THEORETICAL PROBABILITY DISTRIBUTIONS FOR FLOOD PEAKS

by

Emir Zelenhasic

November 1970

HYDROLOGY PAPERS
COLORADO STATE UNIVERSITY
Fort Collins, Colorado

42

THEORETICAL PROBABILITY DISTRIBUTIONS FOR FLOOD PEAKS

by

Emir Zelenhasic

HYDROLOGY PAPERS

COLORADO STATE UNIVERSITY

FORT COLLINS, COLORADO 80521

November 1970 No. 42

PREFACE

This paper is based on results of the dissertation "A Stochastic Model for Flood Analysis" submitted by Mr. Emir Zelenhasic in partial fulfillment of the requirements for the Degree of Doctor of Philosophy, Colorado State University, Fort Collins, Colorado, August 1970. He has been guided in his research by the members of his graduate committee: Dr. Petar Todorovic, Associate Professor of Civil Engineering, who has given counsel and guidance in the probabilistic formulation of the properties of maximum exceedances of a random number of random variables; Dr. Moinuddin Siddiqui, Professor of Mathematics and Statistics, who has given counsel in the statistical aspects of the study; Mr. David Dawdy, Affiliate Staff Member of Civil Engineering Department and Research Hydrologist, U.S. Geological Survey, who has given counsel in the phenomenological aspects of the study, and Dr. Vujica Yevjevich, Professor of Civil Engineering, and Chairman of the Graduate Committee, who has given general counsel as the student's major professor and assisted in editing the dissertation and this text. Graduate student Mr. Tom Croley assisted in the preparation of this manuscript for printing.

The investigations leading to this paper are included under the research project "Stochastic Processes in Hydrology," sponsored by U.S. National Science Foundation, Grant GK-11444.

LIST OF FIGURES AND TABLES - Continued

LIST OF FIGURES AND TABLES - Continued

ABSTRACT

Following the theory of the supremum of a random number of random variables a stochastic model is presented for interpretation, analysis, and prediction of the largest flood peak discharge above a given base level concerning a time interval $[o,t]$, at a given location of a river. Although the analysis of floods is the main objective of the developed stochastic model, it has a broader scope. The model can be applied to any kind of data of an intermittent process having a substantial stochastic component for which probabilities of the largest value are desired.

The model has been applied in this study to data from gaging stations on the Susquehanna River at Wilkes-Barre, Pa., and the Greenbrier River at Alderson, W. Va. Results were compared to those obtained by Gumbel's method; they indicate that the introduced model fits the data better.

THEORETICAL PROBABILITY DISTRIBUTIONS FOR FLOOD PEAKS

by

Emir Zelenhasic*

Chapter 1

INTRODUCTION

In recent years many analysts of flood frequency distributions, and the many inferences made about flood probability distribution functions, have indicated that there is an increasing need for improvement in methods used for flood analysis. Various branches of engineering and water resources conservation and development represent the basis for this need. According to Benson [4], who was with the Work Group on Flow-Frequency Methods, Hydrology Committee, Water Resources Council of the U.S. Government, which studied the most commonly used methods of flood-frequency analysis, "....the range of uncertainty in flood analysis, regardless of the method used, is still quite large so that there is still a need for continued research and development to solve the many unresolved questions." The Work Group recommended continued study of the problem of analyzing floods. For a more general methodology of flood analysis, progress can be made only through a better understanding of the stochastic nature of the flood phenomenon.

This paper presents a new theoretical approach to the analysis of stochastic properties of floods. The approach is based on some recent developments in the theory of extremes by Todorovic [48]; a general stochastic model for the extremes of a random number of random variables applicable to the problem of flood peak flows is developed. The new method uses data on flood peaks above a given base level, with the stochastic process $x(t)$, defined as the maximum term among a random number of random observations in an interval of time $[o,t]$, as the basis. The distribution of the largest flood peak flow in a given time interval $[o,t]$, with the number of flood peaks exceeding a certain base value in a given time interval and the magnitude of these peaks considered as random variables, as studied.

Practical aspects suggest considering only the sequence of flood peaks above a given base level instead of considering all instantaneous discharges of flood hydrographs or of flood-hydrograph tops above a given base flood discharge. In this way, a sequence of random variables $\xi_1, \xi_2, \ldots, \xi_\nu$ is obtained as flood peaks of all hydrograph tops above a given base discharge. It then becomes feasible to investigate the maximum flood peak distribution as is done in classical extreme value theories.

Because both the number of flood peaks in each subsequent time interval $[o,t]$, say with the year as the interval, and the time when a peak flow occurs are random variables, this problem actually transcends the framework of the classical extreme value theories. This relatively restricted applicability is the major shortcoming of the methods of classical extreme value theories. All previous results obtained are related to problems where the number of observations n is given. However, for many naturally occurring phenomena this number depends on chance. Because, in the classical theories of extremes, the number of observations n is always given, the largest and smallest values are functions of this n. In many practical investigations, particularly in flood control problems, it is important to know how large a maximum flood peak one can expect in a given interval of time. Following the theory of a random number of random variables, the number of observations in the interval $[o,t]$ is a random variable, so that the extremes considered in the time interval $[o,t]$ are functions of t. The approach used in this study takes these stochastic properties of flood peaks into consideration simultaneously.

Even though this method deals with flood peaks above a certain level, instead of total flood peaks, the generality of the method is not limited. If a constant is added to a random variable the largest value in the interval $[o,t]$ increases by the same amount.

Flood data are usually extracted either in the form of annual flood series or partial duration series. The former method defines the annual flood as the highest momentary peak discharge in a water year. Langbein [34] states: "Only the greatest flood in each year is used. An objection most frequently encountered with respect to the use of annual floods is that it uses only one flood in each year. Infrequently, the second highest flood in a given year, which the above rule (i.e., annual floods) omits, may outrank many annual floods." Many hydrologists do not consider a maximum annual peak discharge that is so small that its level does not exceed a certain stage, say the bankfull stage, as a flood. Highest annual peak discharges during the dry years of some rivers in arid and semiarid regions may be so small that an analyst calling them floods may question this approach himself. However, Gumbel [28] defines the flood as "the largest mean daily discharge, measured in volume per unit time, among 365 observations of a calendar year. Whereas any year might produce several inundations, or none at all, there is one, and

*Ph.D. graduate, Civil Engineering Department, Colorado State University.

only one annual flood which need not be an inundation and might even be a drought." Another shortcoming of the annual flood series is that a small number of floods is considered, or a limited number of information about the flood phenomenon is used for a given discharge time series.

The partial duration flood series appears more useful for theoretical analysis than do annual flood series. The preceding objections to the annual flood series are not valid for the partial duration series which considers each flood peak individually. The only drawback of this method of data extraction is that the sequence of flood events might not be considered as fully independent. Consecutive flood hydrographs may be sometimes so close that one flood sets the starting stage for the next flood. However, there are ways to surmount this difficulty, as shown in the insuing text.

The hydrologic part in the design of bridges, culverts, spillways, levees, highway drainage, storm sewers, and similar structures is based on high flows exceeding a certain critical magnitude. This magnitude will be called the base level, the base flood flow, or simply the base in this paper. Discharges below this base are excluded from consideration because of their insignificant effect on a structure. The U.S. Geological Survey determines the base flood stage at the locations of river gauging stations. Wisler and Brater [53, p. 320] state, "Frequently, however, especially on the more important rivers, an arbitrary elevation has often been established, either by the U.S. Army Engineers or by others in authority, that is called flood stage." This is reasonable procedure for distinguishing between large discharges which are floods from those which are not. Kirby [33] also divides hydrographs into floods and nonfloods. Partial duration flood series meets the practical considerations needed for this study and therefore the data from this series will be used for computation.

The method presented here, using the theory of extreme values of a particular class of stochastic processes given by Todorovic [48], discusses the development of a probabilistic model that describes the flood phenomenon. The model is sufficiently general to be applicable to numerous naturally or artificially intermittent stochastic processes important in the geophysical sciences. The simplest form of the model, with flood peak exceedances $\xi_1, \xi_2, \ldots, \xi_\nu$, represents a sequence of independent and identically distributed random variables independent of the sequence of times of the occurrences of these exceedances $\tau(1), \tau(2), \ldots,$ as applied to data from the gauging stations on the Susquehanna River at Wilkes-Barre, Pa. and the Greenbrier River at Alderson, W. Va. Both gauging stations have long records of homogeneous data on flood flows in the form of partial duration series. Observed and theoretical results agree fairly well. The simplest form of the model, however, might not be adequate for some rivers. In these instances, it will be necessary either to develop a new particular model from the general model given or to use periods shorter than a year (seasons, for example) for which the simplest form of the model would still be justifiable.

Acceptance of a certain model for analysis of the largest flood peak must be based on the goals and conditions that are to be fulfilled and satisfied by the model; goodness of fit is a necessary but not a critical condition for acceptance. If goodness of fit were the controlling criterion, (as many engineers are often apt to believe) high-degree polynomials to fit empirical curves would have to be used; this is not the standard practice. The most important criterion is that the model have a sound theory describing the phenomenon, and a maximum extraction of information by the proper estimation techniques. The method presented in this study represents an effort to offer a sound, general approach for analyzing the largest flood peak.

Chapter 2

BRIEF REVIEW OF PERTINENT LITERATURE

The history of the problem of extreme values began with the pioneering work of E. L. Dodd and L. H. C. Tippett. Dodd first studied the largest value for other than the normal distribution (1923), and Tippett calculated the probabilities of the largest value taken from the normal initial distribution for different sample sizes (1925). In 1927, M. Fréchet obtained the second asymptotic distribution of the largest value. He also introduced the notion of a type of initial distribution and showed that a common asymptotic distribution of the largest value may exist for different initial distributions having a common property. In 1928, R. A. Fisher and Tippett presented all three asymptotic distributions of the largest values. Their result concerning the second asymptote was independent of Fréchet's. R. von Mises (1936) and B. Gnedenko (1943) made further contributions by classifying the initial distributions which possess asymptotic distributions of the largest values and by giving the necessary and sufficient conditions for validity of asymptotic distributions of the largest values.

The first book completely devoted to the statistics of extremes was written by E. J. Gumbel in 1958 [29]. Gumbel did much to bring the theory of extreme values to engineers and scientists working in different disciplines. The first asymptotic distribution of the largest value (also called the double exponential distribution) has been used in hydrologic practice much more than the other two asymptotes. Its application to the problem of floods has been particularly widespread.

It is well to mention at this point the limitations of the asymptotic distributions of extremes. They are [29, p. 346]:

(1) The observations from which the extreme values are drawn should be independent.

(2) The observations must be reliable and be made under identical conditions. The initial distribution and the parameter it contains must be the same for each sample.

(3) The number of observations, n , from which the extremes are taken must be large. How large n must be depends on the initial distribution and the degree of precision sought. Unfortunately, one is not always free in the choice of the sample size. In meteorology and hydrology, for example, the day and the year are natural units of periodicity, and the choice of $n = 365$ (days) is imposed.

(4) The initial distributions from which the extreme values are taken must belong to one of the three described types.

Fuller [25] extensively studied the problem of floods in the U.S.A. Using a purely empirical approach he concluded that flood flows should increase as the logarithm of the return period. This was in agreement with an approximate formula for floods derived later by Coutagne [12].

In 1939, W. P. Creager [14] published the result of an investigation of maximum recorded flood peaks in the U.S.A. His purely empirical result, presented in a form of a curve embracing all the records of maximum floods, relates the magnitude of flood to the drainage area. His investigation sought to determine the hydrologic aspects of spillway design. His curve is heavily dependent on the length of flood flow records.

Moran [38] has dealt with the problem of estimating a flood corresponding to a given probability. In his published work he discusses the sources of errors in estimating the shape of the tail (large flows) of a streamflow distribution and presents what he considers the most efficient procedure in estimating the parameters of the assumed distribution.

Hall and Howell [30] discusses the probability that a flood of certain magnitude will be equalled or exceeded one or more times in a given time period. They considered floods as independent events occurring according to the Poisson time invariant distribution.

In 1964, Shane and Lynn [45] presented a probability model for use in the statistical analysis of a partial duration series. Design equations relating three measures of risk to design discharge (recurrence-interval distribution, encounter probability, and expected recurrence interval) were presented. Analysis was based on the time independent Poisson process and the probability theory of sums of a random number of random variables.

A succinct description of the present situation regarding the methods of flood frequency analysis is given by the Water Resources Council of the U.S. Government [4] in a study they made to find a consistent approach for estimating flood frequencies: "Methods of flood frequency analysis, which started about 1914, have developed along divergent lines, with resulting nonuniformity in methods of analysis and, hence, in results. The present state of the art is such that no general agreement has been reached as to preferable techniques, and no standards have been established for design purposes, as have been done in other branches of engineering." In the report, the results obtained by the most commonly used distributions used in flood frequency analysis were compared. The six distributions reviewed were: (1) two-parameter gamma distribution, (2) Gumbel distribution, (3) log-Gumbel distribution, (4) log-normal distribution, (5) log-Pearson Type III distribution, and (6) the Hazen distribution. The distributions were applied to a selected group of ten long-record representative sites in different parts of the U.S.A., and records of maximum annual discharges were used. A quotation from that study is significant, "The statistical consultants had indicated that no unique procedures could be specified as correct for any one method of flood frequency analysis. No single method of testing the computed results against the original data was acceptable to all those on the Work Group, and the statistical consultants could not offer a mathematically rigorous method. It appeared, consequently, that if a choice could not be made solely on statistical grounds, a choice on administrative grounds, for which compelling reasons existed, was justified. This administrative choice was largely governed by the relative values of the results and the tests of conformance that were made." This "administrative choice" resulted in the adoption of the log-Pearson Type III distribution, or the Pearson Type III distribution applied to logarithms of flood peak discharges, as the base method for analyzing flood flow frequencies for federal agencies. It is also stated in the same source [4], "The present state of the art of frequency analysis does not warrant the specification of best procedures for any one method." The Work Group also recommended, "That the choice of a base method should not be considered as final and should not freeze hydrologic practice into any set pattern, either now or in the future. That in view of the increasing importance of frequency analysis in water-resources development, studies should be continued for the purpose of resolving uncertainties, improving methods of analysis, and reviewing all work in this field. That when considered desirable, new techniques or methods should be recommended."

In 1969 Kirby [33] discussed the random occurrence of major floods. He considered flood peaks as the successes or exceedances in a sequence of randomly spaced Bernoulli trials representing the occurrence of hydrograph peaks. The event that a hydrograph peak is a flood is called an exceedance. Kirby adopted a criterion for classifying hydrograph peaks into floods and nonfloods. His model showed that, at sufficiently small exceedance probabilities, the probability distributions of the times between exceedances and the number of exceedances approach those implied by trials from a Poisson process. Kirby, therefore, justified Poisson models of flood occurrence and gave an explanation of observed distributions. However, the model could have been better if a time dependent process for the arrival of flood peaks was used instead of a time independent process. Kirby [33] stated, ".... it does no harm to ignore seasonal variations of parameters and thus assume that the times between the hydrograph peaks are identically distributed as well as independent random variables, ..." The times between hydrograph peaks cannot be considered as identically distributed random variables, because the average number of flood peaks, $\lambda(t)$, in a unit interval of time is, in most cases, a nonlinear function of time. Denoting with $\eta(t)$ the number of flood peaks in an interval of time $[o,t]$, the above statement expressed mathematically reads $P\{\eta(t_1+\Delta t) - \eta(t_1) = k\} \neq P\{\eta(t_2+\Delta t) - \eta(t_2) = k\}$ for $t_1 \neq t_2$.

In 1969 Todorovic [48] obtained the distribution functions of extreme values of a random number of random observations which were valid for any given interval of time. Todorovic's article presents a new approach to the theory of extreme values that is of particular interest in the analysis of extremes of naturally occurring pheonmena where an element of probability is involved. This approach offers new possibilities for a more general analysis of extreme values. Although the number of observations occurring in an interval of time and the results of these observations are both random variables, the distribution functions of extremes are uniquely determined functions. This study will apply this new approach to the problem of flood peaks.

Chapter 3

THEORETICAL CONSIDERATIONS

3.1 Phenomenological Analysis. Given a streamflow hydrograph (Fig. 1) at a specific point along a river, consider only those peaks Q_k , $k = 1, 2, \ldots,$ ν , in some interval of time $[o,t]$ that exceed the base flood flow Q_b.

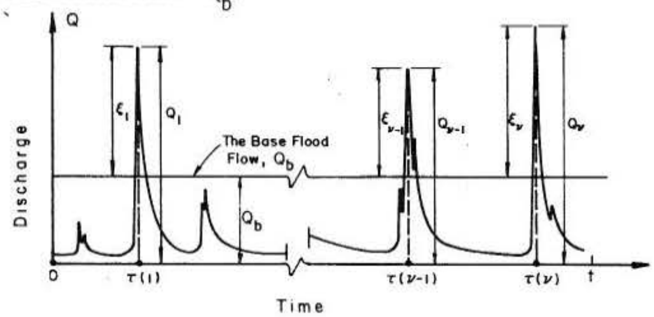

Fig. 1. General streamflow hydrograph of instantaneous discharges at a given point on a river for the time interval $[o,t]$.

As stated in Chapter 1, the ν-th flood peak exceedance ξ_ν in an interval of time $[o,t]$ is defined as

$$\xi_\nu = Q_\nu - Q_b \ , \qquad (1)$$

in which Q_ν is the ν-th total flood peak which has occurred in the time interval $[o,t]$. By definition, $Q_\nu > Q_b$ and $\nu = 0, 1, 2, \ldots$.

In the case of a multipeaked flood hydrograph, such as the hydrograph at $\tau(\nu-1)$ in Fig. 1, only the largest discharge is considered to be the flood peak. This treatment is an approximation of the concept of the independence of flood peaks and the effect on the final result, using this method, is minor. It is possible to separate a complex hydrograph to obtain the independent flood peak but this method would complicate the approach and add nothing significant to its applicability.

Because hydrograph peaks smaller than the base flood flow, Q_b , are not considered as flood peaks, all flows are excluded except the flood peaks (Fig. 2).

The barrier Q_b is the lowest level of the bounded process $\{\xi_t \ ; \ t \geq 0\}$. Therefore, the intermittent process of flood hydrograph tops seen in Fig. 2 is a one-boundary non-negative stochastic process with a period of one year as the time interval.

According to the nature of flood phenomena, the number of flood peak exceedances in an interval of time $[o,t]$, as well as the magnitudes of these exceedances are random variables. Not only the number of flood peak exceedances in $[o,t]$ is a random variable but the times when these exceedances occur are also random variables. With each random variable ξ_ν , where $\xi_\nu > 0$ for all $\nu = 1, 2, \ldots\ldots$, a time $\tau(\nu)$ is associated with the corresponding exceedance (Fig. 3).

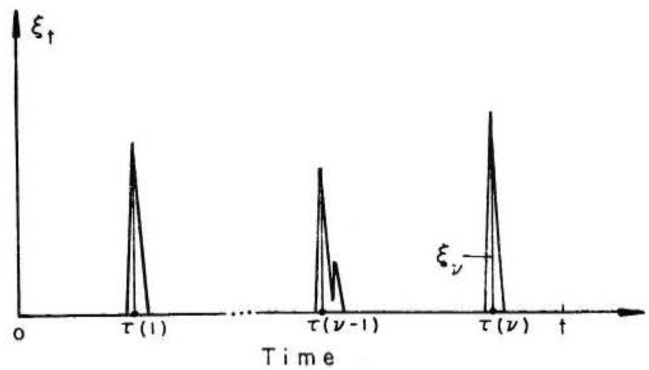

Fig. 2. Intermittent process of flood hydrograph tops, with ξ_t the discharge above the base Q_b .

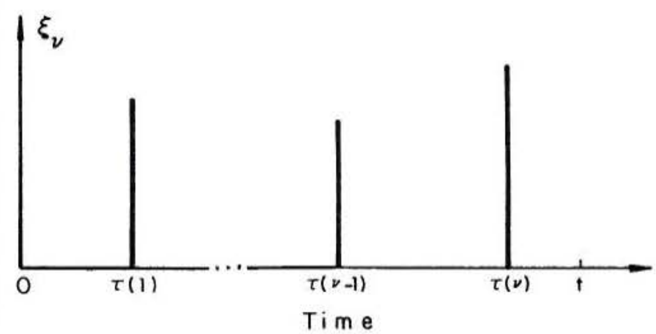

Fig. 3. A realization of the stochastic (discrete, non-negative) process of flood peak exceedances in an interval of time $[o,t]$, with ξ_ν the flood peak exceedance.

Stochastic process of flood peak exceedances is a discrete-parameter stochastic process, $\{\xi_\nu \ ; \ \nu = 0, 1, 2, \ldots\}$. For simplicity, flood peak exceedances ξ_ν, $\nu = 0, 1, 2, \ldots$ will simply be called exceedances.

3.2 Distribution of the Number of Exceedances. Noting the distribution of the number of exceedances plays an important role because this method considers simultaneously the magnitude of these exceedances and their number within a given time interval $[o,t]$. Using the results obtained by Todorovic [48], the distribution function of the number of exceedances is summarized in this section.

Denote by $\eta(t)$ the number of exceedances in an interval of time $[o,t]$. By definition, $\eta(t)$ may be 0, 1, ..., and for all $t \geq 0$ and $\Delta t > 0$, $\eta(t) \leq (t+\Delta t)$. In general, $\eta(t)$ depends also on the parameter Q_b, and for a fixed t, $\eta(t)$ is a non-increasing function of Q_b. However, for a flood analysis at a given point on a river, Q_b can have a physical meaning. As defined in the previous section, Q_b represents the base flood flow and therefore is treated as a selective parameter throughout this study.

In the event that exactly ν exceedances occur in $[o,t]$, denoted as

$$E_\nu^t = \{\eta(t) = \nu\} \ , \tag{2}$$

then

$$E_i^t \cap E_j^t = \theta \text{ and } \bigcup_{\nu=0}^{\infty} E_\nu^t = \Omega \quad \text{for all} \quad i \neq j = 0,1, \ldots$$

in which ν is a particular numerical value of the random variable $\eta(t)$, θ stands for the impossible event, and Ω stands for the certain event. Let $\Lambda(t)$ stand for the expected value of $\eta(t)$. Then

$$\Lambda(t) = \sum_{\nu=1}^{\infty} \nu \, P(E_\nu^t) \quad . \tag{3}$$

Because of seasonal variation, $\Lambda(t)$ is, in most cases, a nonlinear function of time.

