A United States Contribution to the International Hydrological Decade

HEC-IHD-0300

Hydrologic Engineering Methods For Water Resources Development

Volume 3
Hydrologic Frequency Analysis

April 1975

Approved for Public Release. Distribution Unlimited.

IHD-3

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704-0188*

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	April 1975	IHD Volume 3

4. TITLE AND SUBTITLE
Hydrologic Data Management

5. FUNDING NUMBERS

6. AUTHOR(S)
Leo R. Beard, Augustine J. Fredrich

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
US Army Corps of Engineers
Institute for Water Resources
Hydrologic Engineering Center
609 Second Street
Davis, CA 95616-4687

8. PERFORMING ORGANIZATION REPORT NUMBER
IHD-3

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
N/A

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
N/A

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release. Distribution of this document is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
This is Volume 3 of the 12 volume report prepared by the Hydrologic Engineering Center of the U.S. Army Corps of Engineers as a contribution to the International Hydrological Decade.

This volume describes the use of probability theory and statistical methods as they apply to hydrologic engineering problems. The volume discusses general probability concepts and definitions, hydrologic frequency curves, graphical and analytical techniques for computing frequency curves, analytical adjustments to hydrologic frequencies, volume-frequency computations, and regional frequency analysis.

14. SUBJECT TERMS
International Hydrological Decade, Frequency Analysis

15. NUMBER OF PAGES
176

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNLIMITED

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

Hydrologic Engineering Methods for Water Resources Development

Volume 3
Hydrologic Frequency Analysis

April 1975

US Army Corps of Engineers
Institute for Water Resources
Hydrologic Engineering Center
609 Second Street
Davis, CA 95616

(530) 756-1104
(530) 756-8250 FAX
www.hec.usace.army.mil

IHD-3

FOREWORD

The material presented in this volume is essentially a revision of a publication entitled "Statistical Methods in Hydrology," January 1962, by Leo R. Beard. This latter work has served as a basic reference for hydrologic engineers in the Corps of Engineers and other organizations for more than a decade. The revision has consisted primarily of adding three chapters, 2, 3, and 6, on probability theory, statistical concepts, and the effect of development on frequency curves, and additional explanations throughout the text of the volume with the objective of providing the reader with more information on the principles underlying the techniques. Some additional texts published since the publication of "Statistical Methods" have been added to the list of references.

The revision was accomplished by Augustine J. Fredrich and was reviewed by Bill S. Eichert, Harold E. Kubik, John Peters and Leo R. Beard. Darryl W. Davis performed review, final editing and was responsible for production of the report.

CONTENTS

CONTENTS

CONTENTS

Page

LIST OF FIGURES

CONTENTS

Page

LIST OF FIGURES

EXHIBITS

Chapter 1

Introduction

CHAPTER 1. INTRODUCTION

Section 1.01. Nature and Objectives of Statistical Analysis

Statistical analysis as applied in hydrologic engineering consists of (a) estimating the future frequency or probability of hydrologic events based on information contained in hydrologic records and (b) correlating interrelated hydrologic variables. In probability analyses, statistical methods permit coordination of observed data to yield a more accurate estimate of future frequencies than is indicated by the observed data, and also provide criteria for judging the reliability of the frequency estimates. In correlation analyses, statistical methods provide means for deriving the best relationship for predicting the value of one variable from known values of other variables, and also provide criteria for judging the reliability of forecasts or estimates based on the derived relationship.

Section 1.02. Purpose and Scope

The purpose of this volume is to describe and illustrate the application of statistical analysis in hydrologic engineering. The scope of the volume includes the following items:

a. A review of the basic concepts of probability that are applicable in hydrologic engineering, with a guide to supplemental reading for further study.

b. Presentation of detailed computation procedures and supporting justifications and computation aids for derivation of probability or frequency estimates based on analysis of hydrologic records that have been adjusted to be consistent with selected reference base conditions.

c. A summary of procedures for developing "regionalized" hydrologic frequency estimates, based on analyses of hydrologic records available at stream gaging stations, adjusted to provide generalized flood-frequency relations that are considered most representative of long-period hydrologic characteristics in various drainage areas in the region. Also, illustrations and explanations of simple generalization procedures for use where these are adequate and advantageous are given.

Section 1.03. References

There are many textbooks on statistics, probability, and correlation, and a multitude of technical papers on the statistical aspects of hydrologic engineering. Selected references considered particularly applicable and useful for the purpose of supplementing material contained herein are given at the end of the text. Reference 33 is a user document for a computer program that can be used for most of the statistical computations described in this volume.

Chapter 2

General Probability Concepts and Procedures

CHAPTER 2. GENERAL PROBABILITY CONCEPTS AND PROCEDURES

Section 2.01. Introduction

The concepts of probability and frequency of occurrence are familiar in at least an informal, nonmathematical way, to almost everyone. Most people are familiar with some type of game of chance which involves the use of probability theory to compute the "odds" or the pay-off function that relates possible outcomes of the game to the return to the players. In most games of chance, it is possible to identify every possible outcome of the game and to compute the relative frequency or probability of each outcome through the use of exact mathematical formulae. Another example of the use of probability and frequency is the association of probabilistic estimates with weather forecasts, e.g., "30 percent chance of rainfall exceeding 1.5 inches of rain today." However, in this latter instance, the probability estimate has a somewhat different implication because it is not possible to identify discretely every possible outcome of today's weather and it is certainly not possible to compute the relative frequency of occurrence of each possible outcome in a precise mathematical way. These two common examples of the use of probability illustrate two fundamentally different concepts in the application of probability theory. In the first case, the set of all possible outcomes can be identified and the probability of occurrence of each outcome can be computed from known physical conditions. In the second case, the set of possible outcomes cannot be identified exactly, but can only be inferred from examination of a large number of occurrences (e.g., examining a long record of climatological data). Moreover, even if a set of possible outcomes is identified for the second case, the only method for obtaining an estimate of the probability of occurrence for each outcome is to calculate the relative frequency of occurrence by counting for each outcome the number of occurrences of that outcome and dividing by the total number of observations of all outcomes.

The set of all possible outcomes is called the parent population or, more commonly, the population. This term is used to describe the possible outcomes of a phenomenon regardless of whether the outcomes are known (i.e., can be identified exactly) or unknown (i.e., must be inferred from observations of occurrences). As indicated in the two cases discussed above, identification of the parent population is not, in itself, adequate for developing probability estimates. In addition to identifying the possible events in the parent population, it is necessary to have some knowledge of the frequency of occurrence of each event comprising the parent population. A complete description of the frequency of occurrence of each event in a parent population is called a frequency distribution or, more commonly, a distribution. Given a parent population and its distribution, it is possible to determine the probability of occurrence of any specified event in the future.

In order for probability theory to be strictly applicable in making probability estimates for future events, the parent population must have certain characteristics. First, the events or elements comprising the population must be homogeneous with respect to some specified property; that is, there must be at least one unifying characteristic which can be used as a basis for specifying not only elements that are included in the population, but also elements that are excluded from the population. Also, the individual outcomes comprising the population must be independent; that is, the occurrence or nonoccurrence of any event must not depend on or be related in any way to the occurrence or nonoccurrence of any other event in the population. Furthermore, the order of occurrence of outcomes must be random; that is, whether a particular event occurs or does not occur at a given time must be completely a matter of chance. Finally, the population and its frequency distribution must be stationary or time-invariant; that is, the events comprising a population and their associated probabilities of occurrence cannot change with time. Consequently, the physical processes that underlie the frequency of occurrence of the outcomes cannot change with time. If any one of these four properties is not an attribute of a population, the use of probability theory in making probabilistic estimates of future events may result in erroneous results. However, it may be possible to obtain useful probabilistic estimates--even when the population does not conform strictly to the limitations imposed by these properties--if sound judgment is incorporated into the analysis of the population and the development of the estimates.

In engineering problems, it is usually either impossible or infeasible to identify and analyze the entire population, because most problems involve physical or natural phenomena, and the recorded data on these phenomena are relatively limited. Any time the set of observed data does not constitute the entire population, that set of observed data is called a sample. For example, a 100-year record of river stages at a particular point on a particular stream is a sample of the population of all possible river stages at that point on that stream. By analyzing the characteristics exhibited by a sample, it is possible to make inferences about the characteristics of the population from which the sample is drawn.

Section 2.02. Populations and Distributions

The science concerned with collecting, organizing, analyzing and interpreting numerical data--particularly the analysis of population and sample characteristics--is called statistics. Before any statistical analysis of a population or its characteristics can be performed, it is necessary to define unambiguously the population to be studied. In the river-stage example cited previously, the population is defined by identifying river stages as the element or variable to be studied by specifying the stream and the location of the point on the stream where the river-stage data were collected. It is worthwhile to note at this point that river-stage data are an example of a population that is likely to be nonstationary, because the stage observation is dependent upon the control characteristics at the gaging location, and on many streams the control can change with time. Consequently, a stage observation of 10.0 meters in 1923 will probably not represent the same physical conditions as a stage observation of 10.0 meters in 1947 if the geometry of the stream channel at the gaging station changes between 1923 and 1947.

The terms element and outcome have been used to describe the constituents of a population. In engineering practice, the terms variable and event are often used with the same meaning as element or outcome. In defining a population it is necessary to specify not only the homogeneous characteristics of the population but also the particular attribute of the event that is to be studied and the unit of measurement that is associated with the attribute. In the river-stage illustration, for example, the attribute to be studied could be annual maximum river stage, annual minimum stage, or any other attribute that possesses the properties of independence, randomness and stationarity. The unit of measurement selected should be stationary and unambiguous. Both the attribute and the unit of measurement should be related as directly as possible to the physical or natural phenomenon underlying the population.

The variables or events comprising a population or sample can be either continuous or discrete. A continuous variable is one which can take on any value within one or more specified intervals. A discrete variable is one which can take on only a finite, countable number of values within one or more specified intervals. Annual maximum discharge is an example of a continuous variable, since the annual maximum discharge could be any value between zero and infinity. The number of coliform bacteria per 1000-milliliter water sample is an example of a discrete variable since the number can only assume integer values between zero and infinity.

As previously indicated, when dealing with natural phenomena it is almost always impossible to identify specifically each possible outcome in a population and to calculate on a theoretical basis its relative frequency of occurrence. This impossibility stems from an inability to identify all of the physical phenomena underlying a particular natural event and an inability to specify the interactions among physical phenomena that produce natural events of a given magnitude. For example, if one were interested in the population composed of coliform bacteria count per 1000-milliliter water sample at a given sampling location, it would be reasonable to state that the population consists of all integer numbers between zero and infinity, since there is no physical basis for excluding any specific integer within that range. However, it is not possible to calculate theoretically the exact frequency of occurrence of a particular integer value because the coliform bacteria count depends upon physical relationships between source and location of coliform bacteria entering the water from which the sample is drawn, temperature-controlled and nutrient-controlled reactions within the water body, natural purification processes and other processes. Our understanding of these processes and their interactions is not adequate to permit us to calculate how frequently they will produce a given coliform bacteria count. So, although a general description of the outcomes comprising the population can be deduced, the second important element required to completely describe the population--the frequency of occurrence of each outcome (the frequency distribution)--cannot be calculated or deduced. When the frequency distribution associated with a given population cannot be completely described _a priori_, it is necessary to resort to sampling to obtain an estimate of the distribution. Sampling and samples are discussed in more detail in Section 2.04.

