waterways, Harbors and Coastal Engineering Division, ASCE, Vol. 89, pp. 1-35.

Box, G. E. P. and Jenkins, G. M., 1976. Time Series Analysis, Forecasting and Control. Holden Day, San Francisco, California.

Burg, J. P., 1975. Maximum entropy spectral analysis. Unpublished Ph.D. Dissertation, Stanford University, University Microfilms, pp. 75-25, 499.

Collins, R. and Wragg, A., 1977. Maximum entropy histograms. Journal of Physics A: Mathematical and General, Vol. 10, No. 9, pp. 1441-1464.

Gupta, V. K., Duckstein, L. and Peebles, R. W., 1976. On the joint distribution of the largest flood and its time of occurrence. Water Resources Research, Vol. 12, No. 2.

Jaynes, E. T., 1957. Information theory and statistical mechanics, Part I. Physical Review, Vol. 106, pp. 620-630, Part II, Vol. 108, pp. 171-191.

Jaynes, E. T., 1979. Concentration of distributions and entropy maxima. in: E. T. Jaynes, Papers on Probability, Statistics and Statistical Physics, ed. by R. D. Rosenhratz, D. Reidel Publishing Co., 1983.

Jaynes, E. T., 1982. On the rationale of maximum entropy methods. Proceedings of IEEE, Vol. 70, No. 9, pp. 939-952.

Jowitt, P. W., 1979. The extreme-value type I distribution and the principle of maximum entropy. Journal of Hydrology, Vol. 42, pp. 23-38.

Kavvas, M. L., 1982. Stochastic trigger model for flood peaks: 1. Development of the model, 2. Applications of the model. Water Resources Research, Vol. 18, No. 2, pp. 383-411.

Kavvas, M. L. and Delleur, J. W., 1975. The stochastic and chronological structure of rainfall sequences: applications to Indiana. Tech. Report 57, Water Resources Research Center, Purdue University, West Lafayette, Indiana.

Kavvas, M. L., Cervantes, J. E. and Delleur, J. W., 1983. A cluster model for flood analysis. Water Resources Research, Vol. 19, No. 1, pp. 209-224.

North, M., 1980. Time dependent stochastic model of floods. Journal of Hydraulics Division, ASCE, Vol. 106, No. HY5, pp. 649-665.

Rajagopal, A. K., Teitler, S. and Singh, V. P., 1986. Some new perspectives of maximum entropy in water resources research. Paper presented at International Symposium on Flood Frequency and Risk Analyses held at Louisiana State University, Baton Rouge, Louisiana, May 1986.

Shane, R. M. and Lynn, W. R., 1964. Mathematical model for flood risk evaluation. Journal of the Hydraulics Division, ASCE, Vol. 80, pp. 1-20.

Shannon, C. E., 1948a. A mathematical theory of communication I and II. Bell System Tech. Journal, Vol. 27, pp. 379-423.

Shannon, C. E., 1948b. A mathematical theory of communication III and IV. Bell System Tech. Journal, Vol. 27, pp. 623-656.

Singh, V. P. and Krstanovic, P. F., 1985. A stochastic model for sediment yield. Paper presented at the IVth International Hydrology Symposium, "Multivariate Analysis of Hydrologic Processes," Colorado State University, Fort Collins, Colorado.

Singh, V. P. and Krstanovic, P. F., 1986. A stochastic model for sediment yield and phosphorus loading. Submitted to Journal of Environmental Quality.

Singh, V. P., Rajagopal, A. K. and Singh, K., 1986. Derivation of some frequency distributions using the principle of maximum entropy. Advances in Water Resources, Vol. 9, pp. 91-106.

Singh, V. P., Singh, K. and Rajagopal, A. K., 1985. Application of the principle of maximum entropy (POME) to hydrologic frequency analysis. Louisiana Water Resources Research Institute, Louisiana State University, Baton Rouge, Louisiana.

Skilling, J. and Gull, S. F., 1984. The entropy of an image in 'inverse problems.' SIAM-AMS Proceedings, Vol. 14, pp. 167-189

Smith, C. R. and Grandy, Jr., W. T. (editors), 1985. Maximum Entropy and Bayesian Methods in Inverse Problems. D. Reidel Publishing Co., 492 pp.