Writing $F_k(t) = P\{\tau(k) \leq t\}$, where $\tau(k)$ is the time of occurrence of the k-th exceedance, then [48]

$$P(E_k^t) = F_k(t) - F_{k+1}(t) \quad . \tag{4}$$

From Eq. 4, one obtains

$$F_k(t) = \sum_{j=k}^{\infty} P(E_j^t) \quad . \tag{5}$$

Denote by E_k^t the event that exactly k exceedances occur in a fixed time interval $[o,t]$, and denote by $E_1^{t,t+\Delta t}$ the event that only one exceedance occurs in a time interval $[t,t+\Delta t]$, in which Δt is the length of the interval. Under certain very general assumptions one may show that the probabilities $P(E_k^t)$, k=0, 1, ..., satisfy the following system of differential equations,

$$\left.\begin{aligned} \frac{dP(E_k^t)}{dt} &= \lambda_{k-1}(t)P(E_{k-1}^t) - \lambda_k(t)P(E_k^t) \quad k=1,2,\ldots \\ \frac{dP(E_o^t)}{dt} &= -\lambda_o(t)P(E_o^t) \end{aligned}\right\} \tag{6}$$

in which

$$\lambda_k(t) = \lim_{\Delta t \to 0} \frac{P(E_1^{t,t+\Delta t} \mid E_k^t)}{\Delta t} \tag{7}$$

and $\quad E_1^{t,t+\Delta t} = \{\eta(t+\Delta t)-\eta(t)=1\}$.

It is not difficult to verify that system (6) has the solution

$$P(E_o^t) = \exp\{-\int_o^t \lambda_o(s)ds\} \ , \tag{8}$$

$$P(E_k^t) = \exp\{-\int_o^t \lambda_k(s)ds\} \int_o^t \lambda_{k-1}(t_1)$$

$$\exp\{\int_o^{t_1} [\lambda_k(s) - \lambda_{k-1}(s)]ds\}$$

$$\int_o^{t_1} \ldots \int_o^{t_{k-1}} \lambda_o(t_k) \exp\{\int_o^{t_k} [\lambda_1(s)-\lambda_o(s)]ds\} dt_k \, dt_{k-1} \ldots dt_1 \ . \tag{9}$$

Generally, a simple expression for each $P(E_k^t)$ in terms of $\{\lambda_k(t)\}$ is not possible. However, several special cases have been solved and are given in reference [49].

The case considered to be of revelance in flood analysis is when

$$\lambda_k(t) \equiv \lambda(t) \quad \textbf{(independent of } k\textbf{)} \quad .$$

Under this condition

$$P(E_k^t) = \{\int_o^t \lambda(s)ds\}^k \exp\{-\int_o^t \lambda(s)ds\}/k! \tag{10}$$

which is the time dependent Poissonian process. From the mathematical expectation given by Eq. 3, $\Lambda(t)$ becomes

$$\Lambda(t) = \int_o^t \lambda(s)ds \quad . \tag{11}$$

Equation 10 can also be written as

$$P(E_k^t) = [\Lambda(t)]^k \exp\,[-\Lambda(t)]/k! \ . \tag{12}$$

In the preceding equations, $\lambda(t)$ is the mean number of exceedances in a time unit. It can also be called the density of the number of exceedances in a unit of time. Hereafter, for the two rivers used as examples in the application of the method presented, $\lambda(t)$ represents the mean number of exceedances per day, and is a deterministic periodic function of time having a one-year period.

Equation 5 represents the distribution function of the time of the k-th exceedance, which can also be written as

$$F_k(t) = 1 - \sum_{j=0}^{k-1} P(E_j^t) \quad . \tag{13}$$

Denote by $f_k(t)$ the corresponding density function. Taking into account Eq. 10, and after the differentiation of the function $F_k(t)$ with respect to t, it follows [49] that

$$f_k(t) = \frac{\lambda(t)}{\Gamma(k)} \{\int_0^t \lambda(s)ds\}^{k-1} \cdot \exp\{-\int_0^t \lambda(s)ds\} \quad , \text{ for } t \geq 0 \quad . \tag{14}$$

3.3 Distribution of the Largest Exceedance. Another random variable of interest in flood analysis is the largest exceedance. Consider an interval of time $[o,t]$ and denote by $\chi(t)$ the largest exceedance, ξ_ν, in this interval. Because the number of exceedances in $[o,t]$ is a random variable depending on time t, $\chi(t)$ is defined as

$$\chi(t) = \sup_{\tau(\nu) \leq t} \xi_\nu \quad . \tag{15}$$

By definition it follows that for every $t \geq 0$ and $\Delta t > 0$,

$$\chi(t) \leq \chi(t+\Delta t) \quad .$$

in which $\chi(t)$ is a stochastic process of non-decreasing (step) sample functions (Fig. 4). The process $\chi(t)$ represents the essence of the theoretical considerations used in this study. The corresponding distribution function of $\chi(t)$ is denoted as $F_t(x)$, i.e.,

$$F_t(x) = P\{\chi(t) \leq x\} \quad , \text{ for } t \geq 0 \quad , \text{ and } x \geq 0 \quad .$$

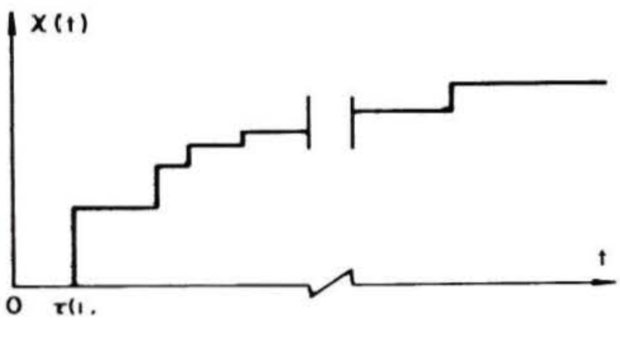

Fig. 4. A sample function of the process $\chi(t)$, as the largest exceedance.

Todorovic [48] obtained the expression for the function $F_t(x)$ as the mathematical expectation of the conditional probability $P\{\sup_{\tau(\nu) \leq t} \xi_\nu \leq x | \eta(t)\}$. His result is [48, Eq. 3.4, p. 1060]

$$F_t(x) = P(E_o^t) + \sum_{k=1}^{\infty} P\left[\bigcap_{\nu=1}^{k} (\xi_\nu \leq x) \cap E_k^t\right] \quad . \tag{16}$$

The graph of the distribution function $F_t(x)$ is given in Fig. 5. The same result may be derived by a simpler approach:

$$F_t(x) = P\{\chi(t) \leq x\} = P\{[\chi(t) \leq x] \cap \Omega\} \quad ;$$

because the events E_0^t, $E_1^t, \ldots$, are mutually exclusive and exhaustive one can continue and obtain

$$F_t(x) = P\left\{[\chi(t) \leq x] \cap \left[\bigcup_{k=0}^{\infty} E_k^t\right]\right\} = \sum_{k=0}^{\infty} P\left\{[\chi(t) \leq x] \cap E_k^t\right\}$$

or

$$F_t(x) = \sum_{k=0}^{\infty} P\left\{\left[\bigcap_{\nu=0}^{k} (\xi_\nu \leq x)\right] \cap E_k^t\right\} \quad , \tag{16a}$$

which equals the expression of Eq. 16. The distribution function given by Eq. 16 can be interpreted as the probability that all exceedances, ξ_ν, in an interval of time $[o,t]$ will be less than or equal to x. Equation 16 represents the most general expression for the distribution function of the largest exceedance within any given time interval $[o,t]$. If $x=0$, it follows from Eq. 16 that

$$F_t(o) = P(E_o^t) \quad , \tag{17}$$

which can be interpreted as the probability that there will be no exceedances $(\nu = 0)$ in a given time interval $[o,t]$.

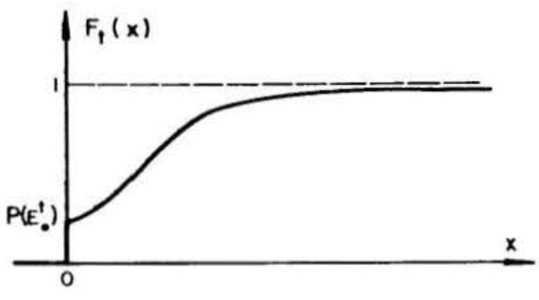

Fig. 5. Distribution function of the largest exceedance for a given interval of time $[o,t]$.

It follows from the foregoing discussion that the distribution function $F_t(x)$ is not differentiable at point $x = 0$. The boundary conditions of the distribution function $F_t(x)$ at $x = 0$ and $x = \infty$ are satisfied, namely

$$F_t(o) = P(E_o^t) \qquad \text{and} \qquad F_t(\infty) = 1 \quad .$$

Equation 16, however, cannot be used directly for a specific problem unless it is reduced to one of its particular forms, i.e., unless one determines the probabilities $P\left[\bigcap_{\nu=1}^{k} (\xi_\nu \leq x) \cap E_k^t\right]$ pertaining to a given case.

Consider now a particular case in which the exceedances ξ_ν, $\nu = 0, 1, 2, \ldots$, occurring in an interval of time $[o,t]$ are independent and identically distributed random variables with the random vectors $\{\xi_1, \xi_2, \ldots\ldots, \xi_k\}$ and $\{\tau(k), \tau(k+1)\}$ mutually independent for all $k = 1,2,\ldots.$ For this case, Eq. 16 becomes

$$F_t(x) = \sum_{k=0}^{\infty} \{[H(x)]^k \cdot P(E_k^t)\} \quad , \tag{18}$$

or

$$F_t(x) = P(E_o^t) + \sum_{k=1}^{\infty} \{[H(x)]^k \cdot P(E_k^t)\} \quad , \tag{19}$$

in which $H(x)$ is the distribution function of all exceedances in a given interval of time $[o,t]$. This case can also be explained another way. Given some interval of time $[o,t]$ and assuming that there exists the common distribution function $H(x)$ of all exceedances within $[o,t]$, the event that there will be k exceedances is independent of the event that all k exceedances will be less than or equal to x.

Considering the applicability of Eq. 19 in flood analysis, the first question requiring attention is the independence of the events $\bigcap_{\nu=1}^{k} (\xi_\nu \leq x)$ and E_k^t. Popularly speaking, one might ask the question: Does the event E_k^t that exactly k exceedances have occurred in an arbitrary but fixed interval of time $[o,t]$ permit any inference about the magnitude of each of the k exceedances? or, does the event that all k exceedances in $[o,t]$ less than or equal to x permit any inference about the time of occurrence of the k-th exceedance? In general, the two events might not be fully independent in some instances. However, even in this case, a question remains whether the degree of dependence of the two events is significant from the position of practical applications. The general case involving the dependence of the events $\bigcap_{\nu=1}^{k} (\xi_\nu \leq x)$ and E_k^t is not considered in this study, but is left for future investigations. The case pertaining to the present study is the case when the two events are, or can be assumed to be independent. The second question arising in this study is related to distribution functions of the exceedances. This problem can be handled using available data for any particular example considered. It is possible that identical distribution of exceedances exists throughout the year for some rivers but not for others. This depends whether the flood peaks are produced only by rainstorms, by rainstorms and the melting of snow and ice, or only by the melting of snow or ice. Therefore, any particular case requires individual investigation. It is not injudicious to treat exceedances occurring in a short interval of time as identically distributed random variables. The problem is to determine, for a particular case, the length of this interval within which the notion of identical distribution of exceedances is justifiable. For some rivers this interval may be a month, or a season, but for others it may extend over the whole year.

The case analyzed in this study is the exceedances ξ_ν, $\nu = 0,1,2,\ldots$, which are independent and identically distributed random variables, with the random vectors $\{\xi_1, \xi_2, \ldots\ldots, \xi_k\}$ and $\{\tau(k), \tau(k+1)\}$ mutually independent for all $k = 1,2,\ldots$

A theoretically derived expression for the distribution function, $P(E_k^t)$, of the number of exceedances in an interval of time $[o,t]$ is given by Eq. 12 as a time-dependent Poissonian process. The other distribution function that requires investigation is the distribution function, $H(x)$, of the magnitude of all exceedances for the same given interval of time $[o,t]$. Determining the distribution function $H(x)$ may or may not be purely a problem of estimation. At the present state of flood analysis there are no, or few, theoretical grounds that indicate the form of the distribution of exceedances. Two probability laws have played an important role in connection with the magnitude of flood peaks. These are the gamma and the exponential probability laws, [4], [29], [38], [45], etc. Because the exponential distribution is a particular case of the gamma distribution, the family of (two-parameter) gamma distributions is used in the sequel as the common distributions of exceedances. Therefore, using Eq. 19, the family of gamma distributions in combination with time-dependent Poissonian process is used in the ensuing text for the study of the theoretical distributions of the largest exceedance.

Chapter 4

STOCHASTIC PROCESS OF THE LARGEST EXCEEDANCE IN GAMMA-DISTRIBUTED EXCEEDANCES

The distribution function, $H(x)$, of exceedances in two-parameter gamma probability distribution is considered in this chapter. The gamma probability distribution with parameters α and β, and both parameters greater than zero, is generally specified by the probability density function

$$h(x) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \quad \text{for} \quad x \geq 0, \quad \text{and} \quad h(x) = 0.$$

For a particular case one can estimate parameters α and β from observations using the method of maximum likelihood. However, it is difficult to obtain mathematically convenient expressions for the distribution function and its moments of the largest exceedance for a given time interval $[o,t]$ if parameter α takes on noninteger values. In the case of noninteger values one can resort to digital computer integrations or similar techniques to obtain approximate solutions. However, the development of a computer integration scheme for a noninteger α is outside the scope of this study. This paper presents the analytical solution of the considered case when parameter α takes on integer values.

4.1 General Solution When Parameter α is a Positive Integer. The common distribution function of exceedances ξ_{ν}, $\nu = 0, 1, 2, \ldots$, in a given interval of time $[o,t]$, when α is a positive integer and $\Gamma(\alpha) = (\alpha-1)!$ is

$$H(x) = \int_{o}^{x} \frac{\beta^{\alpha}}{(\alpha-1)!} u^{\alpha-1} e^{-\beta u} du , \qquad \text{for} \qquad x \geq 0 , \tag{20}$$

in which α is the shape parameter taking on values of positive integers, and β^{-1} is the scale parameter. The integration of the above expression gives

$$H(x) = 1 - e^{-\beta x} \sum_{i=0}^{\alpha-1} \frac{(\beta x)^i}{i!} . \tag{21}$$

Therefore, the distribution function of the largest exceedance for a given interval of time $[o,t]$ is

$$F_t(x) = e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} \left[1 - e^{-\beta x} \sum_{i=0}^{\alpha-1} \frac{(\beta x)^i}{i!}\right]^k , \tag{22}$$

which can be also written as

$$F_t(x) = \exp \left[-\lambda t\, e^{-\beta x} \sum_{i=0}^{\alpha-1} \frac{(\beta x)^i}{i!}\right] . \tag{23}$$

The distribution function $F_t(x)$ is discontinuous at point $x=0$ for all $t>0$. For $x>0$ the function $F_t(x)$ is continuous. It follows from Eq. 22 or Eq. 23 that

$$F_t(o) = \exp(-\lambda t) . \tag{24}$$

The distribution function $F_t(x)$ is a mixture, because it is not always continuous nor is it a step function. Accordingly, the distribution function $F_t(x)$ is not differentiable at $x=0$, and the corresponding density function $f_t(x)$ is defined in a special way at this point. The probability density function of the largest exceedance, for a given interval of time $[o,t]$, is

$$f_t(x) = 2\, e^{-\lambda t}\, \delta(x) + \frac{\partial F_t(x)}{\partial x} , \tag{25}$$

which is valid for $x \geq 0$. In Eq. 25, $\delta(x)$ is the Dirac delta function (or the symmetrical unit - impulse function) of a real variable x defined by

$$\int_a^b f(\xi)\delta(\xi-X)d\xi = \begin{cases} 0 & \text{if } X<a \text{ or } X>b \\ \frac{1}{2} f(X) & \text{if } X=a \text{ or } X=b \\ f(X) & \text{if } a<X<b \end{cases} , \quad (a < b), \tag{26}$$

in which $f(x)$ is an arbitrary function continuous for $x = X$, with

$$\delta(x) = 0 \qquad x \neq 0 , \qquad \int_{-\infty}^{0} \delta(\xi)d\xi = \int_{0}^{\infty} \delta(\xi)d\xi = \frac{1}{2}$$

and

$$\int_{-\infty}^{\infty} \delta(\xi)d\xi = 1 .$$

At $x = 0$, the derivative of $F_t(x)$ may be defined arbitrarily.

The distribution function, $F_t(x)$, of the largest exceedance, $\chi(t)$, for this case can also be specified as

$$\begin{aligned} dF_t(x) &= 0 , && x < 0 \\ &= e^{-\lambda t} , && x = 0 \\ &= \frac{\partial F_t(x)}{\partial x} dx , && x > 0 . \end{aligned}$$

The moment generating function of the largest exceedance for a given time interval $[o,t]$ is

$$\psi_t(u) = \int_{o}^{\infty} e^{ux}\, dF_t(x) .$$

After considering this expression one may write,

$$\psi_t(u) = \int_0^\infty e^{ux}\, 2e^{-\lambda t}\, \delta(x)dx + \int_{0+}^\infty e^{ux} \frac{\partial F_t(x)}{\partial x} dx \quad ,$$

which, with Eq. 26 in mind, gives

$$\psi_t(u) = e^{-\lambda t} + \int_{0+}^\infty e^{ux} \frac{\partial F_t(x)}{\partial x} dx \quad . \tag{27}$$

The same equation can be obtained also by the Stieltjes integral:

$$\int_0^\infty e^{ux}\, dF_t(x) = e^{-\lambda t}\, [e^{ux}]_{x=0} + \int_{0+}^\infty e^{ux} \frac{\partial F_t(x)}{\partial x} dx$$

$$\int_0^\infty e^{ux}\, dF_t(x) = e^{-\lambda t} + \int_{0+}^\infty e^{ux} \frac{\partial F_t(x)}{\partial x} dx \quad ,$$

which equals Eq. 27. Continuing, the moment generating function of the process $\chi(t)$ is obtained:

$$\psi_t(u)=e^{-\lambda t} + \frac{\beta^\alpha}{(\alpha-1)!}\, e^{-\lambda t} \sum_{k=1}^\infty \frac{(\lambda t)^k}{(k-1)!} \int_{0+}^\infty e^{ux-\beta x}$$

$$\left[1-e^{-\beta x} \sum_{i=0}^{\alpha-1} \frac{(\beta x)^i}{i!}\right]^{k-1} x^{\alpha-1}\, dx \quad , \tag{28}$$

or

$$\psi_t(u)=e^{-\lambda t} + \frac{\lambda t}{(\alpha-1)!}\, \beta^\alpha \int_{0+}^\infty x^{\alpha-1} \exp\left[ux-\beta x-\lambda t e^{-\beta x}\right.$$

$$\left.\sum_{i=0}^{\alpha-1} \frac{(\beta x)^i}{i!}\right] dx \quad . \tag{29}$$

All absolute moments, for this case of positive integers, are deduced from the moment generating function $\psi_t(u)$:

$$E\chi^m(t)= \frac{\lambda t}{\beta^m s!}\, e^{-\lambda t} \sum_{k=1}^\infty \frac{(\lambda t)^{k-1}}{(k-1)!} \int_0^\infty y^{s+m}\, e^{-y}$$

$$\left[1 - e^{-y} \sum_{i=0}^{s} \frac{y^i}{i!}\right]^{k-1} dy \quad . \tag{30}$$

Here the substitutions $s = \alpha - 1$, and $y = \beta x$ are introduced for simplicity, and $m = 1, 2, \ldots,$ is the order of the absolute moment. Equation 30 can be also written as

$$E\chi^m(t) = \frac{\lambda t}{\beta^m s!} \int_0^\infty y^{s+m} \exp\left[-y - \lambda t\, e^{-y} \sum_{i=0}^{s} \frac{y^i}{i!}\right] dy \quad . \tag{31}$$

The asymptotic value of the m-th absolute moment when α is very large is

$$E\chi^m(t) \doteq \frac{\lambda t}{\beta^m}\, e^{-\lambda t}\, (\alpha + m)^m \quad . \tag{32}$$

The derivation of Eq. 32 is given in Appendix 1.

4.2 The Special Case of Exponentially Distributed Exceedances ($\alpha=1$). Exponential distribution plays an important role in the theory of extreme values. In the classical theory of extremes [29], all statements which are exact for the exponential distribution are asymptotically valid for all distributions belonging to the exponential type. Exponential distribution is encountered in many areas of geophysics. It describes the decay process of many phenomena in modern physics, hydrology, sanitary engineering, etc. This distribution is also observed in certain flood phenomena.

For the case under discussion, the common distribution function of all exceedances ξ_ν, ν 0,1,2,..., in some interval of time $[o,t]$ is

$$H(x) = 1 - e^{-\beta x} \qquad , \; x \geq 0 \quad . \tag{33}$$

Therefore, the distribution function of the largest exceedance in the same time interval $[o,t]$ is

$$F_t(x) = e^{-\lambda t} \sum_{k=0}^\infty \left[\lambda t(1-e^{-\beta x})\right]^k \frac{1}{k!} \quad ,$$

which is reduced to

$$F_t(x) = \exp[-\lambda t\, e^{-\beta x}] \quad . \tag{34}$$

It can be observed that when $t=\lambda=\beta=1$, Eq. 34 gives the first asymptotic distribution of the largest reduced extreme given in [29], and if this first asymptotic distribution is written in the form

$$F(x) = \exp\{-\exp[-\beta(x-u]\} = \exp[-\, e^{\beta u}\, e^{-\beta x}] \quad ;$$

it then represents a particular case of Eq. 34 obtained for $\lambda t = \exp(\beta u)$.

It can be easily shown that

$$\int_0^\infty f_t(x)dx = 2e^{-\lambda t} \int_0^\infty \delta(x)dx + (1 - e^{-\lambda t}) = 1 \quad .$$

The condition $\frac{\partial}{\partial x}\left[\frac{\partial F_t(x)}{\partial x}\right] = 0$ gives the value of the mode $\tilde{x}$, in the continuous part of $F_t(x)$, of the largest exceedance:

$$\tilde{x} = \frac{1}{\beta} \ln(\, t) \quad , \tag{35}$$

with ln denoting the natural logarithm, for which

$$f_t(\tilde{x}) = \beta e^{-1} \tag{36}$$

and

$$F_t(\tilde{x}) = e^{-1} \quad . \tag{37}$$

Using the condition $\frac{\partial^2}{\partial x^2}\left[\frac{\partial F_t(x)}{\partial x}\right] = 0$, one obtains the expression $\lambda t \exp(-\beta x_{1,2}) = \frac{1}{2}(3 \pm \sqrt{5})$, and the abscissas of the inflection points as

$$x_{1,2} = \frac{1}{\beta}\, [\ln(2\lambda t) - \ln(3 \pm \sqrt{5})] \quad , \tag{38}$$

for which

$$f_t(x_1) = 0.19098\ \beta \qquad (39)$$

$$f_t(x_2) = 0.26070\ \beta\ . \qquad (40)$$

The probability of $x_1 \le x \le x_2$ is

$$P\{x_1 \le x \le x_2\} = 0.60957. \qquad (41)$$

The distance between the abscissas of inflection points, $x_2 - x_1$, is

$$x_2 - x_1 = 1.92484\ \beta^{-1} \qquad (42)$$

and is independent of the time interval $[o,t]$.

The mode $\tilde{x}$ of the largest exceedance is located symmetrically with respect to the abscissas $x_{1,2}$ of the inflection points, because

$$\frac{x_1 + x_2}{2} = \beta^{-1}\ \ln(\lambda t) = \tilde{x}\ . \qquad (43)$$

The median $\breve{x}$ of the largest exceedance is

$$\breve{x} = \tilde{x} + 0.3665\ ^{-1}\ . \qquad (44)$$

The density function $f_t(x)$ of the process $\chi(t)$ is shown in Fig. 6.