The limits, or _bounds_, upon events which comprise a given population are most commonly inferred from knowledge of the physical phenomena involved. For example, in most problems the notion of negative streamflow is nonsensical, so a lower bound of zero is selected for populations which are comprised of variables which are functions of streamflow. Likewise, since there is usually no physical evidence that there is an upper limit on streamflow, the upper bound for populations involving functions of streamflow is usually considered to be infinity. Populations which have an upper limit of infinity are said to be unbounded at the upper end. Populations which contain variables that can take on any value between minus infinity and plus infinity are said to be unbounded. The difference between the largest value and smallest value in a population is called the range.

Once a population is described by specifying the homogeneous characteristic (or characteristics) which uniquely identify the elements comprising the population, all that remains for a complete description of the population is to specify the frequency distribution of the events comprising the population. The frequency distribution is defined by the relative frequency of occurrence of each event in the population. The relative frequency of occurrence is simply the number of occurrences of each particular event divided by the total number of occurrences of all events:

$$f_i = \frac{n_i}{N} \qquad (2\text{-}1)$$

where:

f_i = relative frequency of occurrence of the i^{th} event
n_i = number of occurrence of the i^{th} event
N = total number of occurrences of all events

If the variable describing the events is discrete, and the population is relatively small, it is possible to compute f_i for each event. However, if the variable describing the events is continuous (or if the variable is discrete, but the range of the population is large), it is necessary to establish class intervals (arbitrary subdivisions of the range) and define n_i as the number of events that occur within a class interval. In this case, f_i becomes the relative frequency of occurrence of events within the i^{th} class interval.

The frequency distribution can be represented in a variety of ways. In some cases, a tabular presentation showing the events or class intervals and the corresponding values of f_i is used. A more common way of presenting the frequency distribution is to prepare relative frequency plots such as the ones shown on fig. 2.01(a). This type of plot shows graphically the relative frequency of occurrence of all events in the population. It is not mandatory that all class intervals be of equal length as shown in the example. However, it is essential that the area of the rectangle representing the relative frequency be proportional to the relative frequency of occurrence in the class interval. Still another method of presenting a frequency distribution is to prepare cumulative frequency plots such as the ones shown on fig. 2.01(b). These cumulative plots show the proportion of events that exceed a given magnitude. The histogram type of plot (fig. 2.01(a)) is often used for discrete variables, and the cumulative type of plot (fig. 2.01(b)) is almost always used when working with continuous variables.

The information presented by these curves can be interpreted in a probability context because one definition of probability is the <u>relative frequency</u> definition which states that the probability of a particular event occurring is equal to the relative frequency of occurrence of that event. Thus, based on fig. 2.01, one could state that the probability of occurrence of an annual peak discharge between 3,000 m^3/sec and 3,999 m^3/sec in any future year is .28; the probability of occurrence of an annual peak discharge greater than 4,000 m^3/sec in any future year is about 0.46 or (1.-.54).

Through analysis of sampling experiments and natural phenomena, statisticians and mathematicians have developed theoretical probability

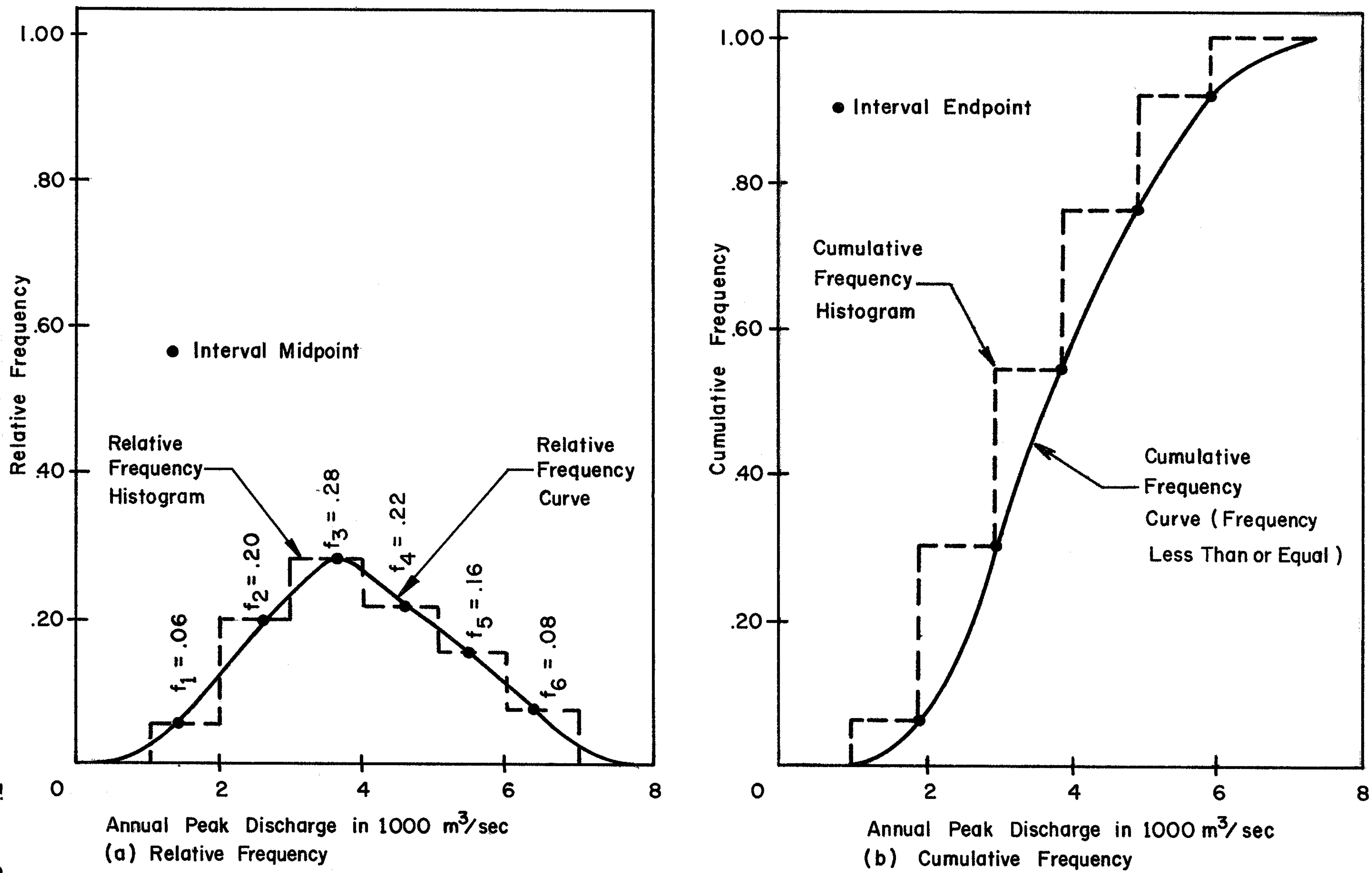


Figure 2.01 Graphical representations of frequency distributions

distributions that can be defined by general mathematical functions. It has been demonstrated that the apparent distribution of a large number of natural phenomena can be closely approximated through the use of these theoretical distributions. One of the most commonly used probability distributions is the normal, or Gaussian distribution. The mathematical function which defines this distribution is:

$$f(X) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(X-\mu)^2/2\sigma^2} \qquad -\infty \leq X \leq +\infty \qquad (2\text{-}2)$$

where:

X = a continuous random variable

σ = the population standard deviation of the random variable X

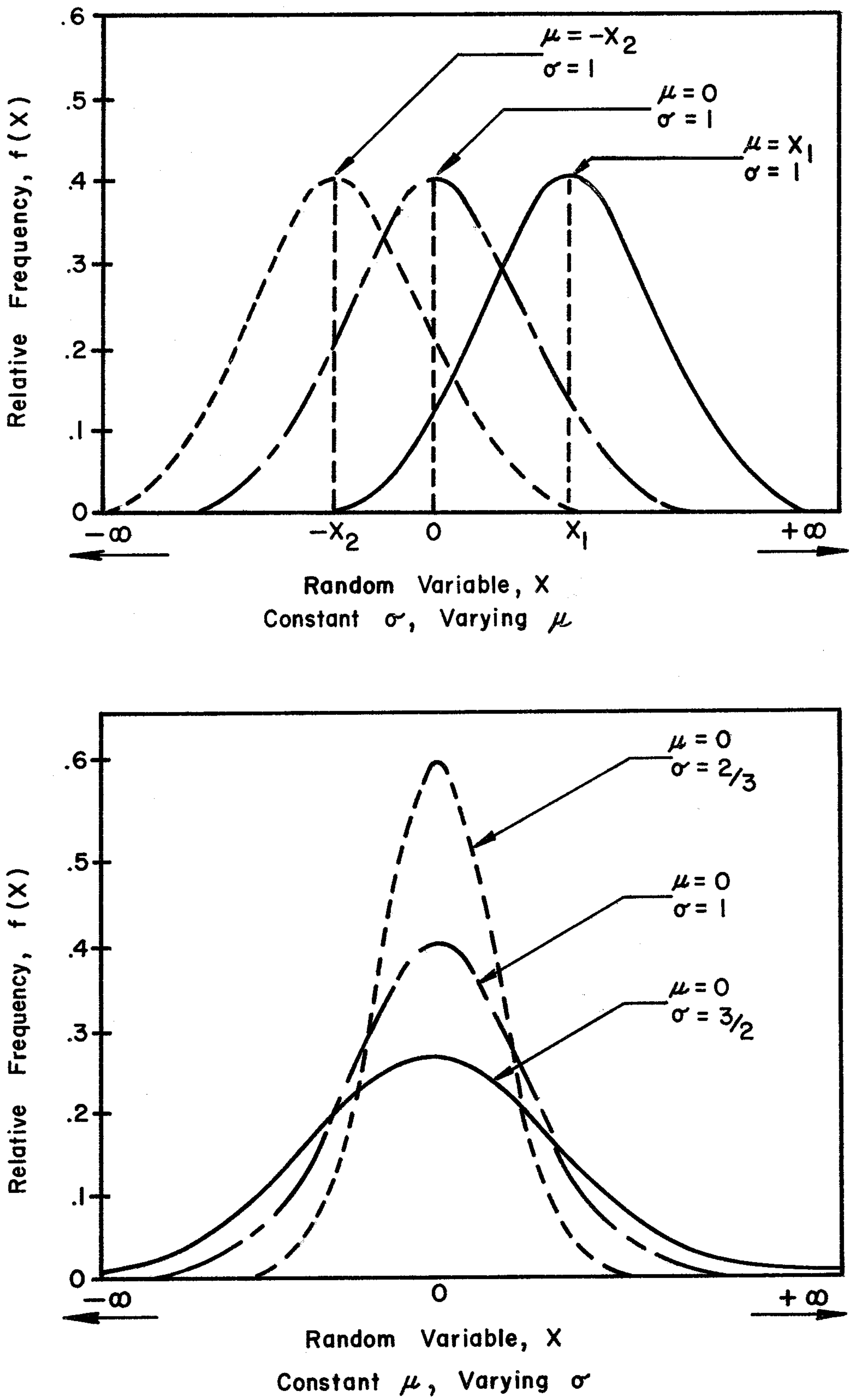
μ = the population mean of the random variable X

e = base of the Napierian logarithm

This function is called the probability density function (PDF) for the normal distribution. It is exactly analogous to the relative frequency curve previously discussed. Two parameters, μ and σ, must be known to define the function for a particular population. The function is unbounded (i.e., extends from $-\infty$ to $+\infty$) and is symmetrical about the parameter μ. Thus the parameter μ is a measure of location (it identifies the point along the axis between $-\infty$ and $+\infty$ at which the function is centered) for the population. The parameter σ is a measure of dispersion (it defines the degree to which the distribution is spread along the axis between $-\infty$ and $+\infty$) of the population. The significance of variations in μ and σ is shown on fig. 2.02.