Sonuga, J. O., 1972. Principle of maximum entropy in hydrologic frequency analysis. Journal of Hydrology, Vol. 27, pp. 177-191.

Sonuga, J. O., 1976. Entropy principle applied to the rainfall-runoff process. Journal of Hydrology, Vol. 30, pp. 81-94.

Thomas, H. A. and Fiering, M. B., 1962. Mathematical synthesis of streamflow sequences for the analysis of river basins by simulation. in: Design of Water Resources Systems, edited by Mass, A., et al., pp. 459-463, Harvard University Press, Cambridge, Massachusetts.

Todorovic, P. and Rousselle, J., 1971. Some problems of flood analysis. Water Resources Research, Vol. 7, No. 5, pp. 1144-1150.

Todorovic, P. and Zelenhasic, E., 1970. A stochastic model for flood analysis. Water Resources Research, Vol. 6, pp. 1641-1648.

Todorovic, P. and Woolhiser, D., 1972. On the time when extreme flood occurs. Water Resources Research, Vol. 8, No. 6, pp. 1433-1438.

Todorovic, P., 1978. Stochastic models of floods. Water Resources Research, Vol. 14, No. 2, pp. 345-356.

Yevjevich, V., 1963. Fluctuations of wet and dry years (Part I). Research Data Assembly and Math. Models, Hydrology Paper I, Colorado State University, Fort Collins, Colorado.

APPENDIX A

By working out a couple of simple examples, we illustrate the importance of the Lagrange multipliers.

For k = 0, using (61b),

$$c_0 = \frac{1}{2\pi} \int_0^{2\pi} \frac{d\theta}{\lambda_0} = \frac{1}{2\pi} \frac{2\pi}{\lambda_0}$$

which yields

$$\lambda_0 = 1/c_0 \tag{A-1}$$

or using (61a):

$$c_0 = - \frac{\partial}{\partial\lambda_0} \ln Z = - \frac{\partial}{\partial\lambda_0} \ln|\Lambda| = - \frac{\partial}{\partial\lambda_0} \ln \lambda_0^{T+1}$$

$$c_0 = \frac{1}{T+1} \frac{T+1}{\lambda_0} = \frac{1}{\lambda_0}$$

which again yields (A-1).

For k = 1,

$$c_0 = \frac{1}{2\pi} \int_0^{2\pi} \frac{d\theta}{\lambda_0 + \lambda_1 \exp(i\theta)}$$

$$c_1 = \frac{1}{2\pi} \int_0^{2\pi} \frac{\exp(i\theta)d\theta}{\lambda_0 + \lambda_1 \exp(i\theta)}$$

Using substitutions: $z = \exp(i\theta)$ and $dz = d\theta i \exp(i\theta)$

$$c_0 = \frac{1}{2\pi i\lambda_1} \int_{|z|=1} \frac{dz}{z[z-(-\lambda_0/\lambda_1)]} = \frac{1}{2\pi i\lambda_1} \int_{|z|=1} \frac{dz}{z(z-z_0)}$$

$$c_1 = \frac{1}{2\pi i\lambda_1} \int_{|z|=1} \frac{dz}{z-z_0}$$

where $z_0 = -\lambda_0/\lambda_1$. The first equation can be solved by the residue theorem, while the second equation by Cauchy integral formula yields:

$$c_1 = \frac{1}{2\pi i\lambda_1} 2\pi i[f(z = z_0)], \text{ where } f(z) = 1$$

$$c_1 = \frac{1}{\lambda_1} 1 = \frac{1}{\lambda_1}$$

Thus,

$$\lambda_1 = 1/c_1 \tag{A-2}$$

For k = 2, sample autocovariances are given as:

$$c_0 = \frac{1}{T+1} \sum_{j=0}^{T} x_j^2$$

$$c_1 = \frac{1}{T+1} \sum_{j=0}^{T-1} x_j\, x_{j-1}$$

and

$$c_2 = \frac{1}{T+1} \sum_{j=0}^{T-2} x_j\, x_{j-2}$$

The Lagrange multipliers are then determined from

$$c_0 = \frac{1}{2\pi} \int_0^{2\pi} \frac{d\theta}{\lambda_0+\lambda_1 \exp(i\theta)+\lambda_2 \exp(2i\theta)} = \frac{1}{2\pi i\lambda_2} \int_{|z|=1} \frac{dz}{(z-z_1)(z-z_2)z}$$

$$c_1 = \frac{1}{2\pi} \int_0^{2\pi} \frac{\exp(i\theta)d\theta}{\lambda_0+\lambda_1 \exp(i\theta)+\lambda_2 \exp(2i\theta)} = \frac{1}{2\pi i\lambda_2} \int_{|z|=1} \frac{dz}{(z-z_1)(z-z_2)}$$

$$c_2 = \frac{1}{2\pi} \int_0^{2\pi} \frac{\exp(2i\theta)d\theta}{\lambda_0+\lambda_1 \exp(i\theta)+\lambda_2 \exp(2i\theta)} = \frac{1}{2\pi i\lambda_2} \int_{|z|=1} \frac{zdz}{(z-z_1)(z-z_2)}$$

Solving the last equation:

$$c_2 = (1/\lambda_2)\ [\mathrm{Res}(z = z_1) + \mathrm{Res}(z = z_2)]$$

$$\mathrm{Res}(z = z_1) = \lim_{z\to z_1} [(z-z_1) \frac{z}{(z-z_1)(z-z_2)}] = \frac{z_1}{z_1-z_2}$$

$$\mathrm{Res}(z = z_2) = \lim_{z\to z_2} [(z-z_2) \frac{z}{(z-z_1)(z-z_2)}] = \frac{z_1}{z_2-z_1}$$

Thus,

$$c_2 = \frac{1}{\lambda_2} [\frac{z_1}{z_1-z_2} + \frac{z_2}{z_2-z_1}] = \frac{1}{\lambda_2}$$

$$\lambda_2 = 1/c_2 \tag{A-3}$$

A similar derivation for c_0, c_1, c_2, c_3 yields:

$$\lambda_3 = 1/c_3$$

or in general the last Lagrange multiplier of the model will always be as:

$$\lambda_k = 1/c_k \tag{A-4}$$

An exact theoretical proof can be shown by mathematical induction.

APPENDIX B

Levinson-Burg Algorithm for Solution of the Toeplitz Matrix

Given the matrix equation of either autocovariances c_k or autocorrelation coefficients r_k:

$$\begin{bmatrix} r_0 & r_{-1} & \cdots & r_{1-m} \\ r_1 & r_0 & & \vdots \\ \vdots & & & r_{-1} \\ r_{m-1} & \cdots & r_1 & r_0 \end{bmatrix} \begin{bmatrix} 1 \\ b_1 \\ \vdots \\ b_{m-1} \end{bmatrix} = \begin{bmatrix} P_{m-1} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{B-1}$$

where P_{m-1} is a constant. For known values of r_k and P_{m-1}, b_m are solved as a system of m equations with m unknowns. To initiate the algorithm we start with m = 1. Then $r_0 = P_0$ and $b_0 = 1$. Every higher m (≥ 2) system is solved by using the solution of previous one with lower m. In general, the system (m+1)x(m+1) will be solved from (B-1) as:

$$\begin{bmatrix} r_0 & r_{-1} & \cdots & r_{-m} \\ r_1 & r_0 & & r_{1-m} \\ \vdots & & & \vdots \\ r_{m-1} & & & \vdots \\ r_m & \cdots & \cdots & r_0 \end{bmatrix} = \begin{bmatrix} 1 \\ b_1 \\ \vdots \\ b_{m-1} \\ 0 \end{bmatrix} + c_m \begin{bmatrix} 0 \\ b^c_{m-1} \\ \vdots \\ b^c_1 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} P_{m-1} \\ 0 \\ \vdots \\ 0 \\ \Delta_m \end{bmatrix} + c_m \begin{bmatrix} \Delta^c_m \\ 0 \\ \vdots \\ 0 \\ P_{m-1} \end{bmatrix} \tag{B-2}$$