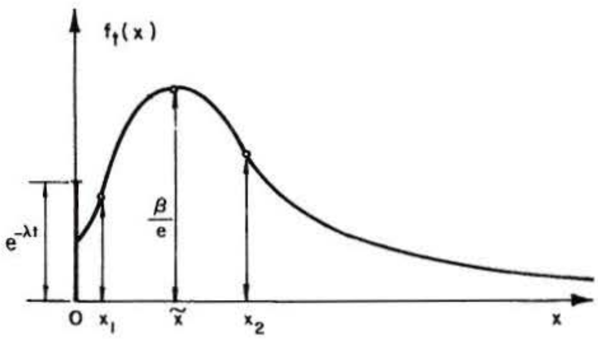

Fig. 6. Probability density function of the largest exceedance for a given time interval $[o,t]$, and for exponentially distributed exceedances.

A further analysis gives

$$g_1(x,t) = \beta\lambda t\ \exp[-\beta x - \lambda t e^{-\beta x}]\ , \qquad (45)$$

with $g_1(o,t) = \beta\lambda t\ \exp(-\lambda t)$, and

$$\lim_{t\to 0} g_1(o,t) = \lim_{t\to\infty} g_1(o,t) = 0\ .$$

Because $\frac{\partial}{\partial t} g_1(o,t) = 0$ for $t = \lambda^{-1}$, it follows that $g_1(o,\lambda^{-1}) = \beta e^{-1}$.

The function $g_1(o,t)$ is convex toward the t-axis when $t\to\infty$ because the function and its second derivative are of the same sign when $t\to\infty$. When $t\to 0$ the function $g_1(o,t)$ and its second derivative are of different signs, therefore the function $g_1(o,t)$ is concave toward the t-axis for small values of t .

The function $g_1(o,t)$ has only one inflection point, in which the abscissa is $2\lambda^{-1}$ and $g_1(o,2\lambda^{-1}) = 2\beta e^{-2}$. With this additional analysis, the function $f_t(x)$ is depicted in Fig. 7 for $x > 0$ and for exponentially distributed exceedances.

The moment generating function of the largest exceedance, for the case under consideration, is

$$\psi_t(u) = e^{-\lambda t} + e^{-\lambda t} \sum_{k=1}^{\infty} \frac{(\lambda t)^k}{k!}$$

$$\int_{0+}^{\infty} k\left(1 - e^{-\beta x}\right)^{k-1} \beta e^{-\beta x}\ e^{ux}\ dx\ .$$

Integrated, this becomes

$$\psi_t(u) = e^{-\lambda t} + e^{-\lambda t} \sum_{k=1}^{\infty} (\lambda t)^k \frac{\Gamma(1-\frac{u}{\beta})}{\Gamma(k+1-\frac{u}{\beta})}\ . \qquad (46)$$

Differentiating the moment generating function $\psi_t(u)$ of the process $\chi(t)$ with respect to u and setting $u=0$, one obtains the absolute moments of the probability distribution given by Eq. 34. The first three absolute moments are

$$E\chi(t) = \beta^{-1}\ e^{-\lambda t} \sum_{k=1}^{\infty} \frac{(\lambda t)^k}{k!} \sum_{i=1}^{k} \frac{1}{i}\ , \qquad (47)$$

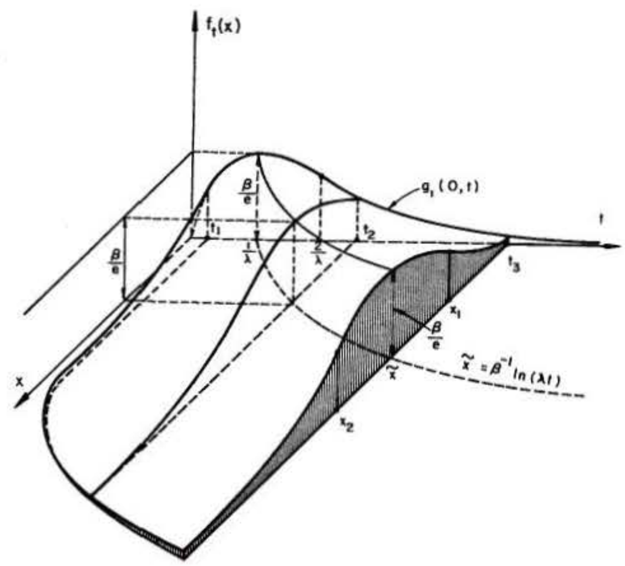

Fig. 7. Probability density function of the largest exceedance, given for $x>0$, for different values of the time interval $[o,t]$ and for exponentially distributed exceedances.

$$E\chi^2(t) = \beta^{-2}\ e^{-\lambda t} \sum_{k=1}^{\infty} \frac{(\lambda t)^k}{k!} (k+1) \sum_{i=1}^{k} \frac{1}{i^2} \qquad (48)$$

and

$$E\chi^3(t) = \beta^{-3} e^{-\lambda t} \sum_{k=1}^{\infty} \frac{(\lambda t)^k}{k!} (k+1) \left[3 \sum_{i=1}^{k} \frac{1}{i^3} + \sum_{j=1}^{k} \sum_{\substack{i=1 \\ i \neq j}}^{k} \frac{1}{j} \frac{1}{i^2} \right] . \tag{49}$$

To check this last computation, the first absolute moment $E_\chi(t)$ of the largest exceedance, for exponentially distributed exceedances, is also computed by definition. The computation is:

$$E\chi(t) = \int_{o+}^{\infty} x \, d\, F_t(x) \quad ,$$

$$E\chi(t) = e^{-\lambda t} \sum_{k=1}^{\infty} \frac{(\lambda t)^k}{k!} \int_{o+}^{\infty} k \left(1 - e^{-\beta x}\right)^{k-1} \beta e^{-\beta x} \, x \, dx$$

$$= e^{-\lambda t} \sum_{k=1}^{\infty} \frac{(\lambda t)^k}{k!} \beta k \sum_{i=0}^{k-1} (-1)^i \binom{k-1}{i} \int_{o+}^{\infty} x \, e^{-\beta x(i+1)} \, dx$$

$$= \beta^{-1} e^{-\lambda t} \sum_{k=1}^{\infty} \frac{(\lambda t)^k}{(k-1)!} \sum_{i=0}^{k-1} (-1)^i \binom{k-1}{i} \frac{1}{(i+1)^2} \quad .$$

Considering the relations

$$\frac{1 - (1-x)^n}{x} = \frac{1}{x} \left[1 - \sum_{i=0}^{n} (-1)^i x^i \binom{n}{i} \right]$$

$$\int_{o}^{1} \frac{1 - (1-x)^n}{x} dx = 1 + \frac{1}{2} + \frac{1}{3} + \ldots$$

$$+ \frac{1}{n} = \sum_{i=1}^{n} (-1)^{i-1} \binom{n}{i} \frac{1}{i}$$

$$\binom{n}{1} \frac{(n-1)}{2^2} - \binom{n}{2} \frac{(n-2)}{3^2} + \binom{n}{3} \frac{(n-3)}{4^2} - \ldots$$

$$+ (-1)^n \binom{n}{n-1} \frac{1}{n^2} = n - \sum_{i=1}^{n} (-1)^{i-1} \binom{n}{i} \frac{1}{i}$$

$$(n+1) \left[\binom{n}{1} \frac{1}{2^2} - \binom{n}{2} \frac{1}{3^2} + \binom{n}{3} \frac{1}{4^2} - \ldots + (-1)^n \binom{n}{n-1} \frac{1}{n^2} \right] =$$

$$\sum_{i=1}^{n-1} (-1)^{i-1} \binom{n}{i} \frac{1}{(i+1)} + n - \sum_{i=1}^{n} (-1)^{i-1} \binom{n}{i} \frac{1}{i}$$

$$\sum_{i=1}^{n-1} (-1)^{i-1} \binom{n}{i} \frac{1}{(i+1)} = 1 - \frac{1 - (-1)^n}{(n+1)} \quad , \text{ and}$$

substituting $k-1$ for n, one obtains

$$\sum_{i=0}^{k-1} (-1)^i \binom{k-1}{i} \frac{1}{(i+1)^2} = \frac{1}{k} \sum_{i=1}^{k} \frac{1}{i} \quad .$$

Finally the first absolute moment of the largest exceedance, for exceedances exponentially distributed is

$$E\chi(t) = \beta^{-1} e^{-\lambda t} \sum_{k=1}^{\infty} \frac{(\lambda t)^k}{k!} \sum_{i=1}^{k} \frac{1}{i} \quad ,$$

which is in agreement with Eq. 47.

The variance $\sigma_t^2(x)$ of the probability distribution given by Eq. 34 was determined with the aid of a computer. The result is given in the form $\beta^2\sigma_t^2(x) = f(\lambda t)$. Summations over k were made for 1 through 100, and also for 1 through 150. Because of the strong convergence of the terms given on the right side of Eq. 47 and Eq. 48, the results by the two approaches were the same for the first three decimal places. The results are given in Fig. 8 and in Table 13 (Appendix 1).

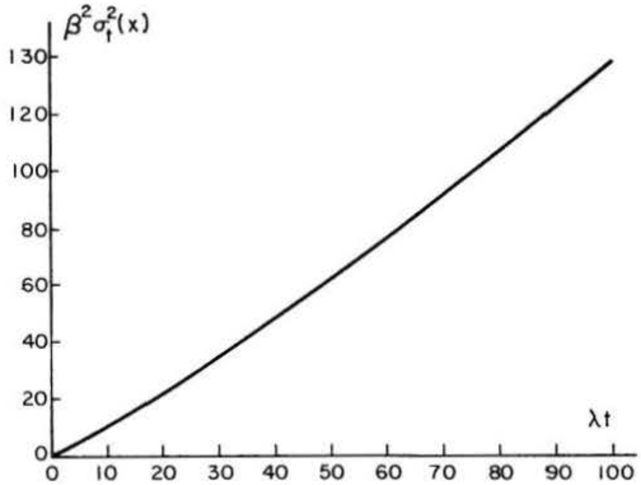

Fig. 8. The variances $\sigma^2{}_t(x)$ of the largest exceedance, for exponentially distributed exceedances, given in the form $\beta^2\sigma_t^2(x)=f(\lambda t)$.

4.3 The Special Case of Gamma-Distributed Exceedances with $\alpha=2$. The common distribution function of all exceedances ξ_ν , $\nu=0,1,2,\ldots$, in an interval of time $[o,t]$, for this case, is

$$H(x) = 1 - e^{-\beta x} (\beta x + 1), \quad x \geq 0 \quad . \tag{50}$$

Therefore, the distribution function of the largest exceedance, in the same time interval $[o,t]$, is

$$F_t(x) = e^{-\lambda t} \sum_{k=0}^{\infty} \left[1 - (1+\beta x) e^{-\beta x} \right]^k \frac{(\lambda t)^k}{k!} \quad . \tag{51}$$

Equation 51 can also be written as

$$F_t(x) = \exp\left[-\lambda t (1+\beta x) e^{-\beta x} \right] \quad . \tag{52}$$

The distribution function given by Eq. 52 is depicted in Fig. 5. Here again, for $x = 0$, the relation

$$F_t(o) = \exp(-\lambda t) \tag{53}$$

represents the discrete part of the function. The term $F_t(o)$ is the probability that no exceedance will occur within an interval of time $[o,t]$.

The probability density function of the largest exceedance, for gamma-distributed exceedances, with $\alpha = 2$, is

$$f_t(x) = 2e^{-\lambda t}\delta(x) + \beta^2 \lambda t x \exp\left[-\beta x - \lambda t (1+\beta x) e^{-\beta x} \right], \tag{54}$$

which is valid for $x \geq 0$. The Dirac delta function, $\delta(x)$, is defined in section 4.2.

It must be verified that the non-negative function $f_t(x)$ satisfies the condition

$$\int_0^\infty f_t(x)dx = 2e^{-\lambda t}\int_0^\infty \delta(x)dx + e^{-\lambda t}\sum_{k=1}^\infty \frac{(\lambda t)^k}{(k-1)!}\int_0^\infty \beta^2 xe^{-\beta x}\left[1-(1+\beta x)e^{-\beta x}\right]^{k-1}dx$$

$$= e^{-\lambda t} + e^{-\lambda t}\sum_{k=1}^\infty \frac{(\lambda t)^k}{(k-1)!}\cdot\frac{1}{k}$$

$$= e^{-\lambda t} + e^{-\lambda t}(e^{\lambda t} - 1) = 1 \quad .$$

The following is an examination of the function $f_t(x)$ for $x > 0$. Denote

$$g_2(x,t) = \beta^2\lambda tx \exp\left[-\beta x-\lambda t(1+\beta x)e^{-\beta x}\right] \quad . \tag{55}$$

Then, $g_2(o,t) = 0$, and

$$\frac{\partial}{\partial x}g_2(x,t) = 0 \quad \text{when} \quad t = \frac{1}{\beta\lambda x}e^{\beta x} \quad . \tag{56}$$

If t from Eq. 56 is inserted into Eq. 55, the result is

$$\max g_2(x,t) = \beta \exp\left[-\frac{(1+\beta x)}{\beta x}\right] \quad . \tag{57}$$

Continuing, one obtains

$$\lim_{x\to 0}\left[\max g_2(x,t)\right] = 0 \quad , \tag{58}$$

$$\lim_{x\to\infty}\left[\max g_2(x,t)\right] = \beta e^{-1} \quad , \tag{59}$$

and the function given by Eq. 57 has no finite extreme because $\frac{d}{dx}[\max g_2(x,t)] \neq 0$. However, the function $\max g_2(x,t)$ has an inflection point at

$$x = \frac{1}{2\beta} \quad . \tag{60}$$

Concavity and convexity of the function is determined by the signs of the function and its second derivative.

The function

$$t = \frac{1}{\beta\lambda x}e^{\beta x} \quad , \text{ for } x \geq 0 \quad , \tag{61}$$

is minimum at $x = \beta^{-1}$, where $\frac{d^2t}{dx^2} > 0$. The function given by Eq. 61 has no inflection points.

The probability density function, $f_t(x)$, of the largest exceedance, for $x > 0$ and for the case of gamma distribution with $\alpha = 2$ for exceedances is depicted in Fig. 9.

The moment generating function of the largest exceedance, for this case, is

$$\psi_t(u) = e^{-\lambda t}+\beta^2 e^{-\lambda t}\sum_{k=1}^\infty \frac{(\lambda t)^k}{(k-1)!}\sum_{i=0}^{k-1}(-1)^i\binom{k-1}{i}\sum_{j=0}^{i}\binom{i}{j}\beta^j\frac{\Gamma(j+2)}{(\beta i+\beta-u)^{j+2}} \quad . \tag{62}$$

All absolute moments, are given by

$$E_X{}^m(t) = \frac{\lambda t}{\beta^m}e^{-\lambda t}\sum_{k=1}^\infty \frac{(\lambda t)^{k-1}}{(k-1)!}\sum_{i=0}^{k-1}(-1)^i\binom{k-1}{i}\sum_{j=0}^{i}\binom{i}{j}\frac{(j+m+1)!}{(i+1)^{j+m+2}} \quad , \tag{63}$$

with $m=1,2,\ldots,$ being the order of the absolute moment. Equation 63 is obtained by routine differentiation of the moment generating function with respect to u and evaluating derivatives at $u = 0$.

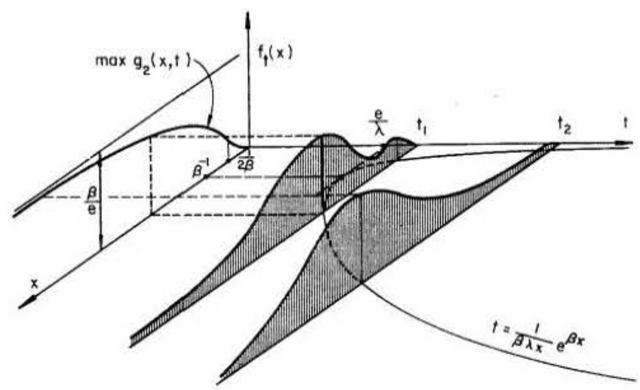

Fig. 9. Probability density function of the largest exceedance; given for $x > 0$, for different values of the time interval $[o,t]$, and for exceedances distributed according to Eq. 50.

Chapter 5

APPLICATIONS TO FLOOD EXCEEDANCES OF THE GREENBRIER RIVER

The first application of the presented method is made on the Greenbrier River at Alderson, W. Va., located in the Ohio River basin. Flood data, in the form of partial-duration series, cover the period of 1896 through 1967. The base flow for the partial-duration series is Q_b = 17,000 cfs. From these data, a series of 205 flood peak exceedances, in the course of 72 years, was obtained, or on the average about three flood peaks per year. The data are given in Table 18, Appendix 2.

5.1 Distribution of the Number of Flood Exceedances.

Seasonal occurrence of exceedances. For every watershed the major portion of the flood exceedances occur during a few specific months of the year. Seasonal occurrence of exceedances for the Greenbrier River at Alderson is shown graphically in Fig. 10 (data given in Table 19, Appendix 2).

To gain insight into the probabilistic structure of the seasonal occurrence of exceedances, the water year was divided into nine periods, eight periods of 40 days each, and one period of 45 days. Observed distributions of exceedances are obtained for all nine periods. The probability distribution governing the occurrence of exceedances is Poissonian, so it remains, at this point, to estimate the parameter λ for each period, using information from the available samples. After this, theoretical frequencies are computed and a goodness of fit test performed for each case. The results are given in Table 1.

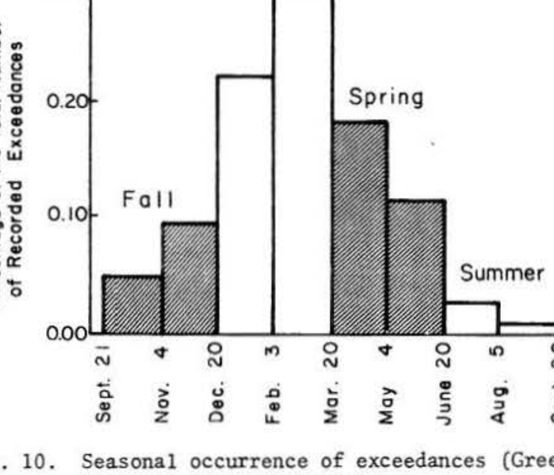

Fig. 10. Seasonal occurrence of exceedances (Greenbrier River at Alderson).

Table 1. Observed distributions and corresponding fitted Poissonian distributions of the number of exceedances for nine nonoverlapping periods of the water year (Greenbrier River at Alderson).

Oct. 1 - Nov. 9

k	$f_{ob.}$	$f_{th.}$
0	63	62.6544
1	8	8.7091
2	1	0.6044
3		0.0281
$\sum$	72	71.9960

$\hat{\lambda} = 0.00347 \frac{\text{exceed.}}{\text{day}}$

$\bar{k} = 0.139$; $\hat{\sigma}_k^2 = 0.149$

$\chi^2 = 0.347 < \chi^2_{cr}$

$\chi^2_{cr} = 5.99$

Nov. 10 - Dec. 19

k	$f_{ob.}$	$f_{th.}$
0	57	56.078
1	12	14.018
2	3	1.750
3		0.144
4		0.009
$\sum$	72	71.999

$\hat{\lambda} = 0.00625$

$\bar{k} = 0.250$; $\hat{\sigma}_k^2 = 0.274$

$\chi^2 = 1.356 < \chi^2_{cr}$

$\chi^2_{cr} = 7.81$

Dec. 20 - Jan. 28

k	$f_{ob.}$	$f_{th.}$
0	41	41.940
1	23	22.690
2	8	6.110
3		1.100
4		0.140
5		0.010
$\sum$	72	71.990

$\hat{\lambda} = 0.01351$

$\bar{k} = 0.541$; $\hat{\sigma}_k^2 = 0.480$

$\chi^2 = 1.863 < \chi^2_{cr}$

$\chi^2_{cr} = 9.49$

Jan. 29 - Mar. 9

k	$f_{ob.}$	$f_{th.}$
0	38	35.940
1	21	24.970
2	10	8.680
3	3	2.010
4		0.340
5		0.050
$\sum$	72	71.990

$\hat{\lambda} = 0.01737$

$\bar{k} = 0.695$; $\hat{\sigma}_k^2 = 0.749$

$\chi^2 = 1.839 < \chi^2_{cr}$

$\chi^2_{cr} = 9.49$

Mar. 10 - Apr. 18

k	$f_{ob.}$	$f_{th.}$
0	36	36.480
1	26	24.810
2	7	8.430
3	3	1.910
4		0.320
5		0.040
$\sum$	72	71.990

$\hat{\lambda} = 0.01700$

$\bar{k} = 0.680$; $\hat{\sigma}_k^2 = 0.672$

$\chi^2 = 1.297 < \chi^2_{cr}$

$\chi^2_{cr} = 9.49$

Apr. 19 - May 28

k	$f_{ob.}$	$f_{th.}$
0	52	52.340
1	17	16.690
2	3	2.660
3		0.283
4		0.022
5		0.001
$\sum$	72	71.996

$\hat{\lambda} = 0.00798$

$\bar{k} = 0.319$; $\hat{\sigma}_k^2 = 0.305$

$\chi^2 = 0.356 < \chi^2_{cr}$

$\chi^2_{cr} = 9.49$

May 29 - July 7

k	$f_{ob.}$	$f_{th.}$
0	64	62.6544
1	6	8.7091
2	2	0.6044
3		0.0281
4		0.0009
$\sum$	72	71.9969

$\hat{\lambda} = 0.00347$

$\bar{k} = 0.139$; $\hat{\sigma}_k^2 = 0.177$

$\chi^2 = 4.124 < \chi^2_{cr}$

$\chi^2_{cr} = 7.81$

July 8 - Aug. 16

k	$f_{ob.}$	$f_{th.}$
0	68	68.112
1	4	3.780
2		0.105
3		0.001
$\sum$	72	71.999

$\hat{\lambda} = 0.00139$

$\bar{k} = 0.055$; $\hat{\sigma}_k^2 = 0.053$

$\chi^2 = 0.119 < \chi^2_{cr}$

$\chi^2_{cr} = 5.99$

Aug. 17 - Sept. 30

k	$f_{ob.}$	$f_{th.}$
0	70	70.033
1	2	1.939
2		0.027
3		0.000
$\sum$	72	71.999

$\hat{\lambda} = 0.00069$

$\bar{k} = 0.028$; $\hat{\sigma}_k^2 = 0.027$

$\chi^2 = 0.028 < \chi^2_{cr}$

$\chi^2_{cr} = 3.84$

Legend:

k = the number of exceedances during the given period.

f_{ob} = observed absolute frequency.

f_{th} = theoretical absolute frequency.

$\hat{\lambda}$ = parameter of the Poisson distribution.

All χ^2_{cr}-values refer to the 5% level of significance.