By performing the following transformation equation 2-2 can be transformed to a standardized normal probability density function:

$$Z = \frac{X - \mu}{\sigma} \qquad (2\text{-}3)$$



Effects of variation in μ and σ on normal probability density function

Figure 2.02

where:

> Z = a standardized random variable, and X, μ and σ are defined as in equation 2-2.

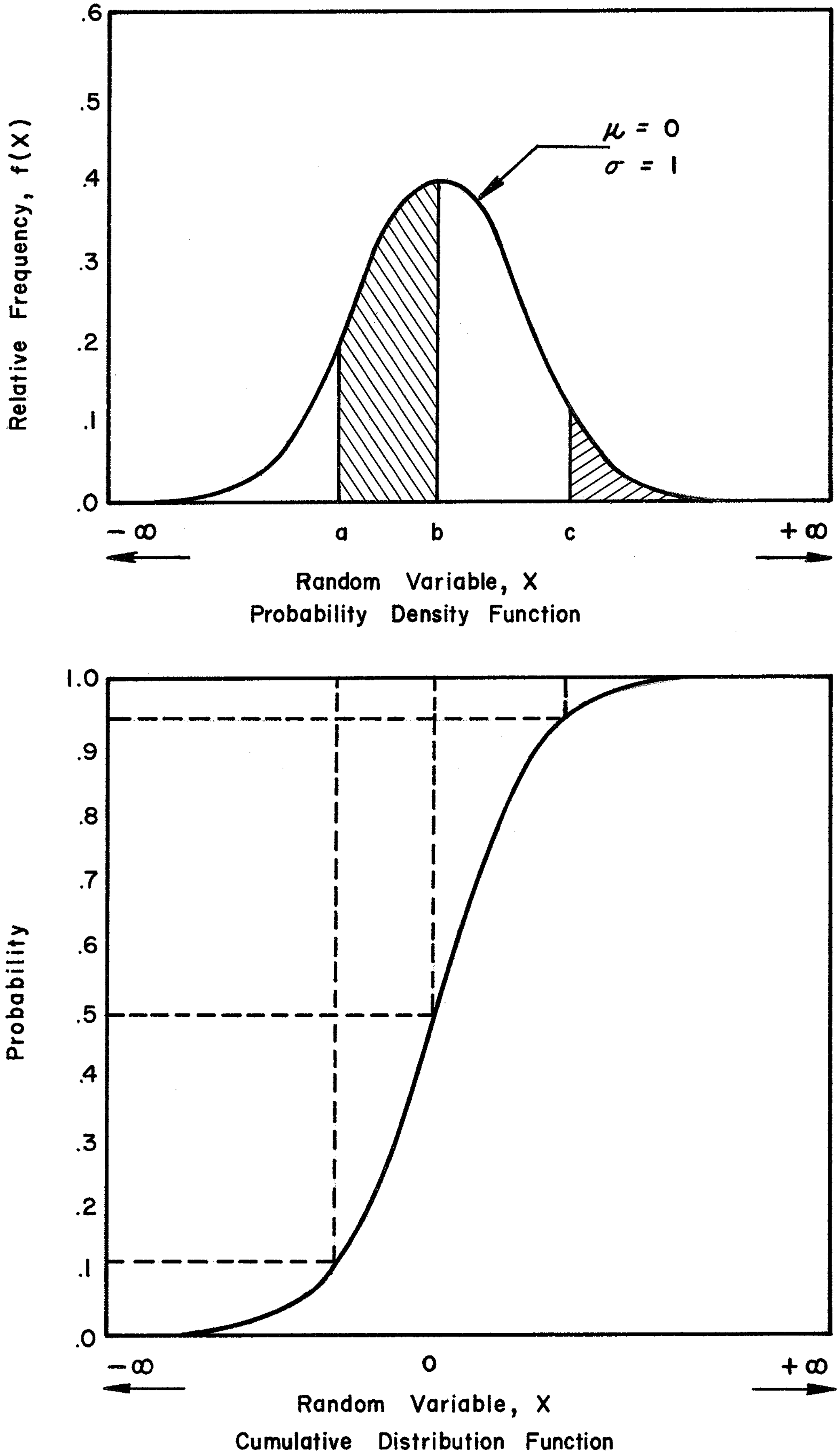
The standardized random variable, Z, is normally distributed--just as the random variable, X. The normal distribution of the random variable, Z, is called the <u>standard normal distribution</u>. For this standardized normal PDF, the parameter, μ_Z (mean of the standard normal distribution), has a value of zero and the parameter, σ_Z (standard deviation of the standard normal distribution), has a value of one. The probability density function for this standard normal distribution has been evaluated, and tables showing the values of $f(Z)$ are available in almost all statistical texts. By using the tables and reversing the transformation of equation 2-3, it is possible to transform the standard PDF to the probability density function for populations that are normally distributed with any values of μ and σ.

At this point, it should be noted that the concept of equivalence between relative frequency and probability is only meaningful when the number of events in the population is finite. When the number of outcomes is infinite, as in the case of a continuous random variable, it can be seen from examination of equation 2-1 that the denominator becomes infinity. This implies that the relative frequency or probability of occurrence of a stated specific value of a continuous random variable is zero. Consequently, the relative frequency and probability concepts only have meaning for a continuous random variable when they are associated with an interval along the X-axis. Therefore, we can speak of the probability of occurrence of some future event having a value between two magnitudes, X_1 and X_2, or greater than X_1 or less than X_1; but not about the probability that the event will have a value of exactly X_1. The interval between X_1 and X_2 can be very small; but as the size of the interval decreases the probability that a future event will have a value within the interval approaches zero.

Since the concept of probability for a continuous random variable only has meaning for intervals along the X-axis and since the normal distribution extends from $-\infty$ to $+\infty$ along the axis (which implies that the interval from $-\infty$ to $+\infty$ contains all possible values of the random variable), it follows that the probability of any future value of the random variable being between $-\infty$ to $+\infty$ must be unity; that is, it is certain that the future value will be within this interval. It can be shown that integration of equation 2-2 between the limits of $-\infty$ and $+\infty$ will produce a result of unity. Consequently, the area under the probability density function must be unity.

The foregoing discussion leads to the observation that probability estimates for continuous variables are related to areas under the PDF rather than to the ordinates of the PDF. The ordinates of the PDF relate the relative frequency of occurrence of values of a given magnitude to the frequency of occurrence of values of other magnitudes within the same population. For example, on fig. 2.02, the relative frequency in the vicinity of $X = 0$ for the population with $\mu = 0$ and $\sigma = 1$ is about 0.4, the relative frequency in the vicinity of $X = -X_2$ is 0.2, and the relative frequency in the vicinity of $X = X_1$ is 0.13. This means that observations in the vicinity of $X = 0$ will occur about twice as frequently as observations in the vicinity of $X = -X_2$ and about three times as frequently as observations in the vicinity of $X = X_1$. It should be noted that this implies nothing about the absolute probability of occurrence of $X = 0$, $X = -X_2$ or $X = X_1$ since it only describes how frequently these values occur with respect to each other.

The probability of occurrence of a value within a specified interval is represented by the area under the PDF within the interval. This is illustrated in fig. 2.03 where the shaded area between points a and b on the X-axis represents the probability of occurrence of a value of the random variable, X, greater than a and smaller than b. Likewise, the shaded area to the right of point c on fig. 2.03 represents the probability of occurrence of a value of X greater than c. The actual



Probability density function and cumulative distribution function for normal distribution with $\mu = 0$ and $\sigma = 1$

Figure 2.03

probability can be calculated by integrating equation 2-2 between the limits a and b for the former case and between c and $+\infty$ for the latter case. Fortunately, it is not necessary to integrate equation 2-2 each time a probability estimate is needed. The integration of equation 2-2 for the standard normal distribution ($\mu = 0$, $\sigma = 1$) has been accomplished and is available in tabular form in most statistics texts. The function which shows the relationship between the random variable, X, and probability of occurrence (area under the PDF) is called the <u>cumulative distribution function</u> (CDF). The value of the CDF for the normal distribution approaches zero as X approaches $-\infty$ and approaches one as X approaches $+\infty$. The relationship between the PDF and CDF for the standard normal distribution is shown on fig. 2.03. This figure shows that the area to the left of point a is 0.1 and the area to the left of point b is 0.50. The area between a and b (and consequently the probability of occurrence of values of X between a and b) is 0.50 minus 0.1 or 0.4. Likewise, the probability of occurrence of a value of X greater than c can be calculated because the area to the left of point c is 0.95 and the area to the left of $+\infty$ is, by definition, 1.0. Therefore, the area between c and $+\infty$ (and the probability) can be calculated as 1.0 - 0.95 or 0.05. Through use of CDF curves (or tables) for the standard normal distribution, probability estimates can be made for any interval along the X-axis, and through conversion by equation 2-3, any normal distribution can be expressed in terms of the standard normal distribution, and consequently probability estimates can be developed for any normal distribution. In most applications, engineers work with the CDF rather than with the PDF, and in most instances, the work involves the graphical form of the CDF.