In the original algorithm b^c_m are the complex conjugates of b_m, but for the real hydrologic values $b^c_m = b_m$. From (B-2),

$$\Delta_m = \sum_{n=0}^{m-1} r_{m-n}\, b_m$$

$$\text{and} \quad c_m = -\frac{\Delta_m}{P_{m-1}} \tag{B-3}$$

The matrix of (m+1)x(m+1) system will have the following coefficients:

$$\begin{bmatrix} 1 \\ a_1 \\ \cdot \\ \cdot \\ \cdot \\ a_{m-1} \\ a_m \end{bmatrix} = \begin{bmatrix} 1 \\ b_1 \\ \cdot \\ \cdot \\ \cdot \\ b_{m-1} \\ 0 \end{bmatrix} + c_m \begin{bmatrix} 0 \\ b_{m-1} \\ \cdot \\ \cdot \\ \cdot \\ b_1 \\ 0 \end{bmatrix}$$

and

$$a_s = b_s + c_{m-s} \tag{B-4}$$

New constant P_m is solved by using P_{m-1} and c_m:

$$P_m = P_{m-1}\,(1 - |c_m|^2) \tag{B-5}$$

Thus, the new system of (m+1)x(m+1) is defined as:

$$\begin{bmatrix} r_0 & \cdot & \cdot & \cdot & \cdot & r_{-m} \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ r_m & \cdot & \cdot & \cdot & \cdot & r_0 \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ \cdot \\ \cdot \\ a_m \end{bmatrix} = \begin{bmatrix} P_m \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix} \tag{B-6}$$

Equation (B-6) is the starting point in solving (m+2)x(m+2) system of equations, after substitution $b_k = a_k$ and $P_{m-1} = P_m$.

We offer an alternate way to compute Lagrange multipliers. Transforming (61b) as:

$$\frac{1}{2\pi} \int_0^{2\pi} \frac{\exp(ik\theta)\, d\theta}{\sum_{k=-m_2}^{+m_2} \lambda_k \exp(ik\theta)} = \frac{1}{2\pi i} \int_{|z|=1} \frac{z^{k-1}\, dz}{\sum_{k=-m_2}^{+m_2} \lambda_k z^k} \tag{B-7}$$

where z is variable in the complex plane. The denominator in (B-7) can be expressed as the product of two polynomials such that the roots of the first lie outside the unit circle and the roots of the second inside the unit circle:

$$\sum_{k=-m_2}^{+m_2} \lambda_k z^k = \frac{1}{P_m} \left(\sum_{k=0}^{m_2} z_k z^k \right) \left(\sum_{k=0}^{m_2} a_k^c z^{-k} \right) \qquad \text{(B-8)}$$

where P_m is a positive constant as defined in (B-5), a_k and a_k^c are coefficients defined in (B-4) where superscript signifies the conjugate. Note that $a_0 = 1$ and $a_k = a_k^c$ for real values. After determining a_k and P_m from (B-4) to (B-6), Lagrange multipliers are obtained from (B-8) by equating the terms of equal powers:

$$\lambda_{-j} = \lambda_j = \frac{a_j + \sum_{k=1}^{m_2-j} a_k a_{k+1}}{P_m} \qquad \text{(B-9)}$$

APPENDIX C

Derivation of the Gumbel Distribution for Annual Flood Series

We derive the Gumbel distribution using the assumption that the maximum term in a sequence of i.i.d. variables with a common distribution function has an exponential tail. We show that this distribution is a special case of the proposed model. Information is given as:

$$g_1 = E[\exp(-x)] = \int_{-\infty}^{+\infty} \exp(-x)\ p(x)dx \tag{C-1}$$

and determined from the sample estimate:

$$g_1 = \frac{\exp(-x_1) + \exp(-x_2) + \ldots + \exp(-x_T)}{T} = \frac{\sum_{i=1}^{T} \exp(-x_i)}{T} \tag{C-2}$$