The null hypothesis was tested for each of the nine periods to determine if the distribution of the number of exceedances in the given period is Poissonian with parameter λ estimated by the statistic $\hat{\lambda}$. As observed from Table 1, the computed χ^2-values are considerably below the critical χ^2-values. Therefore, the null hypothesis should be accepted at the 5 per cent significance level. This is valid for each case considered.

Since the mean equals the variance in a Poisson probability distribution it was also considered advantageous to compare the sample mean and the corresponding sample variance of the number of exceedances for each period These comparisons are given in Fig. 11.

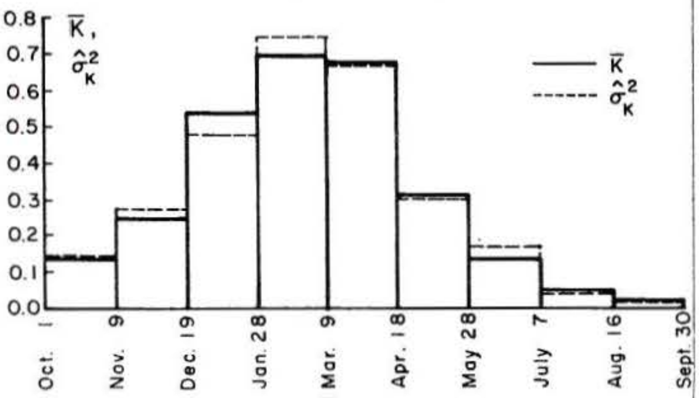

Fig. 11. Sample means, $\bar{K}$, and the corresponding sample variances, $\hat{\sigma}_K^2$, for the number of occurred exceedances during the nine periods of the year considered. (Greenbrier River at Alderson).

The parameter $\lambda(t)$ represents the average number of exceedances per day. The seasonal variations of this parameter are shown in Fig. 12.

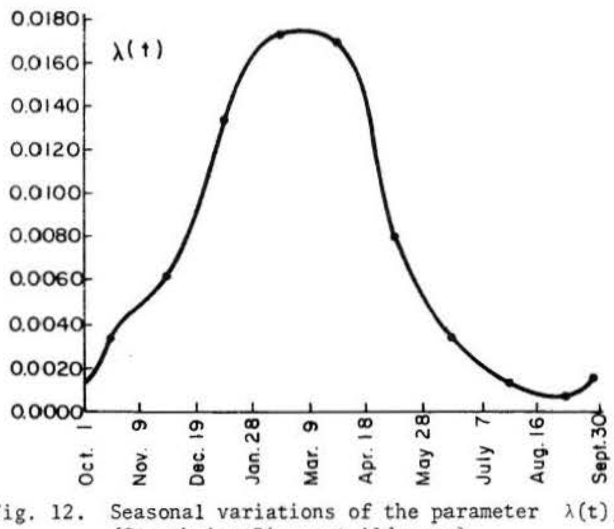

Fig. 12. Seasonal variations of the parameter $\lambda(t)$, (Greenbrier River at Alderson).

Distributions of the number of exceedances for different time intervals. The previous section considered the distributions of the number of exceedances during nine different periods of the year. The periods did not overlap and all were of the same size (only the last one differed slightly). The objective was to examine the seasonal effect on the number of exceedances and the underlying probability distribution of this number.

In this section a somewhat different problem is analyzed. Following the theory given in Chapter 3, the distributions of the number of exceedances for different time intervals are considered. According to Eq. 12, it is necessary to evaluate the function $\Lambda(t)$, which is the average number of exceedances in an interval. The time intervals used in this section have the same origin of October 1, are overlapping, and are of different durations. Empirical and fitted probability distributions of the number of exceedances during these time intervals were obtained in a fashion similar to the one used in the preceding section. The results are given in Table 2. Tests of conformance of the fitted Poissonian distribution to the observed distribution at the 5 percent significance level show good agreement, as can be seen from Fig. 13. Figure 13 shows the change in the distribution of the number of flood exceedances with incremental changes of the time interval for the Greenbrier River. The graph of the observed and fitted function $\Lambda(t)$ of Eq. 11 for the Greenbrier River is given in Fig. 14. The function $\Lambda(t)$ is the mean number of flood exceedances in a time interval $[o,t]$.

The fitted function, $\Lambda_f(t)$, has the expression

$$\Lambda_f(t) = 0.2475 + 0.1583t + 0.5086\cos\left(\frac{2\pi t}{18} + 0.6841\pi\right) + 0.0556\cos\left(\frac{2\pi t}{9} - 0.1476\pi\right) + 0.0154\cos\left(\frac{2\pi t}{6} + 0.7780\pi\right) + 0.0142\cos\left(\frac{2\pi t}{3} + 0.6742\pi\right) \tag{64}$$

and is obtained by using a Fourier-series fit. In Eq. 64, the interval is 20 days. The computations are given in Appendix 1, Table 16 and 17. Therefore, in the Poissonian distribution, $\lambda = \Lambda(t)$ is given by Eq. 64.

5.2 Distribution Magnitudes of Flood Exceedances. The next step in the analysis of exceedances was to investigate the distributions of magnitudes of exceedances occurring during different periods of the year. The periods used are the winter season, the spring season, the summer and fall seasons taken together, and the total period of a year. It was not feasible to analyze the distribution of the magnitude of exceedances occurring only during the summer season, because of the small number of exceedances that occurred during this time (see Table 19, Appendix 2). The observed distributions of the magnitude of exceedances, for the four periods, are given in Table 3 and shown in Fig. 15.

Considering the four sample frequency distributions shown in Fig. 15, the question arises whether they have the same population distribution. In other words, are the magnitudes of exceedances of the given four periods identically distributed? This problem is treated using the Kolmogorov-Smirnov test at the 5 percent significance level. The results show that the differences among the sample distributions are not statistically significant. As an illustration, for the most unfavorable case - the sample frequency distribution pertaining to the winter, and the summer and fall seasons - the result is

$$d_{n_1 n_2} = 0.1533 < 0.2593 = D_{n_1 n_2} ,$$

in which $d_{n_1 n_2}$ is the maximum observed deviation

Table 2. Observed distributions and corresponding fitted Poissonian distributions of exceedances for different time intervals (Greenbrier River at Alderson).

Oct.1-Oct.20 t=20 days

k	$f_{ob.}$	$f_{th.}$
0	68	68.1100
1	4	3.7800
2		0.1048
3		0.0019
$\sum$	72	71.9967

$\bar{k} = 0.055$; $\hat{\sigma}_k^2 = 0.053$
$\chi^2 = 0.120 < \chi^2_{cr}$
$\chi^2_{cr} = 5.99$

Oct.1-Nov.29 t=60 days

k	$f_{ob.}$	$f_{th.}$
0	59	59.260
1	12	11.510
2	1	1.120
3		0.070
$\sum$	72	71.960

$\bar{k} = 0.194$; $\hat{\sigma}_k^2 = 0.187$
$\chi^2 = 0.105 < \chi^2_{cr}$
$\chi^2_{cr} = 5.99$

Oct.1-Jan.8 t=100 days

k	$f_{ob.}$	$f_{th.}$
0	42	36.432
1	16	24.811
2	10	8.431
3	3	1.915
4	1	0.327
5		0.044
$\sum$	72	71.940

$\bar{k} = 0.68$; $\hat{\sigma}_k^2 = 0.92$
$\chi^2 = 6.316 < \chi^2_{cr}$
$\chi^2_{cr} = 9.49$

Oct.1-Feb.17 t=140 days

k	$f_{ob.}$	$f_{th.}$
0	25	21.744
1	21	25.992
2	16	15.552
3	7	6.228
4	3	1.848
5		0.441
6		0.088
$\sum$	72	71.893

$\bar{k} = 1.195$; $\hat{\sigma}_k^2 = 1.310$
$\chi^2 = 2.802 < \chi^2_{cr}$
$\chi^2_{cr} = 11.1$

Oct.1-Mar.9 t=160 days

k	$f_{ob.}$	$f_{th.}$
0	18	14.148
1	17	23.018
2	20	18.684
3	11	10.138
4	3	4.118
5	3	1.333
6		0.362
7		0.084
$\sum$	72	71.885

$\bar{k} = 1.63$; $\hat{\sigma}_k^2 = 1.81$
$\chi^2 = 5.622 < 12.6 = \chi^2_{cr}$

Oct.1-Mar.29 t=180 days

k	$f_{ob.}$	$f_{th.}$
0	13	9.072
1	14	18.792
2	18	19.440
3	16	13.464
4	5	6.948
5	5	2.880
6	1	0.986
7		0.292
$\sum$	72	71.874

$\bar{k} = 2.07$; $\hat{\sigma}_k^2 = 2.29$
$\chi^2 = 5.906 < 12.6 = \chi^2_{cr}$

Oct.1-Apr.18 t=200 days

k	$f_{ob.}$	$f_{th.}$
0	9	7.164
1	15	16.538
2	17	19.080
3	16	14.616
4	6	8.460
5	7	3.910
6	2	1.497
7		0.493
8		0.142
9		0.036
$\sum$	72	71.936

$\bar{k} = 2.31$; $\hat{\sigma}_k^2 = 2.56$
$\chi^2 = 4.967 < 15.5 = \chi^2_{cr}$

Oct.1-May 8 t=220 days

k	$f_{ob.}$	$f_{th.}$
0	8	6.048
1	16	14.962
2	15	18.504
3	14	15.192
4	9	9.374
5	7	4.641
6	2	1.927
7		0.673
8		0.209
$\sum$	72	71.530

$\bar{k} = 2.47$; $\hat{\sigma}_k^2 = 2.83$
$\chi^2 = 3.044 < 14.1 = \chi^2_{cr}$

Oct.1-May 28 t=240 days

k	$f_{ob.}$	$f_{th.}$
0	7	5.184
1	14	13.637
2	15	17.928
3	15	15.746
4	11	10.347
5	6	5.415
6	2	2.383
7	2	0.896
8		0.295
9		0.085
$\sum$	72	71.916

$\bar{k} = 2.63$; $\hat{\sigma}_k^2 = 2.90$
$\chi^2 = 3.066 < 15.5 = \chi^2_{cr}$

Oct.1-Sept.30 t=365 days

k	$f_{ob.}$	$f_{th.}$
0	5	4.154
1	13	11.851
2	15	16.898
3	16	15.984
4	10	11.419
5	8	6.480
6	2	3.078
7	2	1.297
8	1	0.447
9		0.140
10		0.040
$\sum$	72	71.738

$\bar{k} = 2.85$; $\hat{\sigma}_k^2 = 3.18$
$\chi^2 = 2.726 < 16.9 = \chi^2_{cr}$

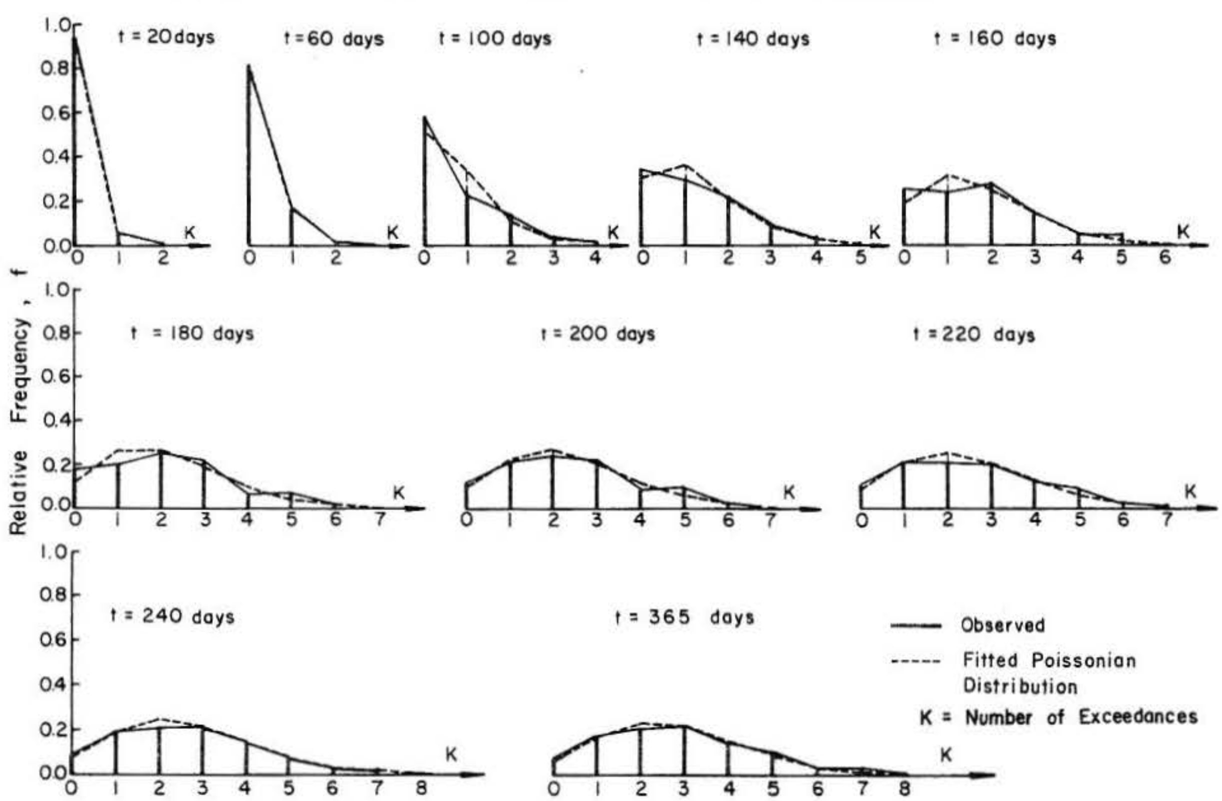

Fig. 13. The observed and corresponding fitted Poissonian distributions of the number of exceedances for intervals of 20, 60,100, 140, 160, 180, 200, 220, 240, and 365 days (Greenbrier River at Alderson).

between the two sample frequency distributions, and $D_{n_1 n_2}$ is the corresponding critical value obtained for the 5 percent level of significance.

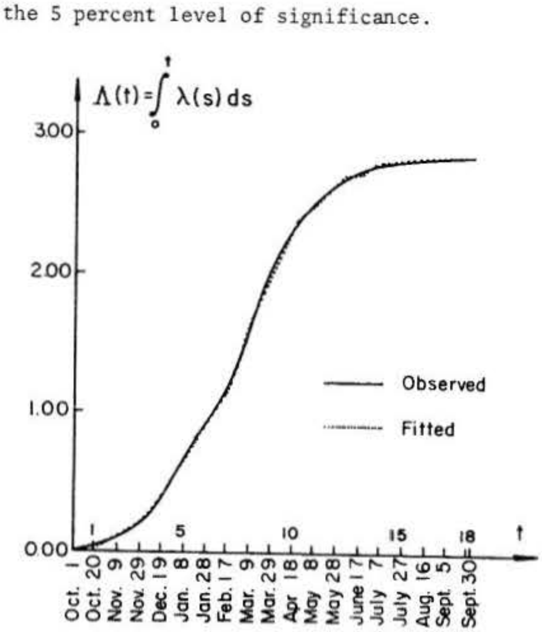

Fig. 14. Observed and fitted function $\Lambda(t)$ (Greenbrier River at Alderson).

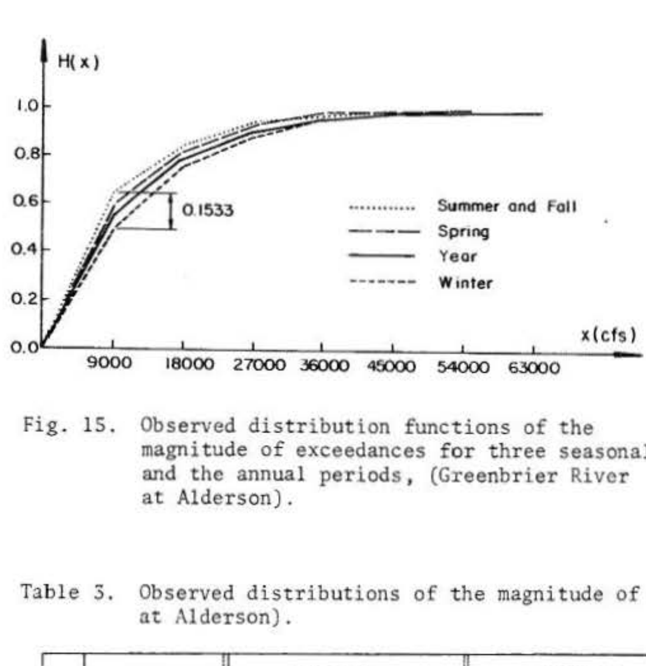

Fig. 15. Observed distribution functions of the magnitude of exceedances for three seasonal and the annual periods, (Greenbrier River at Alderson).

On the basis of the statistical tests performed, the hypothesis that the magnitudes of exceedances for the Greenbrier River at Alderson are identically distributed throughout the year is accepted. Estimation of the corresponding distribution function was made using the one-year sample. This sample had the greatest number of observations, 205. The observed frequency distribution and the corresponding fitted simple exponential distribution function are shown in Fig. 16.

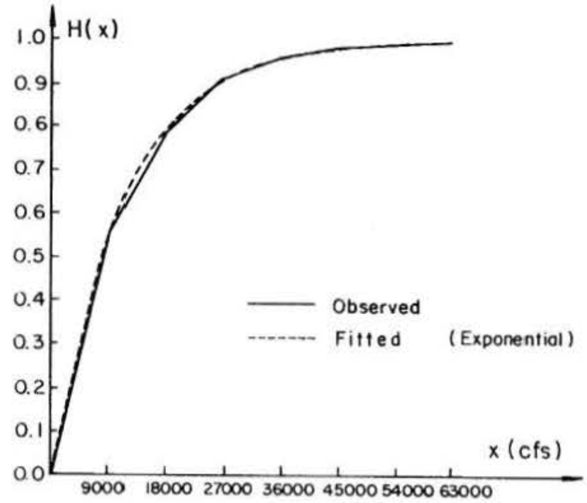

Fig. 16. The fitted simple exponential function and the observed frequency distribution of the magnitude of exceedances during one year (Greenbrier River at Alderson with data from Table 3).

It can be easily seen from Fig. 16 that the simple exponential distribution function gives a very good fit to the observed frequency distribution. Taking the observed frequency distribution of the magnitude of exceedances for the year as representative and estimating the parameter of the corresponding simple exponential distribution function from the data, the expression becomes

$$H(x) = 1 - \exp(-8.821 \cdot 10^{-5} x), \quad x \geq 0 , \qquad (65)$$

with x measured in cfs. Therefore, the parameter β is estimated by the statistic

$$\hat{\beta} = 8.821 \cdot 10^{-5} \ (\text{cfs})^{-1} \ . \qquad (66)$$

Table 3. Observed distributions of the magnitude of exceedances for four seasonal periods (Greenbrier River at Alderson).

		Winter			Spring			Summer & Fall			Year			
No.	x(cfs)	$f_{ob.}$	$f_{rel.}$	Obs. H(x)	$f_{ob.}$	$f_{rel.}$	Obs. H(x)	$f_{ob.}$	$f_{rel.}$	Obs. H(x)	$f_{ob.}$	$f_{rel.}$	Obs. H(x)	Fitted H(x)
1	1- 9000	53	0.4953	0.4953	36	0.5902	0.5902	24	0.6486	0.6486	113	0.55122	0.55122	0.5475
2	9001-18000	28	0.2617	0.7570	14	0.2295	0.8197	7	0.1893	0.8379	49	0.23902	0.79024	0.7957
3	18001-27000	14	0.1308	0.8878	7	0.1148	0.9345	4	0.1081	0.9460	25	0.12195	0.91219	0.9077
4	27001-36000	8	0.0748	0.9626	3	0.0491	0.9836	1	0.0270	0.9730	12	0.05854	0.97073	0.9583
5	36001-45000	3	0.0280	0.9906		0.0000	0.9836	1	0.0270	1.0000	4	0.01951	0.99024	0.9811
6	45001-54000		0.0000	0.9906	1	0.0164	1.0000				1	0.00488	0.99512	0.9914
7	54001-63000	1	0.0094	1.0000							1	0.00488	1.00000	0.9962
	Σ	107	1.0000		61	1.0000		37	1.0000		205	1.00000		

Both the Kolmogorov-Smirnov and the chi-square tests of goodness of fit have shown good agreement between the two distributions shown in Fig. 16.

5.3 Distribution of the Largest Magnitude of Flood Exceedances. The final part of the analysis of exceedances is the analysis of the distribution function, $F_t(x)$, of the largest magnitude of exceedances, called in the ensuing test simply "the largest exceedance". On the basis of the results obtained in previous chapters, the distribution function of the largest exceedance, for a time interval $[o,t]$, is the double exponential function

$$F_t(x) = \exp[-\Lambda(t)\exp(-8.821 \cdot 10^{-5} x)], \quad x \geq 0 \quad . \tag{67}$$

The function $\Lambda(t)$ is given by Eq. 64 and shown in Fig. 14.

Distribution of the largest exceedance using the one-year time interval. The time interval of one year holds the greatest appeal. The probability distribution function of the largest exceedance with a year as the time interval is then, for $\Lambda(t) = 2.85$,

$$F(x) = \exp[-2.85 \exp(-8.821 \cdot 10^{-5} x)] , \quad x \geq 0 \quad . \tag{68}$$

Values of this distribution function are given in Table 4, together with the values of the observed frequency distribution of the largest exceedance, for the same time interval. The graphs of the two distribution functions from Table 4 are given in Fig. 17.

Goodness of fit tests show close conformity between the fitted and the observed distributions of the largest exceedance for this example. According to the chi-square test, $\chi^2 = 4.6834 < 12.6 = \chi^2_{cr}$, and according to the Kolmogorov-Smirnov test, $d = 0.0387 < 0.1602 = D_{cr}$. The statistics χ^2 and d are obtained from Table 4; χ^2_{cr} and D_{cr} are the corresponding critical values for the 5 percent level of significance.

The probability density function of the largest exceedance using the one-year time interval is

$$f(x)=2e^{-2.85}\delta(x)+25.14\cdot10^{-5} \exp[-8.821\cdot10^{-5}x-2.85\exp(-8.821\cdot10^{-5}x)] \quad , \tag{69}$$

for $x \geq 0$. The mode of the largest exceedance is $\tilde{x} = 11.8730 \cdot 10^3$.

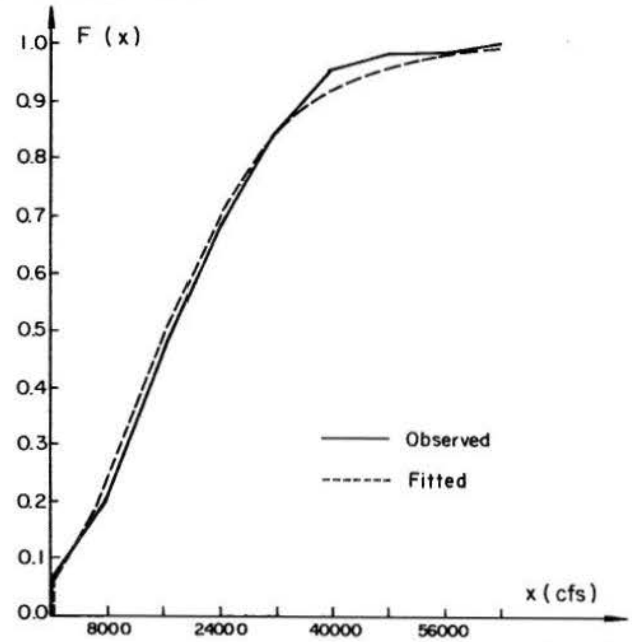

Fig. 17. The double exponential distribution function, and the observed frequency distribution of the largest exceedance, using a one-year time interval (Greenbrier River at Alderson).