<u>Section 2.03. Probability and Frequency</u>

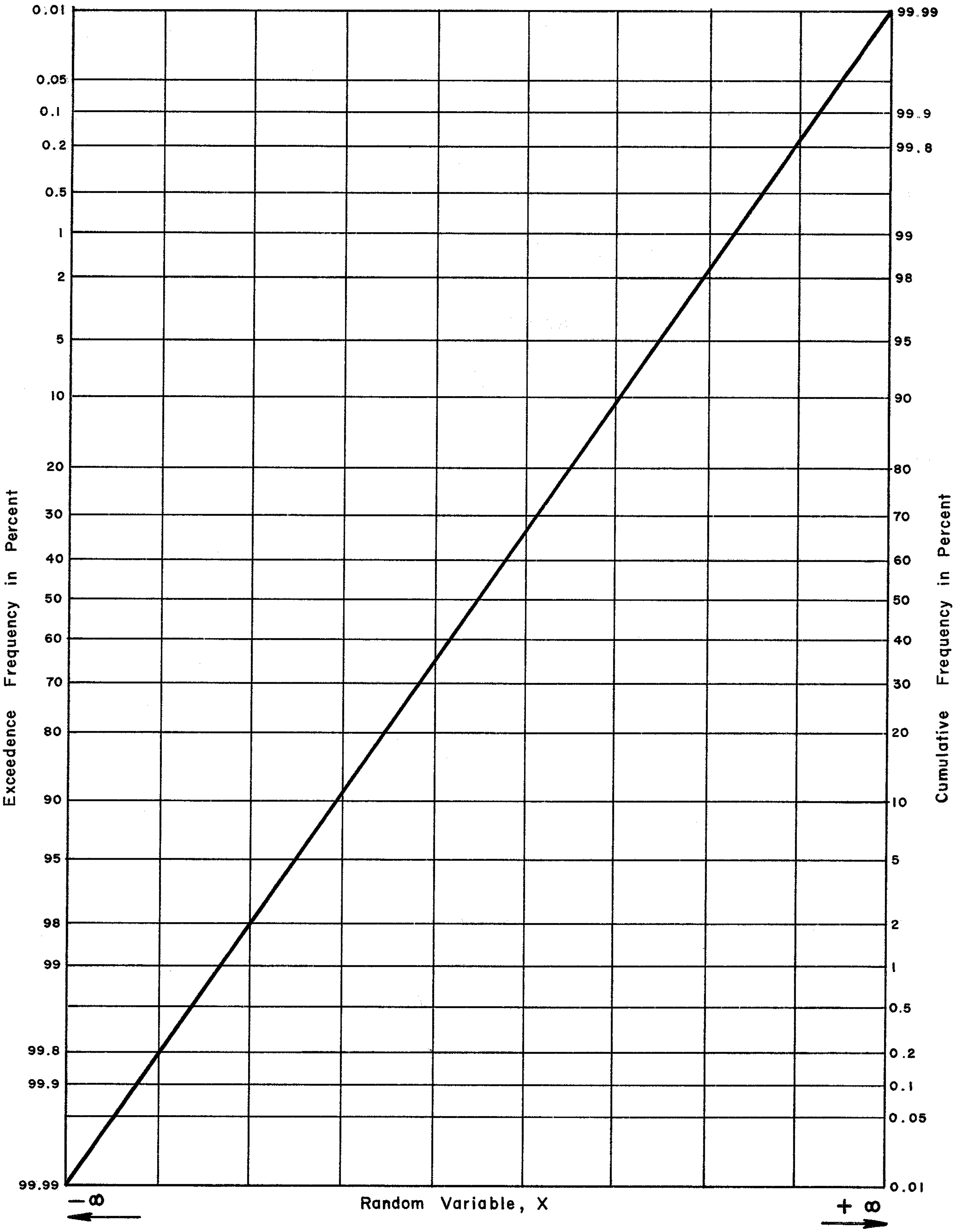
In many practical applications of probability analysis it is desirable to work with <u>frequency</u> of occurrence rather than probability of occurrence.

Frequency of occurrence, which is usually expressed as a percent value, is related to a specified number of occurrences and is directly related to the probability of occurrence. For example, if the probability that a magnitude, X, will be exceeded is .10, it can be stated that there is a 10 percent chance that X will be exceeded for any given future occurrence; or that X will be exceeded, on the average, once in every 10 occurrences; or that X will be exceeded, on the average, 10 times in 100 occurrences. It should be noted that this does not imply that X will be exceeded exactly once in 10 consecutive occurrences, but rather that in a very large number of occurrences (as implied by the phrase "on the average") the ratio of the number of magnitudes greater than X to the total number of events is 1/10. Some sets of 10 consecutive occurrences will have no occurrences greater than X, and other sets will have more than one occurrence greater than X.

In hydrologic analysis, one is frequently concerned with events that are identified with occurrence in the time span of a year, such as the maximum peak discharge in a year or the maximum 6-hour rainfall in a year. For such events, only one event out of an entire year of data (the maximum in the cases cited) is of interest and the event is labeled as the annual event, e.g., the annual maximum peak discharge or the annual maximum 6-hour rainfall. Probability analyses of annual events require that one event be selected out of each year of data, regardless of how many events occurred during the year. That is, if six large floods occur in a single year, and one is analyzing annual maximum floods, only the largest flood is used in the analysis even though some of the remaining five floods are larger than the maximum floods that have occurred in other years. Consequently, there is only one "occurrence" per year, and the total number of occurrences is equal to the number of years of data. Thus, number of years is often substituted for number of occurrences, and engineers speak of events with "10-year exceedence interval" or "50-year exceedence interval" with the understanding that the magnitude of the event with, for example, a 10-year exceedence interval will be exceeded, on the average, 1 year in 10 years or that there is a 10 percent chance that this magnitude will be exceeded in any future year.

The curves on fig. 2.03 are plotted on arithmetic scales. When the CDF for the normal distribution is plotted on arithmetic scales, it takes on an S-shape as shown in fig. 2.03. It is possible, however, to develop a "probability scale" for the vertical axis so that the CDF for a normal distribution will plot as a straight line. Figure 2.04 shows the CDF from fig. 2.03 plotted on probability paper. It should be noted that the probability scales on fig. 2.04 are expressed as percent values rather than as ratios, as on fig. 2.03. Thus, the curve is known as a frequency curve (i.e., frequency per hundred events versus magnitude rather than probability versus magnitude). Also, it should be noted that the complement of the cumulative frequency (i.e., 100 minus the cumulative frequency) is called the exceedence frequency. Many times in engineering problems, the probability estimate needed is the probability that a given magnitude on the X-axis will be exceeded, and this is the exceedence probability or exceedence frequency. Thus, a value with an exceedence frequency of 1 percent (exceedence probability of .01) will be exceeded, on the average, once in each 100 observations. Furthermore, for any given future occurrence (say, next year's annual maximum peak discharge) there is one chance in 100 that the observed magnitude will exceed the magnitude which has an exceedence frequency of 1 percent.

There are many probability distributions other than the normal distribution that are useful in various areas of science and engineering. Some of these distributions are completely specified by one or two parameters, others may require three or more parameters for complete specification. However, many physical and natural phenomena seem to be normally (or log-normally) distributed, and unless the sample data are extensive and clearly not normally distributed, it is usually



Cumulative distribution function for standard normal distribution on probability paper (Frequency curve)

Figure 2.04

desirable to assume that the population is normally distributed. Some discussion of non-normal distributions for hydrologic engineering are contained in subsequent chapters of this report.

Once a distribution form for the population has been selected, it is only necessary to specify the parameters required for the distribution (μ and σ for the normal distribution) in order to completely define the distribution. However, the parameters are unknown and must be estimated from sample data using techniques described in the following section.

Section 2.04. Sampling and Statistics

As indicated in previous sections of this chapter, the true charactertistics of the parent population of most natural phenomena are seldom known; that is, the distribution of the parent population is unknown. In cases where nothing is known about the distribution of the population, it may be possible to analyze samples to gain information about the population. For example, a manufacturer who has produced a million machined bolts to supply a contract may know that human errors and maladjusted machines will cause some of the bolts to be defective (that is, unacceptably larger or smaller than called for in the contract). Typically, a contract would specify that a certain percentage of the bolts could be defective by a certain amount without causing the shipment to be rejected. If the penalties for rejection are severe the manufacturer may measure each of the million bolts to insure that he complies with the contract. More commonly, however, it is advantageous to sample the population of one million bolts to determine whether the lot should be rejected. Say the manufacturer decides to base his decision on accepting or rejecting the whole lot by examining a sample of 10,000 bolts selected at random from the shipment. If he knows nothing at all about how his equipment and personnel perform in producing bolts he might assume that the sample results fully describe the distribution of defective bolts in the lot. He would then compute relative frequencies for various magnitudes of deviations from the specified size using equation 2-1 and draw a relative frequency curve similar to fig. 2.01(a). This decision to accept or reject the lot would be based on the comparison of the results (as exhibited by the curve) with the contract terms. If, for any reason the manufacturer believes that the results of the analysis are incorrect (that is unrepresentative), he simply increases the size of his sample, or takes another sample to see if the results will change.

If the manufacturer has some knowledge about his equipment and personnel (say, from analyses of similar lots of bolts in the past) he may know that a certain type of distribution is always characteristic of his bolt production efforts and therefore it is not necessary to compute relative frequencies and draw the curves to describe the distribution. Instead he can compute from the sample measurements a few indicators that tell him whether the lot is acceptable. As will be shown later in this section, this is often the case.

Through analysis of the frequency of occurrence of various outcomes in a sample, it is possible to infer the characteristics of the population frequency distribution. The reliability of the population characteristics inferred from analysis of a sample is directly dependent upon the extent to which the sample is representative of the population which is, in turn, dependent to a large extent on the size of the sample. For example, if a population were defined as the set of all annual peak discharges at a particular location on a particular stream, that population would include all annual peak discharges that have ever occurred at that location and all annual peak discharges that will occur at that location in the future. It is obvious that a historical record of annual peak discharges--even a record 100 years in length--is a relatively small sample from this population. However, our knowledge of meteorology, hydrology and other sciences that are important in understanding the natural process that produces streamflow makes us aware that a wide variety of combinations of physical and natural phenomena will probably be experienced during a 100-year period and that inferences we make about the population from the 100-year sample should be relatively reliable.

Inferences based on a short record--say, 5 years in length--will, on the other hand, be subject to considerable question. The same knowledge that tells us that a 100-year record should be relatively useful tells us that protracted wet periods or dry periods can easily exceed 5 years in duration. Consequently, we are not likely to observe a representative variety of combinations of physical and natural phenomena during a short period, since conditions during this period might easily reflect only the short-term influences that existed during the period. Unfortunately, in hydrologic engineering the "true" values of population characteristics are almost never known, so there is no way to determine whether a given sample is representative of the population other than to rely on the general principle that the degree of representativeness generally increases as the size of sample (i.e., length of record) increases. Since hydrologic engineering evaluations and the decisions based on these evaluations often cannot be delayed so that a longer record can be obtained, it is frequently necessary to augment the available data at a given location through consideration of data obtained at other locations in the same region. Techniques for this type of analyses are discussed in Chapter 4 of this volume and in Chapter 4 of Volume 2.