Under the i.i.d. assumption, the POME distribution is

$$p(x_i) = \frac{1}{Z} \exp[-\ \lambda_1\ A_1] \tag{C-3}$$

Let $A_1 = T \sum_{i=1}^{T} x_i$. Then

$$p(x_i) = \frac{1}{Z(\lambda)} \exp[-\ \lambda \sum_{i=1}^{n} \exp(-x_i)] \tag{C-4}$$

where $Z(\lambda)$ is determined from $\sum_{i=1}^{n} p(x_i) = 1$ as:

$$Z(\lambda) = \sum_{i=1}^{n} \exp(-\lambda \Sigma x_i) \tag{C-5}$$

$$p(\vec{X}) = \frac{1}{Z(\Lambda)} \exp[-X\Lambda \cdot 1^*] \tag{C-6}$$

where $\vec{X} = [\exp(-x_1), \ldots, \exp(-x_T)]$ is row-vector of annual flood peak exponentials, Λ is diagonal matrix with the unique λ on diagonal and $1 = (1, \ldots, 1)$. $Z(\Lambda)$ is as defined in (C-5). λ is determined numerically from:

$$-\frac{\partial}{\partial\lambda}\ln Z = g_1 \qquad \text{(C-7)}$$

which yields:

$$\exp(-\lambda\Sigma x_i) = \frac{T \sum_{i=1}^{T} x_i}{\sum_{i=1}^{T} \exp(-x_i)} \qquad \text{(C-8)}$$

APPENDIX D

Entropy Representation of the Poisson Process

Consider the number of the flood occurrences N(t) to be a random variable. The information needed as the input to the entropy model is $g_1 = \bar{N}(t)$ which is the average number of occurrences in internal (0,T), e.g., one year. Thus, this derivation is possible only in partial duration-series approach where more than one occurrence is treated in one year interval.

From the POME algorithm:

$$p[N(t) = n_i] = \frac{1}{Z(\lambda)} \exp(-\lambda n_i)$$

where

$$\ln Z(\lambda) = \ln \lambda$$

$$Z(\lambda) = -\bar{N}(t)$$

and

$$\lambda = \frac{1}{\bar{N}(t)}$$

Thus,

$$p(n_i) = \frac{1}{\bar{N}(t)} \exp(-\frac{n_i}{\bar{N}(t)}) \qquad \text{(D-1)}$$

or $\quad p(X) = \lambda \exp(-\lambda x)$

The maximum entropy distribution of the number of occurrences is exponential distribution.

From the probability theory, if the random variable X_i is associated with the time interval τ_i from some fixed origin t = 0 to the subsequent point:

$$p(\tau_i) = \lambda \exp(-\lambda \tau_i)$$

then the times of peak occurrences $\tau_1, \tau_1+\tau_2, \ldots$, are all exponentially distributed with parameter λ.

The N-th point occurs at the time

$$T_r = \sum_{i=1}^{r} \tau_i$$

and it follows that Γ-distribution:

$$P(X=T_r) = \frac{\lambda(\lambda\tau)^{r-1} \exp(-\lambda\tau)}{(r-1)!} \tag{D-2}$$

The number of points in time interval T_r: $N(T_r)$ has Poisson distribution with mean $\lambda_{(0,T_r)}$. Write

$$P\{N(t) < r\} = P\{X_1 + \ldots + X_r > t\} = \int_t^{\infty} \frac{\lambda(\lambda u)^{r-1} \exp(-\lambda u)}{(r-1)!} du$$

By repeated partial integration, we obtain:

$$P\{N(t) < r\} = \sum_{s=0}^{r-1} \frac{\exp(-\lambda t)\ (\lambda t)^s}{s!}$$

or

$$P\{N(t) = r\} = \frac{\exp(-\lambda t)\ (\lambda t)^r}{r!} \tag{D-3}$$

(D-3) is the Poisson distribution that is seen as special case of the exponential distribution considering any number of events occurring in an arbitrary time interval (0,t). In conclusion, $\lambda = 1/\bar{N}(t)$, and the pdf that gives maximum entropy is:

$$P\{N(t) = r\} = \frac{\exp[-t/\bar{N}(t)]\ [t/\bar{N}(t)]^r}{r!} \tag{D-4}$$