The density $f(\tilde{x})$, and the distribution function $F(\tilde{x})$, evaluated at the mode, are $f(\tilde{x}) = 3.2451 \cdot 10^{-5}$ and $F(\tilde{x}) = 0.3679$. The abscissas of the inflection points are $x_1 = 0.9626 \cdot 10^3$ and $x_2 = 22.7836 \cdot 10^3$. The probability densities $f(x)$, evaluated at x_1 and x_2, are $f(x_1) = 1.6846 \cdot 10^{-5}$ and $f(x_2) = 2.2996 \cdot 10^{-5}$. The median of the largest exceedance is $\check{x} = 16.028 \cdot 10^3$. The density function $f(x)$ is given in Fig. 18, with $F^*(o) = 0.0694$.

Table 4. Fitted and observed distributions of the largest exceedance with one year as the time interval (Greenbrier River at Alderson).

No.	x (cfs)	Fitted F(x)	Observed F*(x)	|F(x) - F*(x)|	Absolute Frequency Fitted f	Absolute Frequency Observed f*	$\frac{(f-f^*)^2}{f}$
0	0	0.0577	0.0694	0.0117	4.15	5	0.1738
1	8000	0.2450	0.2083	0.0367	13.48	10	0.8980
2	16000	0.5000	0.4722	0.0278	18.36	19	0.0223
3	24000	0.7092	0.6805	0.0287	15.06	15	0.0002
4	32000	0.8442	0.8472	0.0030	9.73	12	0.5290
5	40000	0.9196	0.9583	→ 0.0387	5.42	8	1.2280
6	48000	0.9595	0.9861	0.0266	2.88	2	0.2670
7	56000	0.9798	0.9861	0.0063	1.46	0	1.4610
8	64000	0.9899	1.0000	0.0101	0.73	1	0.1041
	$\sum$					72	4.6834

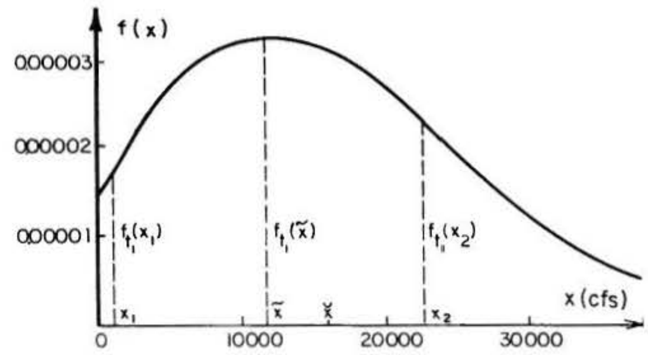

Fig. 18. The probability density function of the largest exceedance using the one-year time interval (Greenbrier River at Alderson).

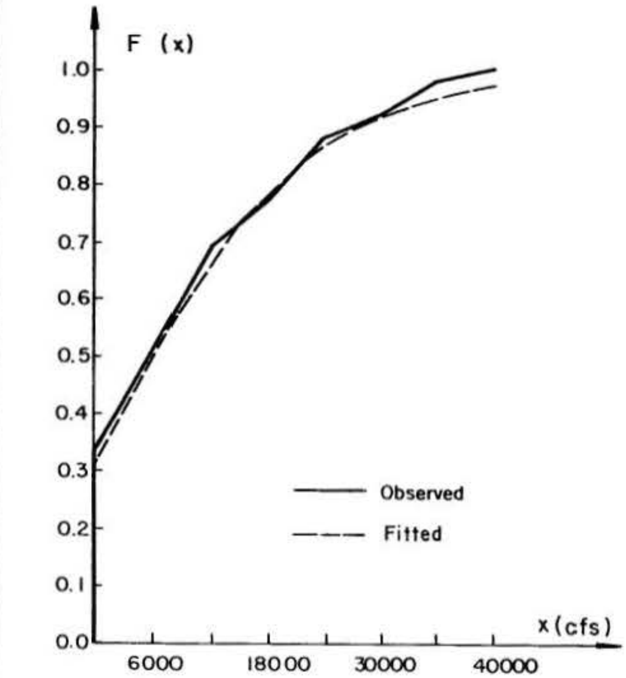

Fig. 19. Fitted and observed distributions of the largest exceedance in the 140-day interval (Greenbrier River at Alderson).

Distributions of the largest exceedance in intervals of 140 and 180 days. The distribution function of the largest exceedance in the 140-day interval counting from October 1, is

$$F(x) = \exp\,[-\,1.195\,\exp(-8.821 \cdot 10^{-5})]\ ,\ x \geq 0. \quad (70)$$

Table 5 gives the values of this function as well as the values of the corresponding observed frequency distribution. Figure 19 graphically shows the data of Table 5. Goodness of fit tests have also shown good agreement between the fitted and observed distributions of the largest exceedance in a 140-day interval. According to the chi-square test, $\chi^2 = 3.9558 < 11.1 = \chi^2_{cr}$, and according to the Kolmogorov-Smirnov test $d = 0.045 < 0.160 = D_{cr}$, in which the statistics χ^2 and d are obtained from Table 5, and χ^2_{cr} and D_{cr} are the corresponding critical values for the 5 percent level of significance.

The fitted distribution function of the largest exceedance in the 180-day interval, counting from October 1, is

$$F(x) = \exp[-2.07\,\exp(-8.821 \cdot 10^{-5}\,x)]\ ,\quad x \geq 0\ . \quad (71)$$

Both the fitted function and the corresponding observed frequency distribution of the largest exceedance in the 180-day interval are given in Table 6. The distributions of Table 6 are shown graphically in Fig. 20. Again, tests of goodness of fit show good agreement between the fitted distribution function of the largest exceedance and the observed frequency distribution in the 180-day interval. According to the chi-square test, $\chi^2 = 8.6327 < 12.6 = \chi^2_{cr}$, and according to the Kolmogorov-Smirnov test $d = 0.0568 < 0.1602 = D_{cr}$. The statistics χ^2 and d are computed in Table 6, and χ^2_{cr} and D_{cr} are the corresponding critical values for the 5 percent level of significance. The distribution functions of all three intervals, one-year, 140-days, and 180-days are in Fig. 21 to accentuate the effect of the time interval on the fitted probability distributions and observed frequency distributions of the largest exceedance.

Table 5. Fitted and observed distributions of the largest exceedance in the 140-day interval (Greenbrier River at Alderson).

No.	x (cfs)	Fitted F(x)	Observed F*(x)	|F(x)-F*(x)|	Absolute Frequency Fitted f	Absolute Frequency Observed f*	$\frac{(f-f^*)^2}{f}$
0	0	0.30200	0.34700	→ 0.04500	21.750	25	0.48500
1	6000	0.49505	0.51388	0.01883	13.890	12	0.25700
2	12000	0.66050	0.69443	0.03393	11.913	13	0.09919
3	18000	0.78302	0.77777	0.00525	8.821	6	0.90218
4	24000	0.86580	0.88888	0.02308	5.960	8	0.69826
5	30000	0.91912	0.93054	0.01142	3.839	3	0.18336
6	36000	0.95138	0.98610	0.03472	2.323	4	1.21063
7	42000	0.97100	1.00000	0.02900	1.412	1	0.12026
	Σ					72	3.95588

Table 6. Fitted and observed distributions of the largest exceedance in the 180-day interval (Greenbrier River at Alderson).

No.	x (cfs)	Fitted F(x)	Observed F*(x)	|F(x)-F*(x)|	Absolute Frequency Fitted f	Absolute Frequency Observed f*	$\frac{(f - f^*)^2}{f}$
0	0	0.12600	0.18056	0.05456	9.075	13	1.7200
1	8000	0.35971	0.31944	0.04027	16.815	10	2.7500
2	16000	0.60419	0.56944	0.03475	17.610	18	0.0086
3	24000	0.77899	0.72222	→ 0.05677	12.590	11	0.2010
4	32000	0.88417	0.87500	0.00917	7.580	11	1.5450
5	40000	0.94091	0.95833	0.01742	4.080	6	0.9040
6	48000	0.97040	0.98611	0.01571	2.122	2	0.0071
7	56000	0.98532	0.98611	0.00079	1.075	0	1.0750
8	64000	0.99266	1.00000	0.00734	0.528	1	0.4220
	∑					72	8.6327

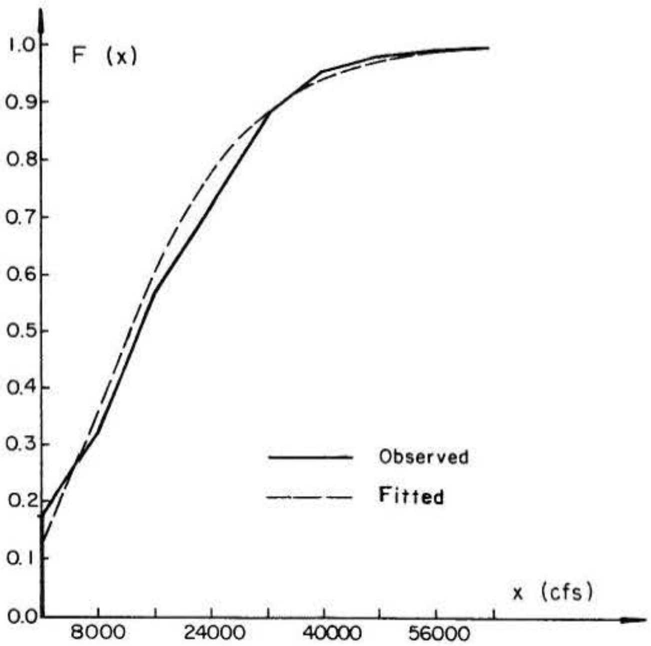

Fig. 20. Fitted and observed distributions of the largest exceedance in the 180-day interval (Greenbrier River at Alderson).

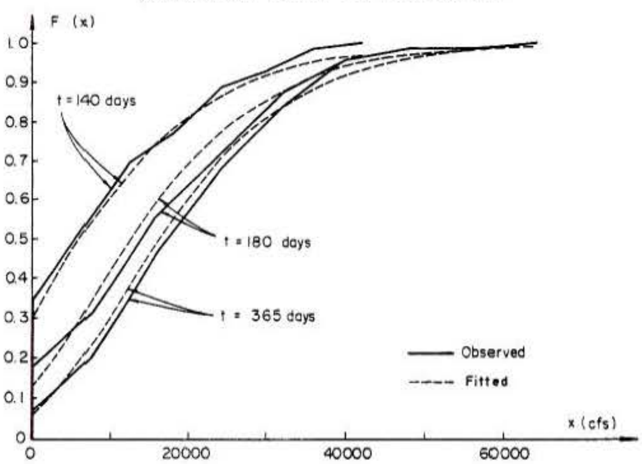

Fig. 21. The fitted and observed distributions of the largest exceedance in 140-, 180-, and 365-day intervals (Greenbrier River at Alderson).

5.4 Comparison of the Results of the Method Developed in this Study with the Method Used By Gumbel. The mean and standard deviation of the maximum annual discharges of the Greenbrier River at Alderson, for the 72 years considered, are $\overline{Q}_a$ = 35412.5 cfs and $\sigma(Q_a)$ = 13351.4 cfs. Therefore, the double exponential distribution function for this case, using Gumbel's estimate, is

$$F(x) = \exp[- 16.86 \exp(- 9.6 \cdot 10^{-5} x)] ,$$

$$x \geq 0 .$$

This function and the corresponding observed frequency distribution of maximum annual peak discharges are given in Table 7 and shown in Fig. 22.

For the Gumbel's estimates of the two parameters the chi-square is χ^2 = 8.34868, which is much larger, for the same number of class intervals, than the chi-square obtained by the estimates of the method used in this study, with χ^2 = 4.0036.

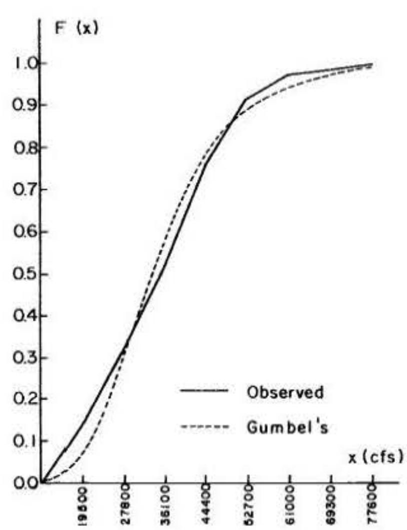

Fig. 22. Double exponential distribution function, with parameters estimated by Gumbel's method, and the corresponding observed frequency distribution of the maximum annual discharge (Greenbrier River at Alderson).

Accordingly, for this example, the method of estimates of parameters in the double exponential distribution function, given in this study, achieved better conformance to the observed distribution than Gumbel's method of estimates.

Table 7. Fitted double exponential distribution function using Gumbel's estimates, and the observed frequency distribution of the maximum annual discharge (Greenbrier River at Alderson).

No.	x (cfs)	Gumbel's F(x)	Observed F*(x)	|F(x)-F*(x)|	Absolute Frequency Gumbel's f	Absolute Frequency Observed f*	$\frac{(f-f^*)^2}{f}$
1	19500	0.0746	0.1389	→ 0.0643	5.371	10	3.98950
2	27800	0.3105	0.3195	0.0090	16.985	13	0.93495
3	36100	0.5900	0.5278	0.0622	20.124	15	1.30468
4	44400	0.7887	0.7639	0.0248	14.306	17	0.50731
5	52700	0.8988	0.9167	0.0179	7.927	11	1.19128
6	61000	0.9529	0.9722	0.0193	3.895	4	0.00282
7	69300	0.9785	0.9861	0.0076	1.843	1	0.38557
8	77600	0.9901	1.0000	0.0099	0.835	1	0.03257
	Σ					72	8.34868

Chapter 6

APPLICATION OF FLOOD EXCEEDANCES TO THE SUSQUEHANNA RIVER

In this chapter, the developed method is applied to the Susquehanna River at Wilkes-Barre, Pa. This river is part of the North Atlantic Slope Basin. Flood data, again in the form of partial-duration series, cover the years 1891 through 1964. Two years, 1898 and 1899, are omitted because of nonhomogeneity of data. The base for the partial-duration series is Q_b = 82,000 cfs. From these data, a series of 136 exceedances, in the course of 72 years, is obtained. The data are given in Table 23, Appendix 2.

6.1 Distribution of the Number of Flood Exceedances.

Seasonal occurrence of exceedances. The seasonal occurrence of exceedances for the Susquehanna River at Wilkes-Barre is shown in Fig. 23 (data given in Table 22, Appendix 2).

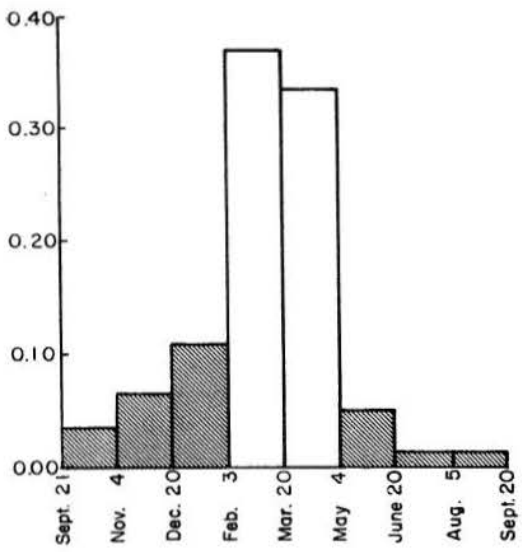

Fig. 23. Seasonal distribution of the number of exceedances (Susquehanna River at Wilkes-Barre).

The water year (Sept. 21-Sept. 20) was divided into the same observation periods as in the previous example using the Greenbrier River. After the estimation of parameter λ for each period, the Poissonian distributions of the number of exceedances were computed, and tests of goodness of fit performed. The results obtained are given in Table 8. For each of the nine considered periods, the null hypothesis that the distribution of the number of exceedances in the given period was Poissonian with parameter λ, estimated by the statistic $\hat{\lambda}$, was tested. As observed from the table, all computed χ^2-values are considerably below the critical χ^2-values. Therefore, the null hypothesis should be accepted at the 5 percent significance level.

The seasonal variation of the parameter λ is shown in Fig. 24.

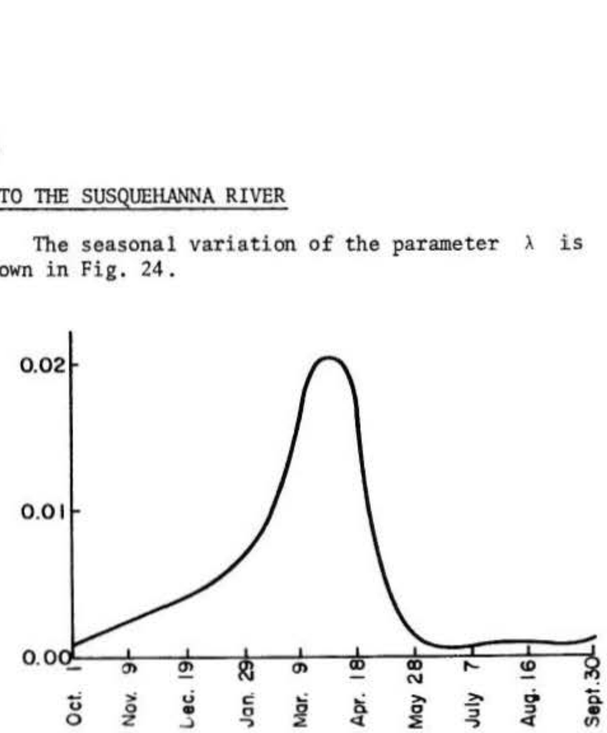

Fig. 24. Seasonal distribution of the parameter $\lambda(t)$, the average number of exceedances per day, (Susquehanna River at Wilkes-Barre).

Distribution of the number of exceedances for varying time intervals. Observed relative frequency distributions of the number of exceedances are determined for time intervals of 20, 60, 100, 140, 160, 180, 200, 220, 240 and 365 days, starting with October 1. Parameters of the corresponding fitted distribution functions are estimated from the available samples. The conformance of the fitted Poissonian distribution functions to the observed frequency distributions is verified by applying the χ^2-test to each case at the 5 percent level of significance. The agreement is very good. The results are given in Table 9 and depicted in Fig. 25.

The graph of the function $\Lambda(t)$ for the Susquehanna River at Wilkes-Barre is given in Fig. 26 The fitted function, $\Lambda_f(t)$, is

$$\Lambda_f(t) = 0.1015 + 0.1050t + 0.3936\cos\left(\frac{\pi t}{9} + 0.6032\pi\right)$$

$$+ 0.1280\cos\left(\frac{2\pi t}{9} - 0.4074\pi\right) + 0.0604\cos\left(\frac{\pi t}{3} + 0.5892\pi\right)$$

$$+ 0.0130\cos\left(\frac{2\pi t}{3} - 0.2041\pi\right) \quad , \qquad (72)$$

and is also shown in Fig. 26.

In Eq. 72, the unit of t refers to a 20-day period. Therefore, the parameter λ in the Poissonian distribution function is given by Eq. 72.

Table 8. Observed and corresponding fitted Poissonian distributions of the number of exceedances for nine non-overlapping periods of the water year (Susquehanna River at Wilkes-Barre).

Oct. 1 - Nov. 9

k	$f_{ob.}$	$f_{th.}$
0	67	67.1000
1	5	4.6500
2		0.1617
3		0.0037
Σ	72	71.9204

$\hat{\lambda} = 0.001735$

$\bar{k} = 0.0694$; $\hat{\sigma}_k^2 = 0.0655$

$\chi^2 = 0.1922 < 5.99 = \chi^2_{cr}$

Nov. 10 - Dec. 19

k	$f_{ob.}$	$f_{th.}$
0	64	63.5000
1	7	7.9400
2	1	0.4950
3		0.0207
4		0.0006
Σ	72	71.9563

$\hat{\lambda} = 0.00312$

$\bar{k} = 0.125$; $\hat{\sigma}_k^2 = 0.139$

$\chi^2 = 0.6517 < 7.81 = \chi^2_{cr}$

Dec. 20 - Jan. 28

k	$f_{ob.}$	$f_{th.}$
0	57	58.4700
1	15	12.1700
2		1.2630
3		0.0877
4		0.0046
Σ	72	71.9953

$\hat{\lambda} = 0.00520$

$\bar{k} = 0.208$; $\hat{\sigma}_k^2 = 0.168$

$\chi^2 = 2.053 < 7.81 = \chi^2_{cr}$

Jan. 29 - Mar. 9

k	$f_{ob.}$	$f_{th.}$
0	48	48.060
1	19	19.380
2	5	3.898
3		0.521
4		0.052
Σ	72	71.911

$\hat{\lambda} = 0.01006$

$\bar{k} = 0.402$; $\hat{\sigma}_k^2 = 0.385$

$\chi^2 = 0.8941 < 7.81 = \chi^2_{cr}$

Mar. 10 - Apr. 18

k	$f_{ob.}$	$f_{th.}$
0	28	31.670
1	32	26.000
2	9	10.640
3	3	2.910
4		0.597
5		0.098
Σ	72	71.915

$\hat{\lambda} = 0.0205$

$\bar{k} = 0.820$; $\hat{\sigma}_k^2 = 0.657$

$\chi^2 = 2.7608 < 9.49 = \chi^2_{cr}$

Apr. 19 - May 28

k	$f_{ob.}$	$f_{th.}$
0	58	59.2500
1	14	11.5000
2		1.1200
3		0.0720
4		0.0036
Σ	72	71.9456

$\hat{\lambda} = 0.00485$

$\bar{k} = 0.194$; $\hat{\sigma}_k^2 = 0.159$

$\chi^2 = 1.7682 < 7.81 = \chi^2_{cr}$

May 29 - July 7

k	$f_{ob.}$	$f_{th.}$
0	71	71.0054
1	1	0.9870
2		0.0068
3		0.0000
Σ	72	71.9992

$\hat{\lambda} = 0.000347$

$\bar{k} = 0.0139$; $\hat{\sigma}_k^2 = 0.0137$

$\chi^2 = 0.0070 < 5.99 = \chi^2_{cr}$

July 8 - Aug. 16

k	$f_{ob.}$	$f_{th.}$
0	70	70.026
1	2	1.946
2		0.027
3		0.000
Σ	72	71.999

$\hat{\lambda} = 0.000695$

$\bar{k} = 0.0278$; $\hat{\sigma}_k^2 = 0.0270$

$\chi^2 = 0.0285 < 5.99 = \chi^2_{cr}$

Aug. 17 - Sept. 30

k	$f_{ob.}$	$f_{th.}$
0	70	70.026
1	2	1.946
2		0.027
3		0.000
Σ	72	71.999

$\hat{\lambda} = 0.000618$

$\bar{k} = 0.0278$; $\hat{\sigma}_k^2 = 0.0270$

$\chi^2 = 0.0285 < 5.99 = \chi^2_{cr}$

Legend:

k = the number of exceedances during the given period.

f_{ob} = observed absolute frequency.

f_{th} = theoretical absolute frequency.