In hydrologic engineering problems, the sample is used primarily to draw inferences about the relative frequency of various events within a population. It should be obvious that a sample would have to be very large before one could expect infrequent extreme events to be represented within the sample.

There are some problems involved in making probability estimates from relative frequency data developed from a sample. By the very nature of the computation used to calculate relative frequency, the relative frequency and therefore the probability of occurrence of a value larger than the largest value in the sample or a value smaller than the smallest value in a sample is zero. It should be clear that this cannot always be true because a sample will not always contain the largest and smallest values from the population. Since it is known that events more extreme than those contained in a sample can occur, there must be some formalism for estimating the probability of occurrence of events larger or smaller than those in the sample. That is, there must be some way of inferring, from the information exhibited by the sample, enough about the population from which the sample was drawn to permit one to estimate the probability of occurrence of events more extreme than those in the sample.

It is possible to infer the characteristics of a population from the frequency characteristics of a sample drawn from the population if something is known (or can be assumed) about the general nature (form) of the population's frequency distribution.

As indicated earlier in this chapter, many natural phenomena seem to exhibit characteristics which indicate that the normal distribution satisfactorily describes the frequency distribution of outcomes in the parent population. Therefore many statistical analyses of natural phenomena are based on the assumption that the events comprising the parent population are normally distributed. While this is a common assumption, it is not one which should be made without due consideration, because the use of an inappropriate distribution form can produce analyses that grossly distort the significance of sample data.

There are techniques for assessing the validity of the assumption of a particular distribution. Most of these techniques require statistical analyses that are beyond the scope of this volume, but are adequately described in statistics texts. Even these techniques do not produce definitive information about the validity of an assumed distribution. Instead, they provide probabilistic estimates of the chance that the assumed distribution is incorrect. A less quantitative assessment of the reasonableness of an assumed distribution is obtained by graphical comparison of observed data with computed data based on the assumed distribution. Techniques for such comparisons are described in Chapter 4 of this volume.

A major consideration in the selection of a distribution for use in statistical studies is whether known characteristics of the natural phenomena under study are consistent with the known characteristics of the distribution. For example, as indicated in the previous section the bounds of the normal distribution are minus infinity and plus infinity; and therefore one cannot expect that this distribution would satisfactorily represent the distribution of events from a natural population with different bounds (for example, phenomena with definite, fixed upper or lower bounds). In some cases it is possible to transform the observed data on natural phenomena that does not conform to known characteristics of a distribution in order to facilitate the use of the distribution. Streamflow data, for example, would not seem to be adequately described by the normal distribution because the lower bound of streamflow is zero rather than minus infinity. However, if the logarithms of streamflow are used in statistical analyses, rather than the streamflows themselves, the resulting data may conform to the known characteristics of the normal distribution (as streamflow approaches zero its logarithm approaches minus infinity). Characteristics of common statistical distributions can be found in most statistics texts, and these characteristics should be studied thoroughly before a decision is made to use any unfamiliar distribution.

In addition to information on the nature of various distributions, statistics texts provide information on the parameters needed to completely specify (in a mathematical sense) the distribution. The normal distribution, for example, can be completely specified by only two parameters: μ, the population mean, and σ^2, the population variance. If μ and σ could be determined for a given population it would be possible to compute precisely the probability of occurrence of events within any range of magnitude associated with the given population. Of course, we cannot determine μ and σ in populations of natural phenomena, because the determination requires complete knowledge of all events comprising the population, and our knowledge is limited to a relatively few recorded events. Consequently, just as the bolt manufacturer estimates characteristics of his population of a million bolts by analyzing a sample of 10,000 bolts, we estimate parameters of a population by analyzing samples (recorded data).

In the case of the normal distribution the sample mean, generally represented by the symbol $\overline{X}$, is computed from the sample data and used as an <u>estimator</u> for the population mean, μ. Similarly, the sample standard deviation, generally represented by the symbol S, is used as an estimator for the population standard deviation, σ. That is, the sample mean is assumed to be equal to the population mean, and the sample standard deviation is assumed to be equal to the population standard deviation. If the sample is representative of the population, these assumptions will be valid, and the probability estimates based on the distribution with $\mu = \overline{X}$ and $\sigma = S$ will be satisfactory. Equations for calculating $\overline{X}$ and S and techniques for comparing the sample data with the assumed distribution based on $\mu = \overline{X}$ and $\sigma = S$ will be discussed in Chapter 4. This comparison is the basis for ascertaining whether the assumed distribution seems to be suitable for probability estimates.

An <u>estimator</u> of a parameter is a <u>statistic</u> computed from sample data and used in lieu of an unavailable parameter to specify a statistical distribution. In general there are many ways of computing statistics from sample data for use as parameter estimators, but statistically-valid estimators have several specific mathematical and statistical properties. Although discussion of the properties is beyond the scope of this text, it should be noted that all meaningful estimators have a property that is of importance in hydrologic studies. This property, consistency, requires that as the size of sample increases the value of the estimator converges on the true value of the parameter being estimated. Because of this property it can be expected that the larger samples (e.g., longer hydrologic records) will produce better definition of the population distribution.

Throughout the remainder of this volume the computational techniques described are consistent with the principles of probabilistic analysis and sampling described in this chapter. In most cases rigorous mathematical and statistical proofs have been omitted. Questions regarding the applicability of techniques to situations which differ significantly from the general principles discussed in this volume or from those conditions specified as being pertinent should be resolved through careful study of relevant material in statistical texts.

Chapter 3

Principles of Probability Analysis in Hydrologic Engineering

CHAPTER 3. PRINCIPLES OF PROBABILITY ANALYSIS IN HYDROLOGIC ENGINEERING

Section 3.01. Introduction

Frequency curves are most commonly used in studies to determine the economic value of flood control projects. Other common uses of frequency curves include the determination of reservoir stage for real-estate acquisition and reservoir-use purposes, the selection of rainfall magnitude for storm-drain design, and the selection of runoff magnitude for interior drainage, pumping-plant, and local-protection project design.

There are two basic types of frequency curves used in hydrologic work. A curve of annual maximum events is ordinarily used when the primary interest lies in the very large events or when the second largest event in any year is of minor concern in the analysis. The partial-duration curve represents the frequency of all events above a given base value, regardless of whether two or more occurred in the same year. This type of curve is ordinarily used in economic analysis, because often there are substantial damages which result from the second largest and third largest floods in extremely wet years and from floods occurring more frequently than the annual event in rural areas. When both the frequency curve of annual floods and the partial-duration curve are used care must be exercised to assure that the two are consistent. A graphic demonstration of the relation between a chronologic record, an annual-event array, and a partial-duration array is shown on fig. 3.01.

In almost all locations there are seasons during which storms or floods do not occur or are not severe, and other seasons when they are more severe. Also, damages associated with a flood often vary with season of the year. In many types of studies, the seasonal variation factor is of primary importance, and it becomes necessary to establish frequency curves for each month or other subdivision of the year.

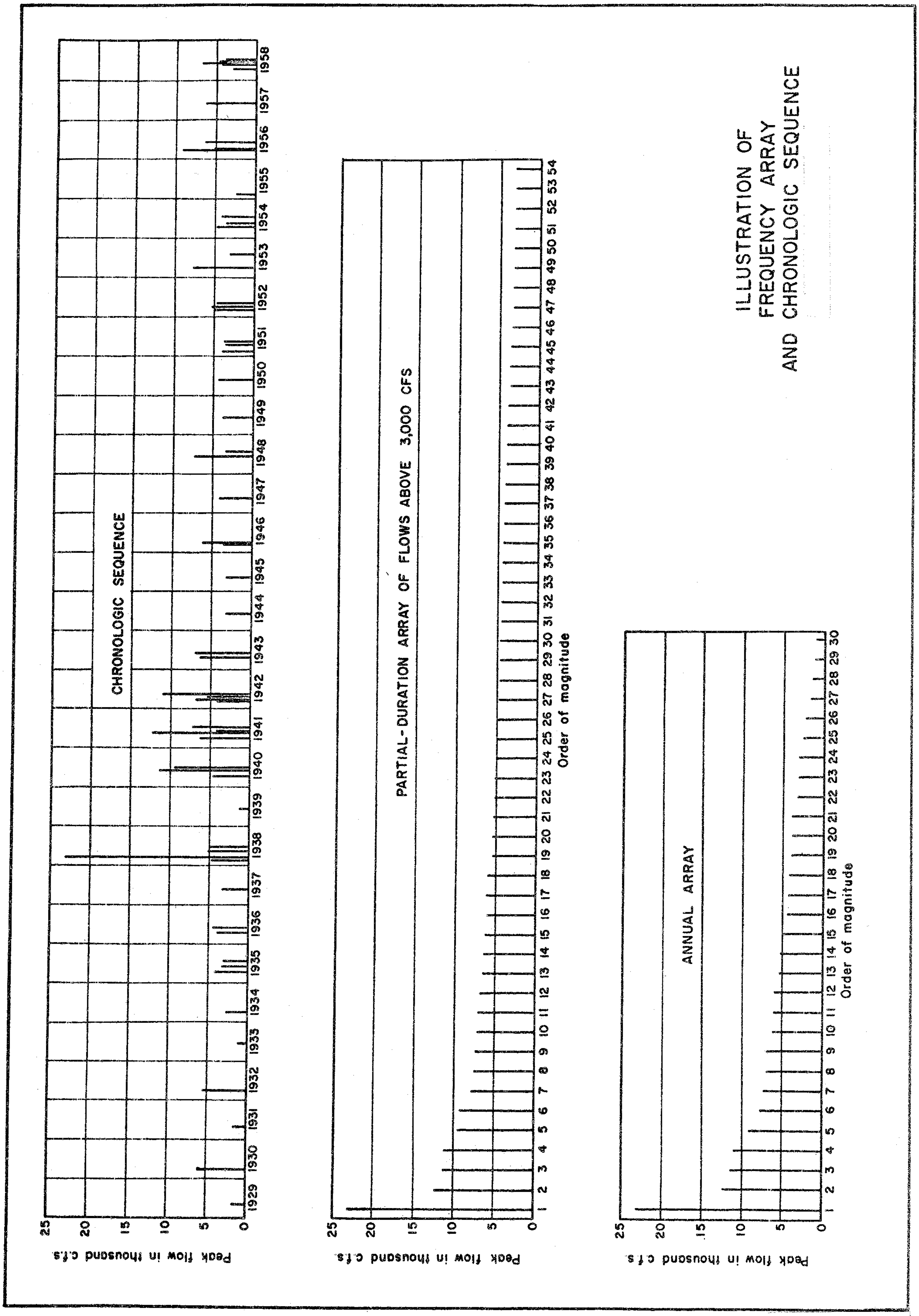


ILLUSTRATION OF
FREQUENCY ARRAY
AND CHRONOLOGIC SEQUENCE

Figure 3.01

For example, one frequency curve might represent the largest floods that occur each January, a second one would represent the largest floods that occur each February, etc. In another case, one frequency curve might represent floods during the snowmelt season, while a second might represent floods during the rain season. Occasionally, when seasons are studied separately, an annual-event curve covering all seasons is also prepared, and care should be exercised to assure that the various seasonal curves are consistent with the annual curve.

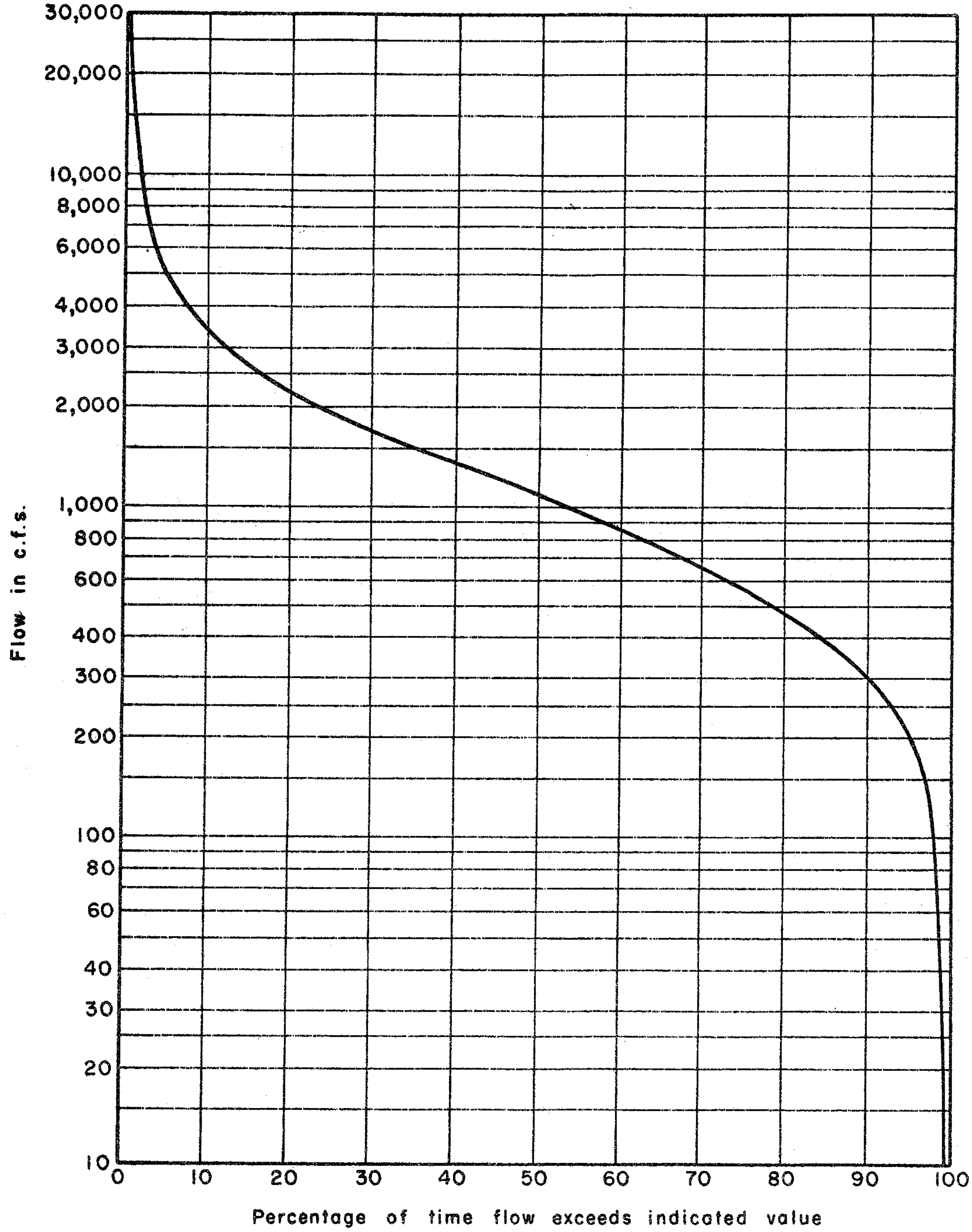
Section 3.02. Duration Curves

In connection with power studies for run-of-river plants particularly, and in some phases of sediment studies, the flow-duration curve serves a useful purpose. It simply represents the percent of time during which specified flow rates are exceeded at a given location. Ordinarily, variations within periods less than 1 day are not of consequence, and the curves are therefore based on observed mean-daily flows. For the purpose served by flow duration curves, the extreme rates of flow are not important, and consequently there is no need for refining the curve in regions of high flow.

The procedure ordinarily used in the preparation of a flow-duration curve consists of counting the number of mean-daily flows that occur within given ranges of magnitude. Then the lower limit of magnitude in each range is plotted against the percentage of days of record that mean-daily flows exceed that magnitude. A typical flow duration curve is shown on fig. 3.02.

Section 3.03. Technical Approach in Estimating Frequency Curves

There are two basic approaches to estimating frequency curves--graphical and analytical. Each of these approaches has several variations in current practice, but the discussion herein will be limited to certain recommended methods.



ILLUSTRATIVE EXAMPLE

FLOW DURATION CURVE

RAPPAHANNOCK RIVER AT FREDERICKSBURG, VIRGINIA

Frequencies are evaluated graphically by arranging observed values in the order of magnitude and representing frequencies by a smooth curve through the array of values. Each observed value represents a fraction of the future possibilities and, when plotting the frequency curve, it is given a "plotting position" that is calculated to give it the proper weight.

In the application of analytical (statistical) procedures, the concept of theoretical populations or distributions is employed. The events that have occurred are presumed to constitute a random sample and accordingly are used to make inferences regarding their "parent population" (i.e., the distribution from which they were derived). Such inferences are necessarily attended by considerable uncertainty, because a given set of observations could result from any of many sets of physical conditions (from any one of many distributions). However, by the use of statistical processes, the most probable nature of the distribution from which the data were derived can be estimated. Since this in all probability is not the true parent population, the relative chance that variations from this distribution might be true must be evaluated. Each range of possible parent population is then weighted in proportion to its likelihood in order to obtain a weighted average. A probability obtained from this weighted average is herein referred to as the expected probability P_N.

Because of the shortness of hydrologic records, frequency determinations for rare events are relatively unreliable where based on a single record. Also, it is often necessary to estimate frequencies for locations where no record exists. For these reasons, regionalized frequency studies, in which frequency characteristics are related to drainage-basin features, are desirable. Regionalized frequency studies are facilitated by the use of analytical methods, as illustrated in Chapter 4.

Section 3.04. Effect of Basin Developments on Frequency Relations

Most hydrologic frequency estimates serve some purpose relating to planning, design or operation of water resources management projects. The anticipated effects of a project on flooding can be assessed by comparing the peak discharge and volume frequency curves with and without the project. Also, projects that have existed in the past have affected the rates and volumes of floods, and the recorded values must be adjusted to reflect uniform conditions in order that the frequency analysis will conform to the basic assumptions of randomness and commom population. In order for a frequency curve to conform reasonably with a generalized mathematical or probability law, the flows must be essentially unregulated by man-made storage or diversion structures. Consequently, wherever practicable, recorded runoff values should be adjusted to natural conditions before a frequency analysis is made. However, in cases where the regulation results from a multitude of relatively small management works that have not changed appreciably during the period of record, it is likely that the general mathematical laws will apply as in the case of natural flows, and that adjustment to natural conditions would be unnecessary and, because of the amount of work involved, undesirable.

If it is practical, the most complete approach to determining frequency curves of regulated runoff consists of (1) routing flows for the entire period of record through the proposed management works, (2) arranging the annual regulated flows in order of magnitude, and (3) plotting the frequency curve of regulated flows using selected plotting positions to determine frequencies of regulated runoff for extreme floods. A less involved method consists of routing of multiples of the largest floods of record or multiples of a large hypothetical flood to yield results approximating those that would be obtained by the more complete procedure. Techniques for estimating project effects are outlined in Chapter 6.

In general, cumulative frequency curves of river stages are determined from frequency curves of flow because of the advantage available in fitting flow frequency data with theoretical curves or because the effects of upstream storage or diversion can best be estimated in terms of flow. In cases where the stage-discharge relation is erratic, a frequency curve of stages can be derived directly from stage data, as illustrated in Chapter 4. Project construction or natural changes in streambed elevation may change the relationship between stage and flow at a location. By forming constrictions, levees may raise stages several feet for some distance upstream. Reservoir or channel work may cause changes in degradation or aggradation of streambeds, and thereby change rating curves. Thus, the effects of projects on river stages often involve the effects on channel hydraulics as well as the effects on streamflow.

Section 3.05. Selection of Frequency Data

The primary consideration in selection of an array of data for a frequency study is the objective of the frequency analysis. If the frequency curve that is developed is to be used for estimating damages that are related to instantaneous peak flows in a stream, peak flows should be selected from the record. If the damages are related to maximum mean-daily flows or to maximum 3-day flows, these items should be selected. If the behavior of a reservoir under investigation is related to the 3-day or 10-day rainflood volume, or to the seasonal snowmelt volume, that pertinent item should be selected. Occasionally, it is necessary to select a related variable in lieu of the one desired. For example, where mean-daily flow records are more complete than the records of peak flows, it may be desirable to derive a frequency curve of mean-daily flows and then, from the computed curve, derive a peak-flow curve by means of an empirical relation between mean daily flows and peak flows. All reasonably independent values should be selected, but the annual maximum events should ordinarily be segregated when the application of analytical procedures discussed in Chapter 4 is contemplated.

Data selected for a frequency study must measure the same aspect of each event (such as peak flow, mean-daily flow, or flood volume for a specified duration), and each event must be controlled by a uniform set of hydrologic and operational factors. For example, it would be improper to combine items from old records that are reported as peak flows but are in fact only daily readings, with newer records where the peak was actually measured. Similarly, care should be exercised when there has been significant change in upstream storage regulation during the period of record so as not to inadvertently combine unlike events into a single series. In such a case, the entire record should be adjusted to a base condition, preferably the unregulated condition.

Hydrologic factors and relationships operating during a winter rain flood are usually quite different from those operating during a spring snowmelt flood or during a local summer cloudburst flood. Where two or more types of floods are distinct and do not occur predominantly in mutual combinations, they should not be combined into a single series for frequency analysis. They should be considered as events from different parent populations. It is usually more reliable in such cases to segregate the data in accordance with type and to combine only the final curves, if necessary. For example, in the mountainous region of eastern California, frequency studies are made separately for rain floods, which occur principally during the months of November through March, and for snowmelt floods, which occur during the months of April through July. Flows for each of these two seasons are segregated strictly by cause--those predominantly caused by snowmelt and those predominantly caused by rain. In desert regions, summer thunderstorms should be excluded from frequency studies of winter rain flood or spring snowmelt floods and should be considered separately. Similarly, in coastal regions it would be desirable to separate floods induced by hurricanes or typhoons from other general flood events.