λ = parameter of the Poisson distribution.

All χ^2_{cr}- values refer to the 5% level of significance.

Table 9. Observed and corresponding fitted Poissonian distributions of the number of exceedances for different time intervals (Susquehanna River at Wilkes-Barre).

Oct.1-Oct.20 t=20 days

k	$f_{ob.}$	$f_{th.}$
0	67	67.1200
1	5	4.6620
2		0.1617
Σ	72	71.9437

$\bar{k} = 0.0694$; $\hat{\sigma}_k^2 = 0.0655$

$\chi^2 = 0.0409 < 3.84 = \chi^2_{cr}$

Oct.1-Nov.29 t=60 days

k	$f_{ob.}$	$f_{th.}$
0	62	62.6400
1	10	8.7100
2		0.6055
3		0.0280
Σ	72	71.9835

$\bar{k} = 0.1390$; $\hat{\sigma}_k^2 = 0.1196$

$\chi^2 = 0.8326 < 5.99 = \chi^2_{cr}$

Oct.1-Jan.8 t=100 days

k	$f_{ob.}$	$f_{th.}$
0	55	55.280
1	15	14.600
2	2	1.927
3		0.170
4		0.011
Σ	72	71.988

$\bar{k} = 0.264$; $\hat{\sigma}_k^2 = 0.253$

$\chi^2 = 0.1943 < 7.81 = \chi^2_{cr}$

Oct.1-Feb.17 t=140 days

k	$f_{ob.}$	$f_{th.}$
0	46	43.63
1	18	21.81
2	6	5.45
3	2	0.91
4		0.11
5		0.01
Σ	72	71.92

$\bar{k} = 0.500$; $\hat{\sigma}_k^2 = 0.591$

$\chi^2 = 2.2764 < 9.49 = \chi^2_{cr}$

Oct.1-Mar.9 t=160 days

k	$f_{ob.}$	$f_{th.}$
0	37	32.110
1	18	25.870
2	12	10.440
3	4	2.800
4	1	0.565
5		0.091
Σ	72	71.876

$\bar{k} = 0.806$; $\hat{\sigma}_k^2 = 1.002$

$\chi^2 = 4.3096 < 9.49 = \chi^2_{cr}$

Oct.1-Mar.29 t=180 days

k	$f_{ob.}$	$f_{th.}$
0	24	20.304
1	20	25.684
2	19	16.243
3	6	6.847
4	1	2.160
5	1	0.547
6	1	0.115
7		0.021
Σ	72	71.921

$\bar{k} = 1.265$; $\hat{\sigma}_k^2 = 1.570$

$\chi^2 = 10.3304 < 12.6 = \chi^2_{cr}$

Oct.1-Apr.18 t=200 days

k	$f_{ob.}$	$f_{th.}$
0	12	14.170
1	26	23.000
2	19	18.700
3	11	10.150
4	1	4.130
5	2	1.340
6	1	0.364
7		0.084
8		0.017
Σ	72	71.955

$\bar{k} = 1.625$; $\hat{\sigma}_k^2 = 1.590$

$\chi^2 = 4.7087 < 14.1 = \chi^2_{cr}$

Oct.1-May 8 t=220 days

k	$f_{ob.}$	$f_{th.}$
0	11	12.5060
1	23	21.8880
2	20	19.1520
3	14	11.2000
4	0	4.9000
5	3	0.9830
6	1	0.4990
7		0.1240
8		0.0158
Σ	72	71.2678

$\bar{k} = 1.750$; $\hat{\sigma}_k^2 = 1.684$

$\chi^2 = 10.6601 < 14.1 = \chi^2_{cr}$

Oct.1-May 28 t=240 days

k	$f_{ob.}$	$f_{th.}$
0	10	11.670
1	23	21.200
2	18	19.300
3	17	11.700
4	0	5.310
5	3	1.937
6	1	0.585
7		0.152
8		0.035
Σ	72	71.889

$\bar{k} = 1.819$; $\hat{\sigma}_k^2 = 1.700$

$\chi^2 = 9.2544 < 14.1 = \chi^2_{cr}$

Oct.1-Sept.30 t=365 days

k	$f_{ob.}$	$f_{th.}$
0	9	10.872
1	22	20.527
2	20	19.425
3	16	12.182
4	1	5.760
5	3	2.174
6		0.681
7	1	0.186
8		0.043
Σ	72	71.850

$\bar{k} = 1.889$; $\hat{\sigma}_k^2 = 1.817$

$\chi^2 = 10.1767 < 14.1 = \chi^2_{cr}$

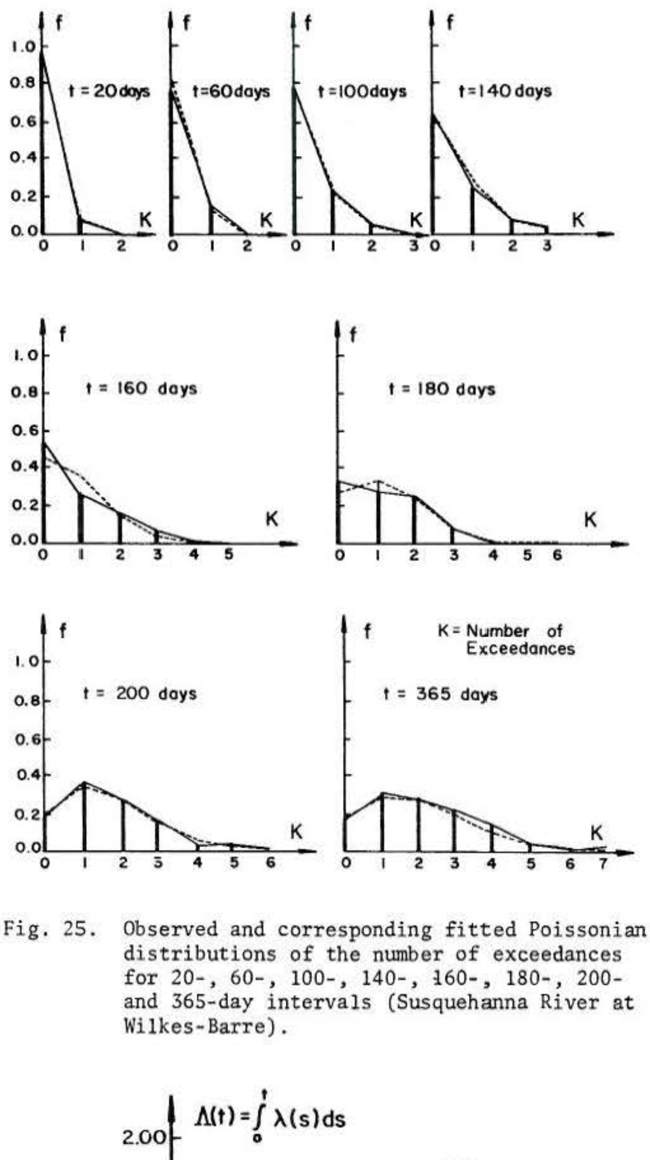

Fig. 25. Observed and corresponding fitted Poissonian distributions of the number of exceedances for 20-, 60-, 100-, 140-, 160-, 180-, 200- and 365-day intervals (Susquehanna River at Wilkes-Barre).

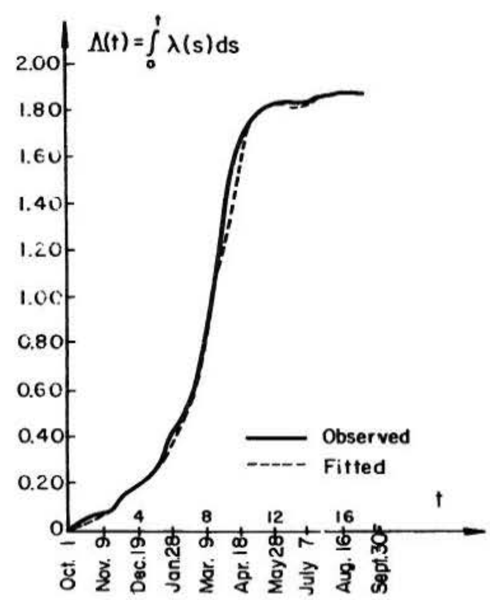

Fig. 26. Observed curve and fitted function $\Lambda(t)$ as the mean number of exceedances in a time interval $[o,t]$ (Susquehanna River at Wilkes-Barre).

6.2 <u>Distribution of magnitudes of flood Exceedances</u>. Most exceedances for the Susquehanna River at Wilkes-Barre occur during the winter and spring seasons (see Table 22, Appendix 2). In 72 years only 14 exceedances occurred during the fall season, and four during the summer season. Therefore, it was feasible to obtain the sample distribution function of exceedances only for the winter season, the spring season, and the year. The observed distributions of the exceedances for these three periods are given in Table 10 and shown in Fig. 27. The statistical hypotheses were tested, applying the Kolmogorov-Smirnov test at the 5 percent level of significance, to see whether the sample distribution functions given in Table 10 had the same population distribution function. The tests performed indicated that the differences between the sample frequency distributions are not statistically significant. Therefore, the null hypothesis that the magnitude of exceedances for the Susquehanna River at Wilkes-Barre is identically distributed through the year should be accepted at the 5 percent level of significance. Estimation of the parameters of the corresponding distribution function was made for the one-year interval by using the sample data. That sample had the greatest number of observations, 136. The sample and the corresponding fitted distribution function are shown in Fig. 28.

The fitted simple exponential distribution function of magnitude of exceedances has the expression

$$H(x) = 1 - \exp(-2.628 \cdot 10^{-5}x) \ , \ x \geq 0, \qquad (73)$$

in which x is measured in cfs. Both the Kolmogorov-Smirnov and the chi-square tests of goodness of fit show good agreement between the two distribution functions shown in Fig. 28.

6.3. <u>Distribution of the Largest Flood Exceedance.</u> The distribution function of the largest exceedance, for an interval of time $[o,t]$, for the Susquehanna River at Wilkes-Barre is

$$F_t(x) = \exp[-\Lambda(t) \exp(-2.628 \cdot 10^{-5}x)], \quad x \geq 0 \ . \qquad (74)$$

The function $\Lambda(t)$ is given by Eq. 72 and shown in Fig. 26.

<u>Distribution of the largest exceedance during the one-year interval</u>. The distribution function of the largest exceedance for the time interval of a year is

$$F(x) = \exp[-1.889 \exp(-2.628 \cdot 10^{-5} x)], \ x \geq 0. \qquad (75)$$

Values of this distribution function are given in Table 11, together with the values of the observed frequency distribution of the magnitude of the largest exceedance for the same time interval. The two distributions of Table 11 are shown in Fig. 29. Goodness of fit tests indicate good agreement between the fitted and the observed distribution for this river. According to the chi-square test, $\chi^2 = 4.206 < 12.6 = \chi^2_{cr}$, and according to the Kolmogorov-Smirnov test $d = 0.0448 < 0.1602 = D_{cr}$. The statistics χ^2 and d are obtained from Table 11, and χ^2_{cr} and D_{cr} are the corresponding critical values for the 5 percent level of significance.

Table 10. Observed winter, spring, and yearly distributions of magnitude of exceedances (Susquehanna River at Wilkes-Barre).

No.	x(cfs)	Winter $f_{ob.}$	Winter $f_{rel.}$	Winter Observed H(x)	Spring $f_{ob.}$	Spring $f_{rel.}$	Spring Observed H(x)	Year $f_{ob.}$	Year $f_{rel.}$	Year Observed H(x)	Year Fitted H(x)
1	1- 15000	29	0.4460	0.4460	12	0.2265	0.2265	45	0.3309	0.3309	0.3270
2	15001- 30000	11	0.1692	0.6152	13	0.2453	0.4718	28	0.2059	0.5368	0.5460
3	30001- 45000	8	0.1231	0.7383	9	0.1698	0.6416	23	0.1691	0.7059	0.6945
4	45001- 60000	3	0.0462	0.7845	7	0.1321	0.7737	11	0.0809	0.7868	0.7940
5	60001- 75000	4	0.0615	0.8460	2	0.0377	0.8114	7	0.0515	0.8383	0.8611
6	75001- 90000	2	0.0308	0.8768	4	0.0755	0.8869	8	0.0588	0.8971	0.9063
7	90001-105000	2	0.0308	0.9076	2	0.0377	0.9246	4	0.0294	0.9265	0.9370
8	105001-120000	2	0.0308	0.9384	2	0.0377	0.9623	4	0.0294	0.9559	0.9575
9	120001-135000	2	0.0308	0.9692	2	0.0377	1.0000	4	0.0294	0.9853	0.9705
10	135001-150000	2	0.0308	1.0000				2	0.0147	1.0000	0.9807
	$\sum$	65			53			136			

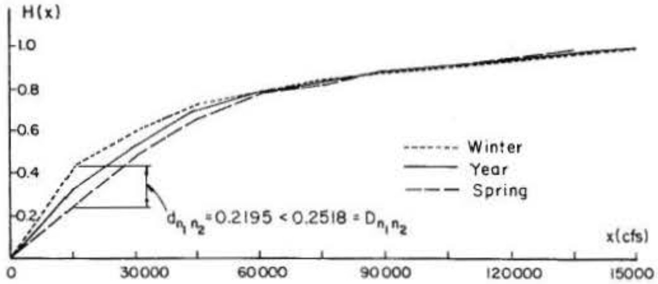

Fig. 27. Observed winter, spring, and yearly distribution functions of magnitudes of exceedances (Susquehanna River at Wilkes-Barre).

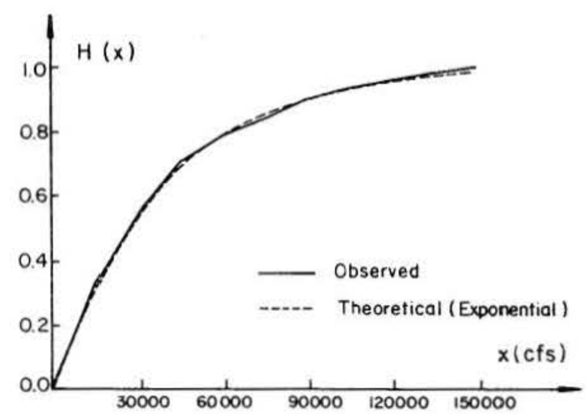

Fig. 28. The fitted function and the observed frequency distribution of the magnitude of exceedances for one year intervals (Susquehanna River at Wilkes-Barre, data from Table 10).

Table 11. Fitted and observed distributions of the largest exceedance using the one-year time interval (Susquehanna River at Wilkes-Barre).

No.	x (cfs)	Fitted F(x)	Observed F*(x)	|F(x) - F*(x)|	Absolute Frequency Fitted f	Absolute Frequency Observed f*	$\frac{(f - f^*)^2}{f}$
0	0	0.1510	0.1250	0.0260	10.880	9	0.3250
1	20000	0.3275	0.3055	0.0220	12.700	13	0.0071
2	40000	0.5170	0.4722	→ 0.0448	13.644	12	0.1981
3	60000	0.6766	0.6805	0.0039	11.491	15	1.0700
4	80000	0.7942	0.7777	0.0165	8.467	7	0.2542
5	100000	0.8726	0.8472	0.0254	5.645	5	0.0737
6	120000	0.9227	0.9167	0.0060	3.607	5	0.5380
7	140000	0.9535	0.9722	0.0187	2.218	4	1.4317
8	160000	0.9723	1.0000	0.0277	1.354	2	0.3082
	$\sum$					72	4.2060

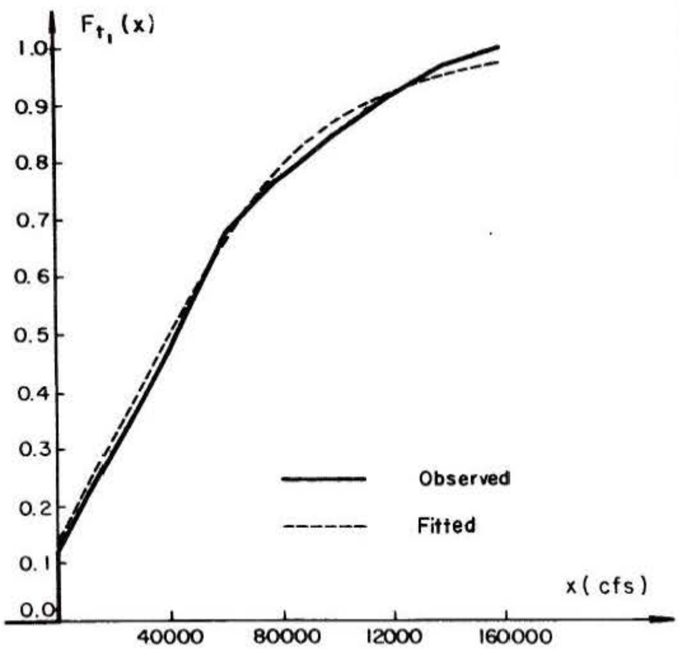

Fig. 29. Fitted and observed distributions of the largest exceedance using the one-year time interval (Susquehanna River at Wilkes-Barre).

Distribution of the largest exceedance for 160- and 200-day intervals. The distribution function of the largest exceedance for the 160-day interval is given by the expression

$$F(x) = \exp[-0.806 \exp(-2.628 \cdot 10^{-5} x)], \quad x \geq 0 \quad . \qquad (76)$$

The 160-day interval was measured from October 1. Table 12 gives the values of the fitted function $F(x)$, as well as the values of the corresponding observed frequency distribution $F^*(x)$. The graphic presentation of the data is given in Fig. 30.

Goodness of fit tests verified good agreement between the fitted and observed distributions for the 160-day interval. According to the chi-square test, $\chi^2 = 5.7785 < 11.1 = \chi^2_{cr}$, and according to the Kolmogorov-Smirnov test, $d = 0.087 < 0.1602 = D_{cr}$, in which the statistics χ^2 and d are obtained from Table 12, and χ^2_{cr} and D_{cr} are the corresponding critical values for the 5 percent level of significance.

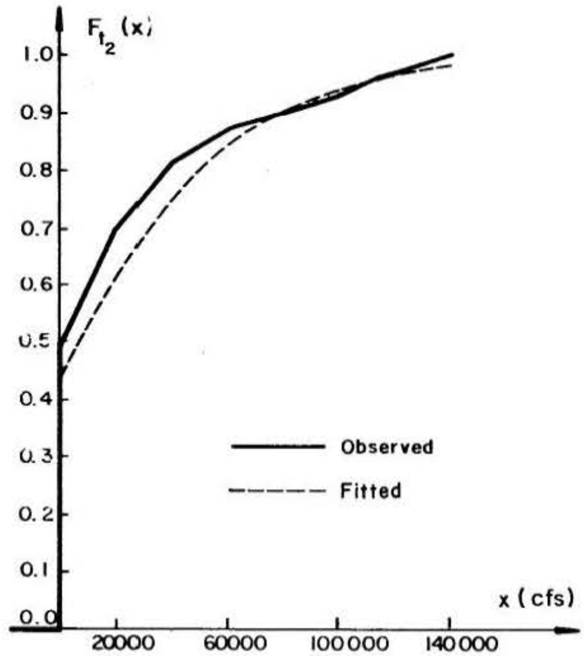

Fig. 30. Fitted and observed distributions of the largest exceedance for the 160-day interval (Susquehanna River at Wilkes-Barre).

The fitted distribution function of the largest exceedance for the 200-day interval, also measured from October 1, is

$$F(x) = \exp[-1.625 \exp(-2.628 \cdot 10^{-5} x)], \quad x \geq 0, \qquad (77)$$

The fitted function, $F(x)$, and the corresponding observed distribution, $F^*(x)$, of the largest exceedance for the 200-day interval, are given in Table 13 and shown in Fig. 31. Goodness of fit tests confirm a good agreement between the fitted and observed distributions for the 200-day interval. According to the chi-square test, $\chi^2 = 5.3524 < 12.6 = \chi^2_{cr}$, and

Table 12. Fitted and observed distributions of the largest exceedance for the 160-day interval (Susquehanna River at Wilkes-Barre).

No.	x (cfs)	Fitted F(x)	Observed F*(x)	|F(x) - F*(x)|	Absolute Frequency Fitted f	Absolute Frequency Observed f*	$\frac{(f - f^*)^2}{f}$
0	0	0.446	0.513	0.067	32.170	37	0.7250
1	20000	0.621	0.708	→ 0.087	12.610	14	0.1532
2	40000	0.755	0.819	0.064	9.610	8	0.2703
3	60000	0.847	0.875	0.028	6.610	4	1.0305
4	80000	0.906	0.902	0.004	4.305	2	1.2330
5	100000	0.943	0.931	0.012	2.670	2	0.1685
6	120000	0.966	0.972	0.006	1.637	3	1.1360
7	140000	0.980	1.000	0.020	0.980	2	1.0620
	$\sum$					72	5.7785

Table 13. Fitted and observed distributions of the largest exceedance for the 200-day interval (Susquehanna River at Wilkes-Barre).

No.	x (cfs)	Fitted F(x)	Observed F*(x)	|F(x) - F*(x)|	Absolute Frequency Fitted f	Absolute Frequency Observed f*	$\frac{(f - f^*)^2}{f}$
0	0	0.1967	0.1667	0.0300	14.160	12	0.3290
1	20000	0.3824	0.3472	0.0352	13.373	13	0.0104
2	40000	0.5667	0.5139	→ 0.0528	13.271	12	0.1217
3	60000	0.7152	0.7222	0.0070	10.686	15	1.7416
4	80000	0.8202	0.7917	0.0285	7.564	5	0.8691
5	100000	0.8895	0.8611	0.0284	4.990	5	0.0000
6	120000	0.9331	0.9305	0.0026	3.138	5	1.1048
7	140000	0.9598	0.9722	0.0124	1.921	3	0.6060
8	160000	0.9762	1.0000	0.0238	1.180	2	0.5698
	Σ					72	5.3524

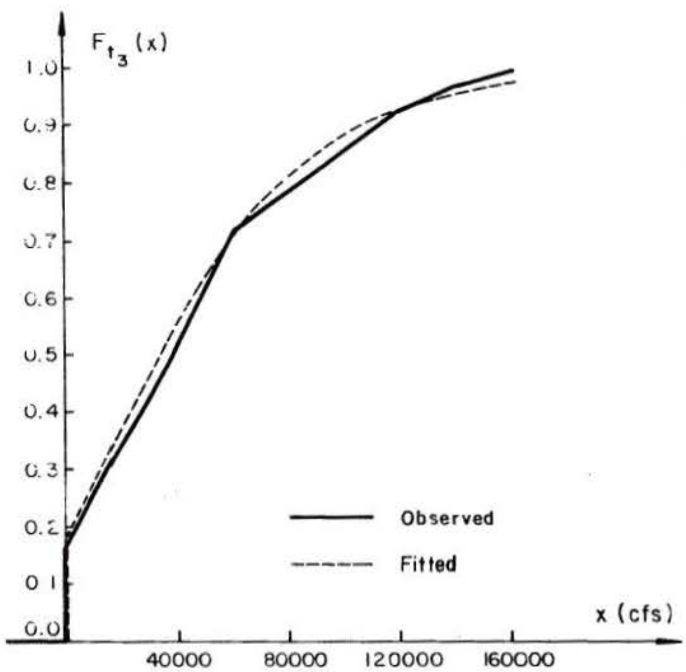

Fig. 31. Fitted and observed distributions of the largest exceedance for the 200-day interval (Susquehanna River at Wilkes-Barre).

according to the Kolmogorov-Smirnov test, $d = 0.0528 < 0.1602 = D_{cr}$.