When practicable, all runoff data should be adjusted to natural hydrologic conditions before making the frequency study, because natural flows are better adapted to analytical methods and are more easily compared within a region. Frequency curves of "present-regulated" conditions (those prevailing under current practices of regulation and diversion) or of "future-regulated" conditions can be constructed from the frequency curve of natural flow by means of an empirical or logical relationship between natural and regulated flows. Where data recorded at two different locations are to be combined for construction of a single frequency curve, the data should be adjusted as necessary to a single location, usually the location of the longer record, accounting for differences of drainage area and precipitation and, where appropriate, channel characteristics between the locations. Where the stream-gage location is somewhat different from the project location, the frequency curve should be constructed for the stream-gage location and subsequently adjusted to the project location.

Occasionally a runoff record may be interrupted by a period of one or more years. If the interruption is caused by destruction of the gaging station by a large flood, failure to fill in the record for that flood would have a biasing effect, which should be avoided. However, if the cause of the interruption is known to be independent of flow magnitude, the entire period of interruption should be eliminated from the frequency array, since no bias would result. In cases where no runoff records are available on the stream concerned, it is possible to estimate the frequency curve as a whole using regional generalizations discussed in Chapter 4. An alternative method is to estimate a complete series of individual floods from recorded precipitation by continuous hydrologic simulation and perform conventional frequency analysis on the simulated record.

Section 3.06. Climatic Variations

Some hydrologic records suggest regular cyclic variations in precipitation and runoff potential, and many attempts have been made to demonstrate that precipitation or stream flows display variations that are in phase with various cycles, particularly the well-established 11-year sunspot cycle. There is no doubt that long-duration cycles or irregular climatic changes are associated with general changes of land masses and seas and with local changes in lakes and swamps. Also, large areas that have been known to be fertile in the past are now arid deserts, and large temperate regions have been covered with glaciers one or more times. Although the existence of climatic changes is not questioned, their effect is ordinarily neglected because long-term climatic changes generally have insignificant effects during the period concerned in water development projects, and short-term climatic changes tend to be self-compensating. For these reasons, and because of the difficulty in differentiating between fortuitous and systematic changes, it is considered that except for the annual cycle the effect of natural cycles or trends during the period of useful project life can ordinarily be neglected in hydrologic frequency studies.

Section 3.07. Frequency Reliability Analyses

The reliability of frequency estimates is influenced by:

a. The amount of information available.

b. The variability of the events.

c. The accuracy with which the data were measured.

In general with regard to item *a*, errors of estimate are inversely proportional to the square root of the number of independent items contained in the frequency array. Therefore, errors of estimates based on 40 years of record would normally be half as large as errors of estimates based on 10 years of record, other conditions being the same.

The variability of events in a record (item *b*) is usually the most important factor affecting the reliability of frequency estimates. For example, the ratio of the largest to the smallest annual flood of record on the Mississippi River at Red River Landing, Louisiana, is about 2.7, whereas the ratio of the largest to the smallest annual flood of record on the Kings River at Piedra, California, is about 100, or 35 times as great. Statistical studies show that as a consequence of this factor, a flow corresponding to a given frequency that can be estimated within 10 percent on the Mississippi River, can be estimated only within 40 percent on the Kings River.

The accuracy of data measurement (item *c*) normally has relatively small influence on the reliability of a frequency estimate, because such errors ordinarily are not systematic and tend to cancel, and because the influence of chance events is great in comparison with that of measurement errors. For this reason, it is usually better to include an estimated magnitude for a major flood; for example, that was not recorded because of gage failure, rather than to omit it from the frequency array, even though its magnitude can only be estimated approximately. However, it is advisable always to use the most reliable sources of data and, in particular, to guard against systematic errors such as result from using an unreliable rating curve.

It should be remembered that the possible errors in estimating flood frequencies are very large, principally because of the chance of having a non-representative sample. Sometimes the occurrence of one or two abnormal floods can change the apparent exceedence frequency of a given magnitude from once in 1,000 years to once in 200 years. Nevertheless, the frequency-curve technique is considerably better than any other tool available for some purposes and represents a substantial improvement over using an array restricted to observed flows only. Reliability criteria useful for illustrating the accuracy of frequency determinations and for applying safety factors in design are described in Chapter 4.

Section 3.08. Presentation of Data and Results of Frequency Analyses

Information provided with frequency curves should indicate clearly the scope of the studies and include a brief description of the procedure used, including appropriate references. When rough estimates are adequate or necessary, the frequency data should be properly qualified in order to avoid misleading conclusions that might seriously affect the project plan. Summary of the basic data consisting of chronological tabulation of values used and indicating sources of data and adjustments made, would be helpful. The frequency data can also advantageously be presented in graphical form, ordinarily on probability paper, along with the adopted frequency curves.

Chapter 4

Computation of Frequency Curves

CHAPTER 4. COMPUTATION OF FREQUENCY CURVES

Section 4.01. Computation by Graphical Method

Every derived frequency relation should be plotted graphically, even though the results can be obtained entirely analytically as described in Section 3.02, in order that observed data may be visually compared with the derived curve. The visual comparison is the simplest method of evaluating whether or not the assumed distribution is consistent with the observed data. The graphical method of frequency-curve determination can be used for any type of frequency study, but analytical methods have certain advantages where they are applicable (see Section 4.02). The principal advantages of graphical methods are that they are generally applicable, that the derived curve can be easily visualized, and that the observed data can be readily compared with the computed results. However, graphical methods of frequency analysis are inferior in accuracy to analytical methods where the latter apply, and do not provide means of evaluating the reliability of the estimates. Comparison of the adopted curve with plotted points is not an absolute index of reliability as in correlation analysis, but it is often erroneously assumed to be, thus implying a much greater reliability than is actually attained. For these reasons, graphical methods should be limited to those cases where analytical methods do not apply (that is, where frequency curves are too irregular to compute analytically) and to use as a visual aid or check on analytical computations.

General principles in the selection of frequency data are discussed in the previous chapter. Data used in the construction of frequency curves of peak flows consist of the maximum flow for each year of record (for annual-event curve) and all of the secondary flows that exceed a selected base value (for partial-duration curve). This base value must be smaller than any flood flow that is of importance in the analysis, and should also be low enough so that the total number of floods in excess of the base equals or exceeds the number of years of record. Ordinarily, the latter criterion controls, and as a matter of convenience the annual and partial-duration series of events are often made equal in number. Figure 4.01 is a sample tabulation of data in the order of magnitude directly from the record, and of corresponding plotting positions.

Recommended plotting positions are tabulated on Exhibit 1. In ordinary hydrologic frequency work, N is taken as the number of years of record rather than the number of events, so that exceedence frequencies will be in terms of events per hundred years. For arrays larger than 100, the plotting position of the largest event, P_1, is obtained by use of the following equation:

$$1 - P_1 = (0.5)^{1/N} \tag{4-1}$$

The plotting position for the smallest event is the complement of this value, and all other plotting positions are interpolated linearly between these two. For partial-duration curves, particularly where there are more events than years (N), plotting positions larger than 50 percent are obtained by use of the following equation:

$$P = (2m - 1)/2N \tag{4-1a}$$

in which m is the order number of the event (that is, for the largest event m = 1, for the second largest m = 2, etc.).

The plotting grid commonly used in studies of flood flow frequencies is the logarithmic normal (logarithmic probability) grid which is designed such that the logarithmic normal frequency distribution will plot as a straight line. A plotting grid that covers the usual range of flows and frequencies concerned in hydrologic studies and designed for

ILLUSTRATIVE EXAMPLE

TABULATION OF PEAK FLOW FREQUENCY DATA
MILL CREEK NEAR LOS MOLINOS, CALIFORNIA

Events in order recorded: Annual maximum, and all other flows above 3000 c.f.s. and separated by 10 days or more from larger flows						Events in decreasing order of magnitude: Annual Maximum		Partial-duration series (flows above 3000 c.f.s.)			
Water Year (1)	Date (2)	Flow (cfs) (3)	Water Year (4)	Date (5)	Flow (cfs) (6)	Plotting Position (7)	Flow (cfs) (8)	Plotting Position (9)	Flow (cfs) (10)	Plotting Position (11)	Flow (cfs) (12)
1928-29	3 Feb	1,520	1943-44	4 Mar	3,220	2.3	23,000	2.3	23,000	91.7	4,650
1929-30	15 Dec	6,000	1944-45	5 Feb	3,230	5.6	12,200	5.6	12,200	95.0	4,600
1930-31	23 Jan	1,500	1945-46	4 Dec	3,660	8.9	11,400	8.9	11,400	98.3	4,540
1931-32	24 Dec	5,440	1945-46	21 Dec	6,180	12.2	11,000	12.2	11,000	101.7	4,430
1932-33	16 Mar	1,080	1946-47	12 Feb	4,070	15.4	9,180	15.4	9,360	105.0	4,380
1933-34	29 Dec	2,630	1947-48	23 Mar	7,320	18.7	7,710	18.7	9,180	108.3	4,330
1934-35	4 Jan	4,010	1947-48	28 Apr	3,380	22.0	7,320	22.0	7,710	111.7	4,250
1934-35	28 Feb	3,190	1948-49	11 Mar	3,870	25.3	6,970	25.3	7,320	115.0	4,240
1934-35	8 Apr	3,040	1949-50	4 Feb	4,430	28.6	6,880	28.6	7,260	118.3	4,130
1935-36	11 Jan	3,930	1950-51	16 Nov	3,870	31.9	6,180	31.9	6,970	121.7	4,070
1935-36	21 Feb	4,380	1950-51	22 Jan	3,510	35.2	6,140	35.2	6,910	125.0	4,010
1936-37	14 Feb	3,310	1950-51	11 Feb	3,660	38.5	6,000	38.5	6,880	128.3	3,970
1937-38	20 Nov	4,700	1951-52	1 Dec	4,930	41.8	5,440	41.8	6,480	131.7	3,930
1937-38	11 Dec	23,000	1951-52	26 Dec	5,280	45.1	5,280	45.1	6,450	135.0	3,870
1937-38	2 Feb	5,050	1951-52	1 Feb	4,650	48.4	4,910	48.4	6,240	138.3	3,870
1937-38	23 Mar	4,950	1952-53	9 Jan	7,710	51.6	4,430	51.7	6,180	141.7	3,660
1938-39	8 Mar	1,260	1952-53	27 Apr	3,070	54.9	4,380	55.0	6,140	145.0	3,660
1939-40	2 Jan	4,600	1953-54	17 Jan	4,910	58.2	4,070	58.3	6,000	148.3	3,510
1939-40	28 Feb	11,400	1953-54	13 Feb	3,300	61.5	4,010	61.7	5,450	151.7	3,380
1939-40	30 Mar	9,360	1953-54	4 Apr	4,240	64.8	3,870	65.0	5,440	155.0	3,310
1940-41	24 Dec	6,240	1954-55	11 Nov	2,480	68.1	3,870	68.3	5,280	158.3	3,300
1940-41	10 Feb	12,200	1955-56	22 Dec	9,180	71.4	3,310	71.7	5,050	161.7	3,230
1940-41	1 Mar	4,250	1955-56	7 Jan	5,020	74.7	3,230	75.0	5,020	165.0	3,220
1940-41	4 Apr	7,260	1955-56	22 Feb	6,480	78.0	3,220	78.3	4,950	168.3	3,190
1941-42	3 Dec	4,130	1956-57	24 Feb	6,140	81.3	2,630	81.7	4,930	171.7	3,070
1941-42	16 Dec	6,910	1957-58	26 Jan	3,060	84.6	2,480	85.0	4,910	175.0	3,060
1941-42	27 Jan	5,450	1957-58	12 Feb	4,330	87.8	1,520	88.3	4,700	178.3	3,040
1941-42	6 Feb	11,000	1957-58	24 Feb	6,880	91.1	1,500				
1942-43	21 Jan	6,450	1957-58	21 Mar	4,540	94.4	1,260				
1942-43	8 Mar	6,970	1957-58	2 Apr	3,970	97.7	1,080				