To examine how the method presented in this study compares with an already existing method, comparison was made with Gumbel's distribution using maximum annual discharges for both rivers. Chi-square values were used as measures of goodness of fit of the observed distributions. The method presented in this study achieved better conformance with the data in both cases.

6.4 Comparison of the Results and the Method Developed in This Study and Gumbel's Method for the Susquehanna River at Wilkes-Barre Pa. The mean and standard deviation of the maximum annual discharges of the Susquehanna River at Wilkes-Barre, for the 72 years considered, are $\overline{Q}_a = 129887.50$ cfs, and $\sigma(Q_a) = 43083.86$ cfs. The two parameters in the double exponential function are estimated by Gumbel's method from 72 annual flood peak discharges.

Values of fitted Gumbel's F(x) function and the corresponding observed frequency distribution F*(x) of maximum annual peak discharges for the Susquehanna River are given in Table 14.

For the Gumbel's estimates of the two parameters the chi-square is $\chi^2 = 6.0130$, which is much greater for the same number of class intervals, than the chi-square obtained by the estimates of the method used in this study, with $\chi^2 = 3.9813$.

Table 14. Fitted Gumbel distribution function and observed frequency distribution of the maximum annual peak discharge for the Susquehanna River at Wilkes-Barre.

No.	x (cfs)	Gumbel's F(x)	Observed F*(x)	|F(x) - F*(x)|	Absolute Frequency Gumbel's f	Absolute Frequency Observed f*	$\frac{(f-f^*)^2}{f}$
1	85000	0.1178	0.1389	0.0211	8.481	10	0.27207
2	106000	0.3183	0.3056	0.0127	14.436	12	0.41106
3	127000	0.5418	0.5417	0.0001	16.092	17	0.05123
4	148000	0.7204	0.6945	0.0259	12.859	11	0.26875
5	169000	0.8390	0.8056	0.0334	8.539	8	0.03402
6	190000	0.9103	0.8750	→0.0353	5.134	5	0.00349
7	211000	0.9509	0.9444	0.0065	2.923	5	1.47585
8	232000	0.9734	1.0000	0.0266	1.620	4	3.49654
	Σ					72	6.01301

Chapter 7

GENERAL SUMMARY

This study presents a stochastic model for the analysis of the largest flood peak. The treatment takes into consideration that, for any given time interval, both the number of flood peaks and their magnitude are random variables. This stochastic model of the flood phenomenon that determines the largest flood peak has no constraints such as the stringent assumption of time-invariance (stationarity) of time series. It is general enough to embrace the concept of seasonality in flood occurrence. Treating flood peaks, obtained from the partial-duration series, as independent events, the theory herein employed adopts an approximation in the case of a complex flood hydrograph. It considers any complex flood hydrograph as one streamflow event, and uses only the highest flood peak of such a hydrograph. This assumption does not seriously affect the accuracy of the results obtained by the method. The distribution of the relatively small number of flood peak exceedances occurring in a time interval closely approximates the time-dependent Poissonian distribution, which deals with small probabilities and gives the number of rare events. Inclusion, in the model, of the distribution of the number of flood peak occurrences to study the largest magnitude of the flood peaks represents a contribution to the generality of methods used for flood analysis. This study concentrates on the largest flood peak exceedance among a random number of flood peak exceedances occurring in an interval of time. The magnitudes of exceedances are random continuous variables of a stochastic, discrete, non-negative process. This stochastic process of exceedances is periodic, with one year as its period.

When using the simple exponential distribution function of the magnitude of all exceedances, this method yields results that show good agreement between the resulting double-exponential distribution function of the magnitude of the largest exceedance and the corresponding observed frequency distribution. The corresponding theoretical double exponential distribution function obtained by a different approach than in the case of asymptotic distributions of extremes, fitted better the observed data for the two examples used than in the case of using the annual flood peak series. For the example used, the agreement is somewhat better for the Greenbrier River at Alderson, than for the Susquehanna River at Wilkes-Barre because a greater number of exceedances were available from the Greenbrier records.

REFERENCES

1. Barndoff, N., (1963). On the Limit Behavior of Extreme Order Statistics, Annals of Mathematical Statistics., Vol. 34, 992-1002.

2. Baxter, G. and M. D. Donsker, (1957). On the Distribution of the Supremum Functional for Processes with Stationary Independent Increments, Transactions of the American Mathematical Society, Vol. 85, 73-87.

3. Benham, A. D., (1965), (1966). Application of the Theory of Extreme Values to the Estimation of Flood Discharge, Water Resources Series No. 34, 100-118.

4. Benson, Manuel A., (1968). Uniform Flood Frequency Estimating Methods for Federal Agencies, Water Resources Research, Vol. 4, No. 5, October.

5. Benson, M. A., and N. C. Matalas, (1967). Synthetic Hydrology Based on Regional Statistical Parameters, Water Resources Research, Vol. 3, No. 4, Fourth Quarter.

6. Berman, S. M., (1962). Limiting Distribution of the Maximum Term in Sequences of Dependent Random Variables, Annals of Mathematical Statistics., Vol. 33, 894-908.

7. Billingsley, P., (1962). Limit Theorems for Randomly Selected Partial Sums, Annals of Mathematical Statistics., Vol. 33, 85-92.

8. Blum, J. R., D. L. Hanson, and J. I. Rosenblatt, (1963). On the Central Limit Theorem for the Sum of Random Number of Independent Random Variables, Wahrscheinlichkeits Theorie, Vol. 1, 389-393.

9. Cartwright, D. E., and M. S. Longnet-Higgins, (1956). The Statistical Distribution of the Maxima of a Random Function, Royal Society of London, Proceedings Series of Mathematics and Physical Sciences, 237, Sept. 25 - Nov. 20.

10. Chow, V. T., (1951). A General Formula For Hydrologic Frequency Analysis, Transactions, American Geophysical Union, Vol. 32, No. 2, 231-237.

11. Chow, V. T., (1964). Editor-in-Chief, Handbook of Applied Hydrology, McGraw-Hill Book Company.

12. Coutagne, A., (1937). Etude Statistique des Débits de Crue. Revue Générale de l'Hydraulique, Paris.

13. Cramer, H., (1962). On the Maximum of a Normal Stationary Stochastic Process, Bulletin of the American Mathematical Society, Vol. 68, 512-516.

14. Creager, W. P., (1939). Possible and Probable Future Floods, Civil Engineering, November.

15. Department of the Army, Office of the Chief of Engineers, (1952). Standard Project Flood Determinations, Washington, D. C., Civil Engineer Bulletin No. 52-58, 26 March.

16. Dobrushin, R. L. (1955). A Lemma on the Limit of a Composite Random Function, Uspekhi Math. Nauk, Vol. 10, No. 2, 157-159.

17. Dodd, E. L., (1923). The Greatest and the Least Variate Under General Laws of Error, Transactions of the American Mathematical Society, 25:525.

18. Dwass, M., (1964). Extremal Stochastic Processes, Annals of Mathematical Statistics, Vol. 35, 1718-1725.

19. Dynkin, E. B., (1955). Some Limit Theorems for Sums of Independent Random Variables with Infinite Mathematical Expectation, Izv. Akad. Nauk SSSR, Vol. 19, 247-266. (English Translation, American Mathematical Society).

20. Eagleson, Peter S. (1970). Dynamic Hydrology, McGraw-Hill Book Company, New York.

21. Feller, W., (1949). Fluctuation Theory of Recurrent Events, Transactions of the American Mathematical Society, Vol. 67, 98-119.

22. Fisher, R. A., and Tippet, L. H. C., (1928). Limiting Forms of the Frequency Distribution of the Largest or Smallest Member of a Sample, Proceedings of the Cambridge Philosophical Society, 24:180.

23. Foster, E. E., Evaluation of Flood Losses and Benefits, American Society of Civil Engineers, Transactions, Paper No. 2149.

24. Fréchet, M., (1927). Sur la Loi de Probabilité de l'Ecart Maximum, Ann. de la Soc. Polonaise de Math. (Cracow), 6:93.

25. Fuller, W. E., (1914). Flood Flows, Transactions of the American Society of Civil Engineers, 77:564.

26. Gnedenko, B. V., (1943). Sur la Distribution limite du Terme Maximum d'une Serie Aleatoire, Annals of Mathematics. Vol. 44, 423-453.

27. Guiasu, S. (1965). Asymptotic Distribution for Stochastic Processes with Random Discrete Time, Fourth Prague Conference, 307-321.

28. Gumbel, E. J., (1945). Floods Estimated by Probability Method, Engineering News-Record, June 14.

29. Gumbel, E. J., (1958). Statistics of Extremes, Columbia University Press, New York.

30. Hall, Warren A., and David T. Howell, (1963). Estimating Flood Probabilities within Specific Time Intervals, Journal of Hydrology 1, 265-271, North-Holland Publishing Co., Amsterdam.

31. Hardy, G. H., (1949). Divergent Series, Oxford.

32. Kimbal, F. Bradford, (1955). Practical Applications of the Theory of Extreme Values, American Statistical Association Journal, June.

33. Kirby, W., (1969). On the Random Occurrence of Major Floods, Water Resources Research, Vol. 5, No. 4, August.

34. Langbein, W. B., (1949). Annual Floods and the Partial-Duration Flood Series, Transactions, American Geophysical Union, Vol. 30, No. 6, Dec.

35. Linsley, R. K., M. A. Kohler and J. H. Paulhus, (1958). Hydrology for Engineers, McGraw-Hill Book Company, New York.

36. McGilchrist, C. A., K. D. Woodyer, and T. G. Chapman, (1969). 1. Recurrence Intervals Between Exceedances of Selected River Levels 2. Alternatives to a Markov Model, Vol. 5, No. 1, Water Resources Research, Feb.

37. Mises, R. Von, (1936). La Distribution de la Plus Grande de n Valeurs, Revue Math. de l'Union Interbalkanique (Athens), 1:1.

38. Moran, P. A. P. (1957). The Statistical Treatment of Flood Flows, Transactions, American Geophysical Union, Vol. 38, No. 4, August.

39. Mygorody, J. (1962). Central Limit Theorem for the Sum of Random Number of Independent Random Variables, Publication of Mathematical Institute of Hungarian Academy of Science., Vol. 7, 409-429.

40. Mygorody, J. (1965). Limit Distributions for Sequences of Random Variables with Random Indices, Fourth Prague Conference, 463-470.

41. Newell, G. F., (1962). Asymptotic Extreme Value Distribution for One-Dimensional Diffusion Processes, Journal of Mathematical Mechanics, Vol. 11, 481-496.

42. Renyi, A., (1960). On the Central Limit Theorem for Sum of Random Number of Independent Random Variables, Hungarica Acta Mathematica Scient. Vol. 11, 97-102.

43. Richter, W., (1965). Limit Theorems for Sequences of Random Variables with Sequences of Random Indices, Theory of Probability and its Application, Vol. 10, 74-84.

44. Seshappa Rao, G. N., J. R. Assenzo and J. F. Harp, (1966). Estimating Peak Rates of Runoff and Design Hydrograph in the State of Oklahoma, Journal of Hydrology 4, 147-170.

45. Shane, R. M. and W. R. Lynn, (1964). Mathematical Model for Flood Risk Evaluation, Journal of the Hydraulics Division, HY 6, 1-20.

46. Tippett, L. H. C., (1925). On the Extreme Individuals and the Range of Samples Taken From a Normal Population, Biometrika, 17:364.

47. Todorovic, P., (1968). A Mathematical Study of Precipitation Phenomena, Colorado State University CET67-68PT65, May.

48. Todorovic, P., (1970). On Some Problems Involving Random Number of Random Variables, Annals of Mathematical Statistics, Vol. 41, No. 3.

49. Todorovic, P., and V. Yevjevich (1969). Stochastic Process of Precipitation, Colorado State University Hydrology Paper No. 35, Sept.

50. Todorovic, P., and E. Zelenhasic, (1968). The Extreme Values of Precipitation Phenomena, Bulletin of the International Association of Scientific Hydrology, XIII, 4-12.

51. Walker, W. R., (1967). Similarities between Flood Control Problems and Other Aspects of Water Management in the United States, Water Resources Research, Vol. 3, No. 1, 11-12.

52. Watson, G. S., (1954). Extreme Values in Samples from m-Dependent Stationary Stochastic Processes, Annals of Mathematical Statistics, Vol. 25, 798-800.

53. Wisler, Chester O., and Ernest F. Brater, (1965). Hydrology, John Wiley & Sons, Inc. Publishers, Second Edition.

54. Yevjevich, V. (1956). Hidrologija, I Deo, (Hydrology, I Part), Hidr. Inst., J. Cerni, Belgrade, Yugoslavia.

APPENDIX 1

Derivation of the Asymptotic Value of the m-th Absolute Moment of the Largest Exceedance When Integer α is Large, for Gamma-Distributed Exceedances

The exact value of the m-th absolute moment of the largest exceedance, where $m = 1, 2, \ldots\ldots.,$ for gamma-distributed exceedances is

$$E\chi^m(t) = \frac{\lambda t}{\beta^m(\alpha-1)!}\int_0^\infty y^{\alpha+m-1}\exp\left[-y-\lambda t\, e^{-y}\sum_{i=0}^{\alpha-1}\frac{y^i}{i!}\right]dy \quad ,$$

where $y = \beta x$, and α is a positive integer. For large positive values of α ,

$$E\chi^m(t) \cong \frac{\lambda t}{\beta^m e^{-\alpha}\alpha^{\alpha-\frac{1}{2}}\sqrt{2\pi}}\int_0^\infty y^{\alpha+m-1}\exp(-y-\lambda t)dy \quad ,$$

$$\cong \frac{\lambda t\,\Gamma(\alpha+m)}{\beta^m\alpha^{\alpha-\frac{1}{2}}\sqrt{2\pi}}\, e^{\alpha-\lambda t} \quad ,$$

$$\cong \frac{\lambda t}{\beta^m}\,\frac{\exp(-\lambda t-m)}{\alpha^{\alpha-\frac{1}{2}}}\,(\alpha+m)^{\alpha+m-\frac{1}{2}} \quad .$$

Since

$$\lim_{\alpha\to\infty}\left[\frac{(\alpha+m)^{\alpha+m-\frac{1}{2}}}{\alpha^{\alpha-\frac{1}{2}}}\right] = e^m\,(\alpha+m)^m \quad ,$$

one can use, for large positive values of α the asymptotic value of the m-th absolute moment of the largest gamma-distributed exceedance,

$$E\chi^m(t) \cong \frac{\lambda t}{\beta^m}\, e^{-\lambda t}(\alpha+m)^m \quad ,$$

for any $m = 1, 2, \ldots\ldots$.

Table 15. Computed values for the relation $\beta^2\sigma_t^2(x)=f(\lambda t)$, for exponentially distributed magnitudes of exceedances.

λt	$\beta^2\sigma_t^2(x)$	λt	$\beta^2\sigma_t^2(x)$	λt	$\beta^2\sigma_t^2(x)$	λt	$\beta^2\sigma_t^2(x)$
1	1.267	26	28.680	51	64.200	76	101.600
2	1.940	27	30.040	52	65.660	77	103.100
3	2.548	28	31.400	53	67.140	78	104.600
4	3.219	29	32.770	54	68.610	79	106.100
5	3.976	30	34.150	55	70.090	80	107.600
6	4.812	31	35.530	56	71.570	81	109.200
7	5.717	32	36.920	57	73.050	82	110.700
8	6.681	33	38.320	58	74.530	83	112.200
9	7.694	34	39.720	59	76.020	84	113.700
10	8.749	35	41.130	60	77.510	85	115.300
11	9.841	36	42.540	61	79.000	86	116.800
12	10.960	37	43.950	62	80.490	87	118.300
13	12.120	38	45.370	63	81.990	88	119.800
14	13.290	39	46.800	64	83.480	89	121.400
15	14.490	40	48.230	65	84.980	90	122.900
16	15.710	41	49.660	66	86.480	91	124.400
17	16.950	42	51.100	67	87.980	92	126.000
18	18.200	43	52.540	68	89.480	93	127.500
19	19.470	44	53.990	69	90.990	94	129.000
20	20.750	45	55.440	70	92.500	95	130.600
21	22.050	46	56.890	71	94.000	96	132.100
22	23.350	47	58.340	72	95.510	97	133.700
23	24.670	48	59.800	73	97.020	98	135.200
24	26.000	49	61.260	74	98.540	99	136.700
25	27.340	50	62.730	75	100.100	100	138.300

Table 16. Computation of the harmonics $Z_j(t)$ of the fourier series used to represent the function $\Lambda(t)$, for the Greenbrier River at Alderson, W. Va.

t	$\hat{\Lambda}(t)$	$0.158\dot{3}t$	$y_t = \hat{\Lambda}(t) - 0.158\dot{3}t$	j = 1		j = 2		j = 3		j = 6	
				$y_t \cos\left(\frac{2\pi t}{18}\right)$	$y_t \sin\left(\frac{2\pi t}{18}\right)$	$y_t \cos\left(\frac{2\pi t}{9}\right)$	$y_t \sin\left(\frac{2\pi t}{9}\right)$	$y_t \cos\left(\frac{2\pi t}{6}\right)$	$y_t \sin\left(\frac{2\pi t}{6}\right)$	$y_t \cos\left(\frac{2\pi t}{3}\right)$	$y_t \sin\left(\frac{2\pi t}{3}\right)$
0	0.000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
1	0.056	0.15833	-0.10233	-0.09616	-0.03500	-0.07839	-0.06578	-0.05116	-0.08862	0.05116	-0.08862
2	(0.120)	0.31667	-0.19667	-0.15066	-0.12642	-0.03415	-0.18582	0.09834	-0.17032	0.09833	0.17032
3	0.194	0.47500	-0.28100	-0.14050	-0.24335	0.14050	-0.24335	0.28100	0.00000	-0.28100	0.00000
4	(0.340)	0.63333	-0.29333	-0.05093	-0.28887	0.27564	-0.10032	0.14666	0.25403	0.14666	-0.25403
5	0.681	0.79167	-0.11067	0.01922	-0.10899	0.10400	0.03785	-0.05533	0.09584	0.05533	0.09584
6	(0.930)	0.95000	-0.02000	0.01000	-0.01732	0.01000	0.01732	-0.02000	0.00000	-0.02000	0.00000
7	1.195	1.10833	0.08667	-0.06639	0.05571	0.01505	-0.08189	0.04333	0.07506	-0.04333	0.07505
8	1.625	1.26667	0.35833	-0.33672	0.12255	0.27449	-0.23033	-0.17916	0.31032	-0.17916	-0.31032
9	2.070	1.42500	0.64500	-0.64500	0.00000	0.64500	0.00000	-0.64500	0.00000	0.64500	0.00000
10	2.310	1.58333	0.72667	-0.68284	-0.24853	0.55666	0.46710	-0.36333	-0.62932	-0.36333	0.62932
11	2.473	1.74167	0.73133	-0.56023	-0.47009	0.12700	0.69097	0.36566	-0.63335	-0.36566	-0.63335
12	2.630	1.90000	0.73000	-0.36500	-0.63220	0.36500	0.63220	0.73000	0.00000	-0.73000	0.00000
13	(2.730)	2.05833	0.67167	-0.11663	-0.66147	-0.63116	0.22972	0.33584	0.58169	-0.33583	0.58169
14	(2.790)	2.21667	0.57333	0.09956	-0.56462	-0.53875	-0.19609	-0.28666	0.49652	-0.28666	-0.49652
15	(2.810)	2.37500	0.43500	0.21750	-0.37672	-0.21750	-0.37672	-0.43500	0.00000	0.43500	0.00000
16	(2.825)	2.53333	0.29167	0.22343	-0.18748	0.05065	-0.27557	-0.14583	-0.25259	-0.14583	0.25259
17	(2.840)	2.69167	0.14833	0.13938	-0.05073	0.11363	-0.09534	0.07416	-0.12846	-0.07416	-0.12845
18	2.850	2.85000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Σ				-2.50197	-3.83353	0.44767	0.22395	-0.10648	-0.08920	-0.06652	-0.10918
$a_j(t)$				-0.27800		0.04974		-0.01183		-0.00739	
$b_j(t)$					-0.42595		0.02488		-0.00991		-0.01213
$c_j(t)$				0.50864		0.05562		0.01543		0.01420	
$\theta_j(t)$				0.68407π		-0.14765π		0.77803π		0.67420π	
$Z_j(t)$				$0.50864 \cos(\frac{2\pi t}{18} + 0.68407\pi)$		$0.05562 \cos(\frac{2\pi t}{9} - 0.14765\pi)$		$0.01543 \cos(\frac{2\pi t}{6} + 0.77803\pi)$		$0.01420 \cos(\frac{2\pi t}{3} + 0.67420\pi)$	

Table 17. The fitting function $\Lambda_f(t)$; The Greenbrier River at Alderson, W. Va.

t	$Z_1(t)$	$Z_2(t)$	$Z_3(t)$	$Z_6(t)$	$\Lambda_f(t) = 0.2475 + 0.158\dot{3}t + \sum_{j=1,2,3,6} Z_j(t)$
0	-0.27802	0.04975	-0.01182	-0.00739	0.00000
1	-0.40693	0.05410	-0.01450	-0.00681	0.02167
2	-0.48676	0.03314	-0.00268	0.01420	0.12105
3	-0.50788	-0.00333	0.01182	-0.00739	0.21570
4	-0.46774	-0.03824	0.01450	-0.00681	0.38252
5	-0.37121	-0.05525	0.00268	0.01420	0.62957
6	-0.22985	-0.04642	-0.01182	-0.00739	0.90200
7	-0.06081	-0.01586	-0.01450	-0.00681	1.15783
8	0.11557	0.02209	-0.00268	0.01420	1.66333
9	0.27802	0.04975	0.01182	-0.00739	2.00468
10	0.40693	0.05410	0.01450	-0.00681	2.29953
11	0.48676	0.03314	0.00268	0.01420	2.52593
12	0.50788	-0.00333	-0.01182	-0.00739	2.63282
13	0.46774	-0.03824	-0.01450	-0.00681	2.70400
14	0.37121	-0.05525	-0.00268	0.01420	2.79163
15	0.22985	-0.04642	0.01182	-0.00739	2.81034
16	0.06081	-0.01586	0.01450	-0.00681	2.83345
17	-0.11557	0.02209	0.00268	0.01420	2.86255
18	-0.27802	0.04975	-0.01182	-0.00739	2.85000

APPENDIX 2

The Greenbrier River
at Alderson, W. Va.

DATA

Base for partial duration
series, Q_b = 17000 cfs

Table 18. Flood peaks and magnitudes of flood peak exceedances.

Water Year	Date	Q (cfs)	ξ (cfs)
1896	Mar. 30, 1896	28800	11800
1897	Nov. 5, 1896	27600	10600
	Feb. 23, 1897	54000	37000
	May 14, 1897	40900	23900
1898	Mar. 30, 1898	17100	100
	May 7, 1898	18600	1600
	Aug. 11, 1898	52500	35500
1899	Oct. 22, 1898	25300	8300
	Jan. 7, 1899	20000	3000
	Feb. 27, 1899	23800	6800
	Mar. 5, 1899	48900	31900
1900	Mar. 21, 1900	17100	100
1901	Nov. 26, 1900	56800	39800
	Jan. 12, 1901	21100	4100
	Apr. 21, 1901	20400	3400
	May 28, 1901	19300	2300
	June 17, 1901	20000	3000
1902	Dec. 15, 1901	36700	19700
	Mar. 1, 1902	43800	26800
1903	Jan. 3, 1903	25300	8300
	Feb. 5, 1903	29600	12600
	Feb. 17, 1903	33500	16500
	Feb. 28, 1903	34400	17400
	Mar. 23, 1903	48900	31900
1904	Jan. 23, 1904	25700	8700
	May 19, 1904	25700	8700
1905	Mar. 10, 1905	29600	12600
	May 12, 1905	37600	20600
1906	Jan. 4, 1906	18200	1200
	Jan. 23, 1906	26000	9000
1907	June 9, 1907	17500	500
	June 14, 1907	52500	35500
1908	Dec. 11, 1907	17800	800
	Dec. 24, 1907	23000	6000
	Jan. 12, 1908	31500	14500
	Feb. 6, 1908	52500	35500
	Mar. 7, 1908	26800	9800
	Apr. 1, 1908	27600	10600
	May 8, 1908	31500	14500
1909	Apr. 15, 1909	20000	3000
1910	June 17, 1910	45900	28900
1911	Jan. 30, 1911	43800	26800
	Apr. 5, 1911	20000	3000
1912	Oct. 18, 1911	23800	6800
	Feb. 22, 1912	18900	1900
	Feb. 27, 1912	18900	1900
	Mar. 16, 1912	35500	18500
	Mar. 29, 1912	27200	10200
	May 12, 1912	20000	3000
	May 17, 1912	21100	4100

Water Year	Date	Q (cfs)	ξ (cfs)
1913	Mar. 15, 1913	21800	4800
	Mar. 27, 1913	64000	47000
	Apr. 13, 1913	20000	3000
1914		0	0
1915	Jan. 7, 1915	34000	17000
	Feb. 2, 1915	40800	23800
1916	Oct. 2, 1915	27200	10200
	Dec. 30, 1915	24400	7400
1917	Dec. 29, 1916	17300	300
	Mar. 4, 1917	43000	26000
	Mar. 13, 1917	28000	11000
1918	Feb. 27, 1918	17900	900
	Mar. 14, 1918	77500	60500
	June 26, 1918	24000	7000
1919	Oct. 31, 1918	28600	11600
	Dec. 23, 1918	24800	7800
	Jan. 2, 1919	49000	32000
1920	Dec. 7, 1919	38000	21000
	Jan 25, 1920	20700	3700
	Mar. 20, 1920	33500	16500
1921		0	0
1922	Nov. 1, 1921	21500	4500
	Dec. 25, 1921	20100	3100
	Feb. 21, 1922	22200	5200
1923	Feb. 2, 1923	19500	2500
1924	Jan. 17, 1924	26500	9500
	Mar. 29, 1924	20400	3400
	May 12, 1924	36200	19200
	Sept. 30, 1924	17900	900
1925		0	0
1926	Jan. 20, 1926	20700	3700
	Feb. 15, 1926	17600	600
1927	Nov. 16, 1926	17900	900
	Dec. 22, 1926	24000	7000
	Dec. 26, 1926	40200	23200
	Feb. 6, 1927	18800	1800
	Feb. 20, 1927	19500	2500
1928	May 1, 1928	18000	1000
1929	Dec. 1, 1928	22800	5800
	Feb. 28, 1929	32700	15700
	Mar. 6, 1929	23800	6800
	May 21, 1929	20000	3000
1930	Nov. 18, 1929	36600	19600
1931		0	0

Water Year	Date	Q (cfs)	ξ (cfs)
1932	Feb. 5, 1932	50100	33100
	Mar. 18, 1932	17600	600
	Mar. 28, 1932	31500	14500
	May 2, 1932	27500	10500
	July 5, 1932	21900	4900
1933	Mar. 20, 1933	26400	9400
1934	Mar. 5, 1934	32300	15300
	Mar. 8, 1934	20500	3500
	Mar. 28, 1934	27900	10900
1935	Nov. 30, 1934	19400	2400
	Jan. 23, 1935	49600	32600
	Mar. 13, 1935	22300	5300
	Mar. 26, 1935	17900	900
	Apr. 1, 1935	24800	7800
	May 7, 1935	20100	3100
	July 9, 1935	24800	7800
	Sept. 6, 1935	20800	3800
1936	Nov. 13, 1935	19400	2400
	Jan. 3, 1936	20800	3800
	Feb. 15, 1936	27100	10100
	Mar. 18, 1936	58600	41600
	Apr. 7, 1936	28300	11300
1937	Dec. 7, 1936	21200	4200
	Jan. 2, 1937	22300	5300
	Jan. 21, 1937	36600	19600
	Apr. 26, 1937	26400	9400
1938	Oct. 20, 1937	21200	4200
	Oct. 28, 1937	32800	15800
	May 25, 1938	22300	5300
1939	Jan. 31, 1939	40200	23200
	Feb. 4, 1939	41600	24600
	Feb. 11, 1939	21200	4200
	Apr. 17, 1939	17200	200
	July 30, 1939	19400	2400
1940	Apr. 20, 1940	29900	12900
	May 25, 1940	21500	4500
	May 31, 1940	19400	2400
	June 28, 1940	18700	1700
1941		0	0
1942	May 17, 1942	35300	18300
1943	Dec. 30, 1942	33600	16600
	Jan. 27, 1943	17200	200
	Mar. 13, 1943	36200	19200
	Apr. 20, 1943	21200	4200
1944	Feb. 23, 1944	25200	8200
	Mar. 1, 1944	17200	200
1945	Dec. 26, 1944	17900	900
	Jan. 2, 1945	19000	2000
1946	Jan. 8, 1946	43600	26600
1947	Jan. 21, 1947	20000	3000
	Mar. 14, 1947	24400	7400

Table 18. Flood peaks and magnitudes of flood peak exceedances - Continued

Water Year	Date	Q (cfs)	ξ (cfs)
1948	Feb. 14, 1948	35200	18200
	Mar. 24, 1948	23500	6500
	Apr. 14, 1948	40300	23300
1949	Dec. 4, 1948	18500	1500
	Dec. 16, 1948	37100	20100
	Jan. 6, 1949	26300	9300
	Apr. 14, 1949	23200	6200
1950	Jan. 31, 1950	31500	14500
1951	Dec. 4, 1950	25600	8600
	Dec. 8, 1950	27800	10800
	Feb. 2, 1951	26700	9700
	Feb. 22, 1951	18500	1500
	Mar. 31, 1951	19800	2800
	June 14, 1951	29300	12300
1952	Jan. 18, 1952	17800	800
	Jan. 28, 1952	19100	2100
	Mar. 12, 1952	27600	10600
1953	Feb. 22, 1953	47100	30100
	Mar. 24, 1953	20100	3100
1954	Mar. 1, 1954	29700	12700
	July 16, 1954	18800	1800

Water Year	Date	Q (cfs)	ξ (cfs)
1955	Oct. 1, 1954	32000	15000
	Feb. 7, 1955	28000	11000
	Mar. 6, 1955	44400	27400
	Mar. 23, 1955	26200	9200
1956	Mar. 15, 1956	18200	1200
1957	Jan. 24, 1957	23900	6900
	Jan. 30, 1957	28900	11900
	Apr. 6, 1957	22000	5000
1958	Dec. 8, 1957	21800	4800
	Dec. 27, 1957	23900	6900
	Mar. 31, 1958	22200	5200
	Apr. 7, 1958	17500	500
	May 6, 1958	26700	9700
1959	Jan. 22, 1959	17200	200
	June 3, 1959	23900	6900
1960	Dec. 13, 1959	17800	800
	Mar. 31, 1960	35500	18500
	Apr. 4, 1960	32500	15500
1961	Feb. 19, 1961	25000	8000
	Feb. 24, 1961	21800	4800
	Feb. 26, 1961	31400	14400
	May 7, 1961	17200	200

Water Year	Date	Q (cfs)	ξ (cfs)
1962	Oct. 21, 1961	34700	17700
	Dec. 13, 1961	20100	3100
	Dec. 19, 1961	21500	4500
	Jan. 7, 1962	17800	800
	Feb. 28, 1962	23200	6200
	Mar. 22, 1962	35500	18500
1963	Jan. 13, 1963	22700	5700
	Mar. 6, 1963	34800	17800
	Mar. 12, 1963	47200	30200
	Mar. 17, 1963	26100	9100
	Mar. 20, 1963	30400	13400
1964	Jan. 26, 1964	19100	2100
	Mar. 6, 1964	39600	22600
	Mar. 9, 1964	22800	5800
1965	Jan. 25, 1965	22000	5000
	Feb. 8, 1965	28400	11400
	Mar. 26, 1965	19800	2800
	Apr. 12, 1965	18600	1600
1966	Feb. 14, 1966	26400	9400
1967	Mar. 7, 1967	54500	37500
	Mar. 15, 1967	39900	22900
	May 7, 1967	20900	3900

Q = total flood peak
ξ = flood peak exceedance.

Table 19. Seasonal occurrence of number of exceedances for the Greenbrier River at Alderson, W. Va.

Season	Period	Absolute Frequency	Absolute Frequency	Relative Frequency
Fall	Sept. 21 - Nov. 4	10	29	0.049
	Nov. 5 - Dec. 20	19		0.093
Winter	Dec. 21 - Feb. 3	46	107	0.224
	Feb. 4 - Mar. 20	61		0.298
Spring	Mar. 21 - May 4	37	61	0.180
	May 5 - June 20	24		0.117
Summer	June 21 - Aug. 5	6	8	0.029
	Aug. 6 - Sept. 20	2		0.010
Total Year	$\sum$	205	205	1.000

Note: Period of 72 years considered.

Table 20. Seasonal occurrence of number of exceedances for the Susquehanna River at Wilkes-Barre, Pa.

Season	Period	Absolute Frequency	Absolute Frequency	Relative Frequency
Fall	Sept. 21 - Nov. 4	5	14	0.0368
	Nov. 5 - Dec. 20	9		0.0661
Winter	Dec. 21 - Feb. 3	15	65	0.1100
	Feb. 4 - Mar. 20	50		0.3676
Spring	Mar. 21 - May 4	46	53	0.3377
	May 5 - June 20	7		0.0524
Summer	June 21 - Aug. 5	2	4	0.0147
	Aug. 6 - Sept. 20	2		0.0147
Total Year	$\sum$	136	136	1.0000

Note: Period of 72 years considered.

The Susquehanna River
at Wilkes-Barre, Pa.

Base for partial duration series, Q_b = 82000 cfs

Table 21. Total flood peaks and flood magnitudes of peak exceedances

Water Year	Date	Q (cfs)	ξ (cfs)
1891	Jan. 24, 1891	164000	82000
	Feb. 19, 1891	130000	48000
	Feb. 27, 1891	125000	43000
1892	Jan. 14, 1892	97100	15100
	Feb. 26, 1892	97100	15100
	Apr. 4, 1892	112000	30000
1893	May 5, 1893	115000	33000
1894	Mar. 8, 1894	88600	6600
	May 21, 1894	97100	15100
1895	Apr. 10, 1895	113000	31000
1896	Jan. 1, 1896	88600	6600
	Feb. 7, 1896	88600	6600
	Apr. 1, 1896	135000	53000
1897	Oct. 15, 1896	88600	6600
1900	Jan. 21, 1900	86800	4800
	Mar. 2, 1900	94500	12500
1901	Nov. 28, 1900	115000	33000
	Mar. 12, 1901	89000	7000
	Mar. 28, 1901	112000	30000
	Apr. 8, 1901	82100	100
	Apr. 23, 1901	90300	8300
1902	Dec. 16, 1901	166000	84000
	Mar. 2, 1902	213000	131000
	Mar. 18, 1902	101000	19000
1903	Dec. 23, 1902	82100	100
	Feb. 5, 1903	92800	10800
	Mar. 2, 1903	110000	28000
	Mar. 10, 1903	93700	11700
	Mar. 12, 1903	91100	9100
	Mar. 25, 1903	119000	37000
	Aug. 30, 1903	101000	19000
1904	Oct. 11, 1903	112000	30000
	Jan. 23, 1904	101000	19000
	Feb. 10, 1904	152000	70000
	Mar. 9, 1904	204000	122000
	Mar. 27, 1904	124000	42000
1905	Mar. 26, 1905	129000	47000
1906		0	0
1907		0	0
1908	Dec. 12, 1907	95400	13400
	Dec. 25, 1907	86100	4100
	Feb. 17, 1908	130000	48000
	Mar. 16, 1908	106000	24000
	Mar. 30, 1908	98800	16800
1909	Feb. 21, 1909	85300	3300
	Feb. 26, 1909	85300	3300
	May 2, 1909	125000	43000
1910	Jan. 23, 1910	93700	11700
	Mar. 3, 1910	157000	75000
	Apr. 25, 1910	112000	30000
1911	Mar. 29, 1911	94500	13400
1912	Mar. 31, 1912	115000	33000
	Apr. 3, 1912	127000	45000
1913	Jan. 9, 1913	97200	15200
	Mar. 28, 1913	184000	102000
1914	Mar. 29, 1914	182000	100000
	Apr. 9, 1914	107000	25000
	May 14, 1914	105000	23000
1915	Jan. 9, 1915	84900	2900
	Feb. 17, 1915	84900	2900
	Feb. 26, 1915	127000	45000
	July 10, 1915	120000	38000
1916	Apr. 2, 1916	160000	78000
1917		0	0
1918	Mar. 2, 1918	85700	3700
	Mar. 15, 1918	124000	42000
1919		0	0
1920	Mar. 13, 1920	155000	73000
1921	Mar. 10, 1921	86600	4600
1922	Nov. 29, 1921	117000	35000
	Mar. 9, 1922	83200	1200
1923	Mar. 5, 1923	91800	9800
1924	Apr. 8, 1924	129000	47000
1925	Oct. 1, 1924	111000	29000
	Feb. 13, 1925	145000	63000
1926	Mar. 26, 1926	90100	8100
	Apr. 10, 1926	83200	1200
1927	Nov. 17, 1926	121000	39000
	Mar. 15, 1927	92700	10700
	May 26, 1927	108000	26000
1928	Oct. 20, 1927	141000	59000
	May 1, 1928	102000	20000
1929	Mar. 17, 1929	127000	45000
	Apr. 22, 1929	159000	77000
1930		0	0
1931		0	0
1932	Apr. 2, 1932	107000	25000
1933	Aug. 25, 1933	99800	17800
1934	Mar. 6, 1934	85500	3500
1935	Jan. 11, 1935	107000	25000
	July 10, 1935	151000	69000
1936	Mar. 13, 1936	184000	102000
	Mar. 20, 1936	232000	150000
1937		0	0
1938		0	0
1939	Feb. 22, 1939	137000	55000
	Mar. 28, 1939	82300	300
1940	Apr. 1, 1940	212000	130000
	Apr. 22, 1940	93000	11000
1941	Apr. 7, 1941	138000	56000
1942	Mar. 11, 1942	111000	29000
	Mar. 19, 1942	94600	12600
	May 24, 1942	82600	600
1943	Jan. 1, 1943	191000	109000
	Mar. 18, 1943	101000	19000
1944	May 9, 1944	90000	8000
1945	Mar. 5, 1945	119000	37000
	Mar. 18, 1945	95800	13800
	Mar. 23, 1945	97600	15600
1946	Mar. 10, 1946	94800	12800
	May 29, 1946	210000	128000
1947	Apr. 7, 1947	151000	69000
1948	Mar. 18, 1948	118000	36000
	Mar. 23, 1948	193000	111000
	Apr. 15, 1948	98700	16700
1949	Dec. 31, 1948	82700	700
1950	Mar. 30, 1950	172000	90000
	Apr. 6, 1950	119000	37000
1951	Nov. 27, 1950	119000	37000
	Dec. 5, 1950	114000	32000
	Apr. 1, 1951	128000	46000
1952	Mar. 13, 1952	124000	42000
1953	Dec. 12, 1952	98000	16000
1954		0	0
1955	Mar. 3, 1955	85900	3900
1956	Oct. 16, 1955	166000	84000
	Mar. 9, 1956	186000	104000
	Apr. 6, 1956	126000	44000
1957	Apr. 7, 1957	107000	25000
1958	Apr. 8, 1958	170000	89000
	Apr. 23, 1958	83800	1800
1959	Jan. 23, 1959	113000	31000
	Apr. 4, 1959	86600	4600
1960	Nov. 29, 1959	88000	6000
	Feb. 12, 1960	90100	8100
	Apr. 2, 1960	201000	119000
1961	Feb. 27, 1961	163000	81000
	Apr. 18, 1961	88000	6000
	Apr. 26, 1961	148000	66000
1962	Apr. 2, 1962	128000	46000
1963	Mar. 19, 1963	90500	8500
	Mar. 28, 1963	131000	49000
1964	Jan. 27, 1964	93900	11900
	Mar. 7, 1964	197000	115000
	Mar. 10, 1964	228000	146000

Q = total flood peak
ξ = flood peak exceedance.

Key Words: Hydrology, Flood Peaks, Flood Exceedances, Distribution of Maximum Exceedances, Random number of Random Variables.

Following the theory of the supremum of a random number of random variables a stochastic model is presented for interpretation, analysis, and prediction of the largest flood peak discharge above a given base level concerning a time interval [o,t], at a given location of a river. Although the analysis of floods is the main objective of the developed stochastic model, it has a broader scope. The model can be applied to any kind of data of an intermittent process having a substantial stochastic component for which probabilities of the largest value are desired.

The model has been applied in this study to data from gaging stations on the Susquehanna River at Wilkes-Barre, Pa., and the Greenbrier River at Alderson, W. Va. Results were compared to those obtained by Gumbel's method; they indicate that the introduced model fits the data better.

Key Words: Hydrology, Flood Peaks, Flood Exceedances, Distribution of Maximum Exceedances, Random number of Random Variables.

Following the theory of the supremum of a random number of random variables a stochastic model is presented for interpretation, analysis, and prediction of the largest flood peak discharge above a given base level concerning a time interval [o,t], at a given location of a river. Although the analysis of floods is the main objective of the developed stochastic model, it has a broader scope. The model can be applied to any kind of data of an intermittent process having a substantial stochastic component for which probabilities of the largest value are desired.

The model has been applied in this study to data from gaging stations on the Susquehanna River at Wilkes-Barre, Pa., and the Greenbrier River at Alderson, W. Va. Results were compared to those obtained by Gumbel's method; they indicate that the introduced model fits the data better.

Key Words: Hydrology, Flood Peaks, Flood Exceedances, Distribution of Maximum Exceedances, Random number of Random Variables.

Following the theory of the supremum of a random number of random variables a stochastic model is presented for interpretation, analysis, and prediction of the largest flood peak discharge above a given base level concerning a time interval [o,t], at a given location of a river. Although the analysis of floods is the main objective of the developed stochastic model, it has a broader scope. The model can be applied to any kind of data of an intermittent process having a substantial stochastic component for which probabilities of the largest value are desired.

The model has been applied in this study to data from gaging stations on the Susquehanna River at Wilkes-Barre, Pa., and the Greenbrier River at Alderson, W. Va. Results were compared to those obtained by Gumbel's method; they indicate that the introduced model fits the data better.

Key Words: Hydrology, Flood Peaks, Flood Exceedances, Distribution of Maximum Exceedances, Random number of Random Variables.

Following the theory of the supremum of a random number of random variables a stochastic model is presented for interpretation, analysis, and prediction of the largest flood peak discharge above a given base level concerning a time interval [o,t], at a given location of a river. Although the analysis of floods is the main objective of the developed stochastic model, it has a broader scope. The model can be applied to any kind of data of an intermittent process having a substantial stochastic component for which probabilities of the largest value are desired.

The model has been applied in this study to data from gaging stations on the Susquehanna River at Wilkes-Barre, Pa., and the Greenbrier River at Alderson, W. Va. Results were compared to those obtained by Gumbel's method; they indicate that the introduced model fits the data better.

PREVIOUSLY PUBLISHED PAPERS

Colorado State University Hydrology Papers

No. 26 "The Investigation of Relationship Between Hydrologic Time Series and Sun Spot Numbers," by Ignacio Rodriquez-Iturbe and Vujica Yevjevich, April 1968.

No. 27 "Diffusion of Entrapped Gas From Porous Media," by Kenneth M. Adam and Arthur T. Corey, April 1968.

No. 28 "Sampling Bacteria in a Mountain Stream," by Samuel H. Kunkle and James R. Meimann, March 1968.

No. 29 "Estimating Design Floods from Extreme Rainfall," by Frederick C. Bell, July 1968.

No. 30 "Conservation of Ground Water by Gravel Mulches," by A. T. Corey and W. D. Kemper, May 1968.

No. 31 "Effects of Truncation on Dependence in Hydrologic Time Series," by Rezaul Karim Bhuiya and Vujica Yevjevich, November 1968.

No. 32 "Properties of Non-Homogeneous Hydrologic Series," by V. Yevjevich and R. I. Jeng, April 1969.

No. 33 "Runs of Precipitation Series," by Jose Llamas and M. M. Siddiqui, May 1969.

No. 34 "Statistical Discrimination of Change in Daily Runoff," by Andre J. Dumas and Hubert J. Morel-Seytoux, August 1969.

No. 35 "Stochastic Process of Precipitation," by P. Todorovic and V. Yevjevich, September 1969.

No. 36 "Suitability of the Upper Colorado River Basin for Precipitation Management," by Hiroshi Nakamichi and Hubert J. Morel-Seytoux, October 1969.

No. 37 "Regional Discrimination of Change in Runoff," by Viboon Nimmannit and Hubert J. Morel-Seytoux, November 1969.

No. 38 "Evaluation of the Effect of Impoundment on Water Quality in Cheney Reservoir," by J. C. Ward and S. Karaki, March 1970.

No. 39 "The Kinematic Cascade as a Hydrologic Model," by David F. Kibler and David A. Woolhiser, February 1970.

No. 40 "Application of Run-Lengths to Hydrologic Series," by Jaime Saldarriaga and Vujica Yevjevich, April 1970.

No. 41 "Numerical Simulation of Dispersion in Groundwater Aquifers, by Donald L. Reddell and Daniel K. Sunada, June 1970.