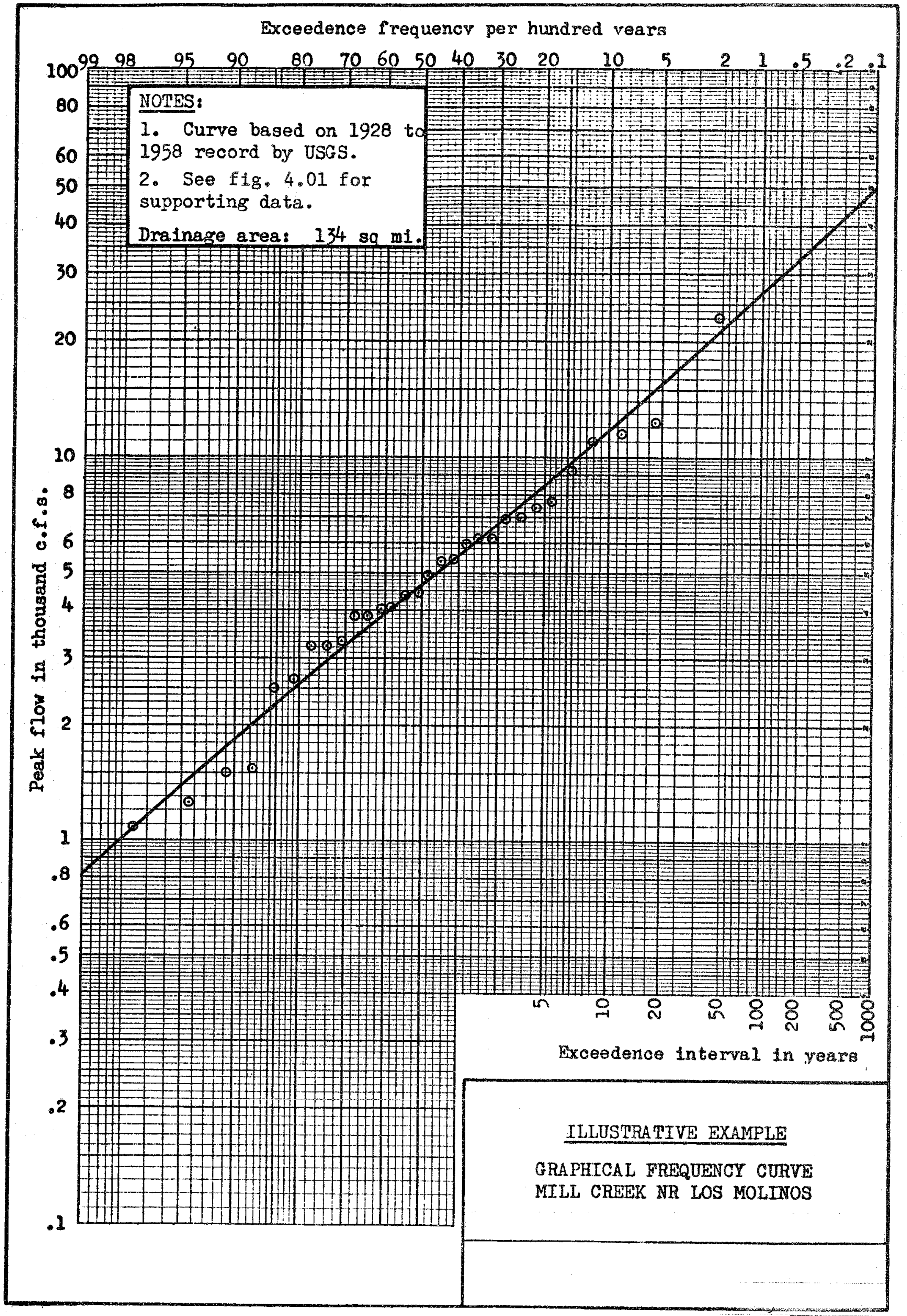
Figure 4.01

presentation on letter-size charts is illustrated on fig. 4.02. The plotting grid used for flood stage frequencies is the arithmetic probability (or arithmetic normal) grid illustrated on fig. 4.03. The plotting grids should contain a horizontal scale of exceedence interval as well as a horizontal scale of exceedence frequency and a vertical scale of flood flow.

Figure 4.02 shows the plotting of a frequency curve of annual peak flows corresponding to those tabulated on fig. 4.01. The curve should be drawn so as to balance out the plotted points to a reasonable degree and so that there is no abrupt break in the frequency curve. Unless computed as discussed in Section 4.02, the frequency curve should be drawn as a straight line on the grid whenever data and conditions warrant.

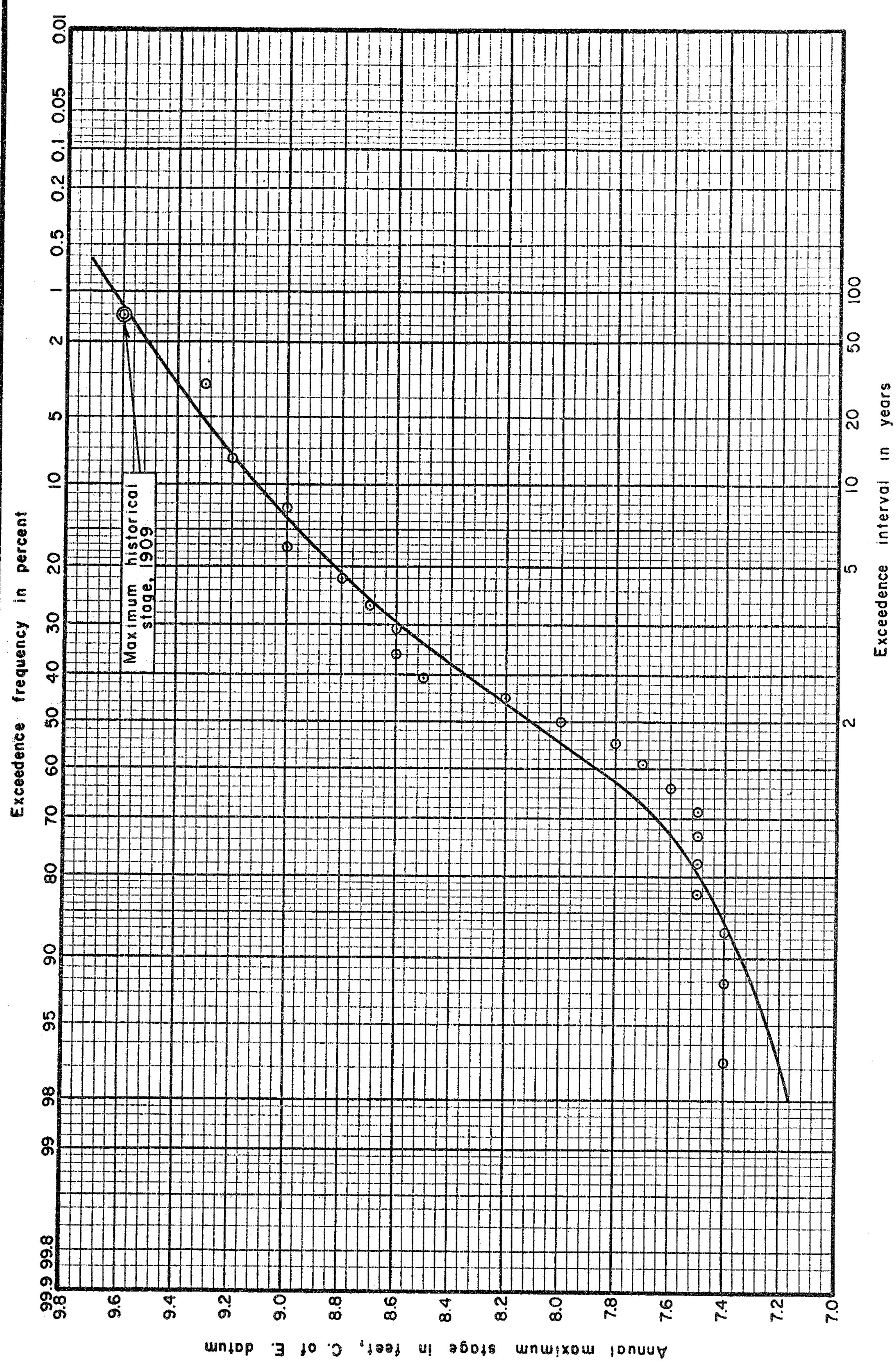
The partial-duration curve corresponding to the partial-duration data on fig. 4.01 has been shown on fig. 4.04. This curve has been drawn by generally balancing out the plotted points, except that it was made to conform with the annual-event curve in the upper portion in general accord with the standard relationship discussed in Section 4.02. When partial-duration data includes more events than there are years of record it will be necessary to use logarithmic paper for plotting purposes, as on fig. 4.04, in order to plot exceedence frequencies greater than 100 percent. Otherwise, the curve can be plotted on probability grid, as illustrated on fig. 5.06.

Figure 4.03 illustrates the graphical construction of a river stage frequency curve. The shape of this curve is dictated by the plotted points and, in some cases, by consideration of the stage at which overbank flows begin. Whenever stage is a consistent function of flow, as is the usual case, the stage frequency curve should be obtained from the flow frequency and stage-discharge curves.



ILLUSTRATIVE EXAMPLE

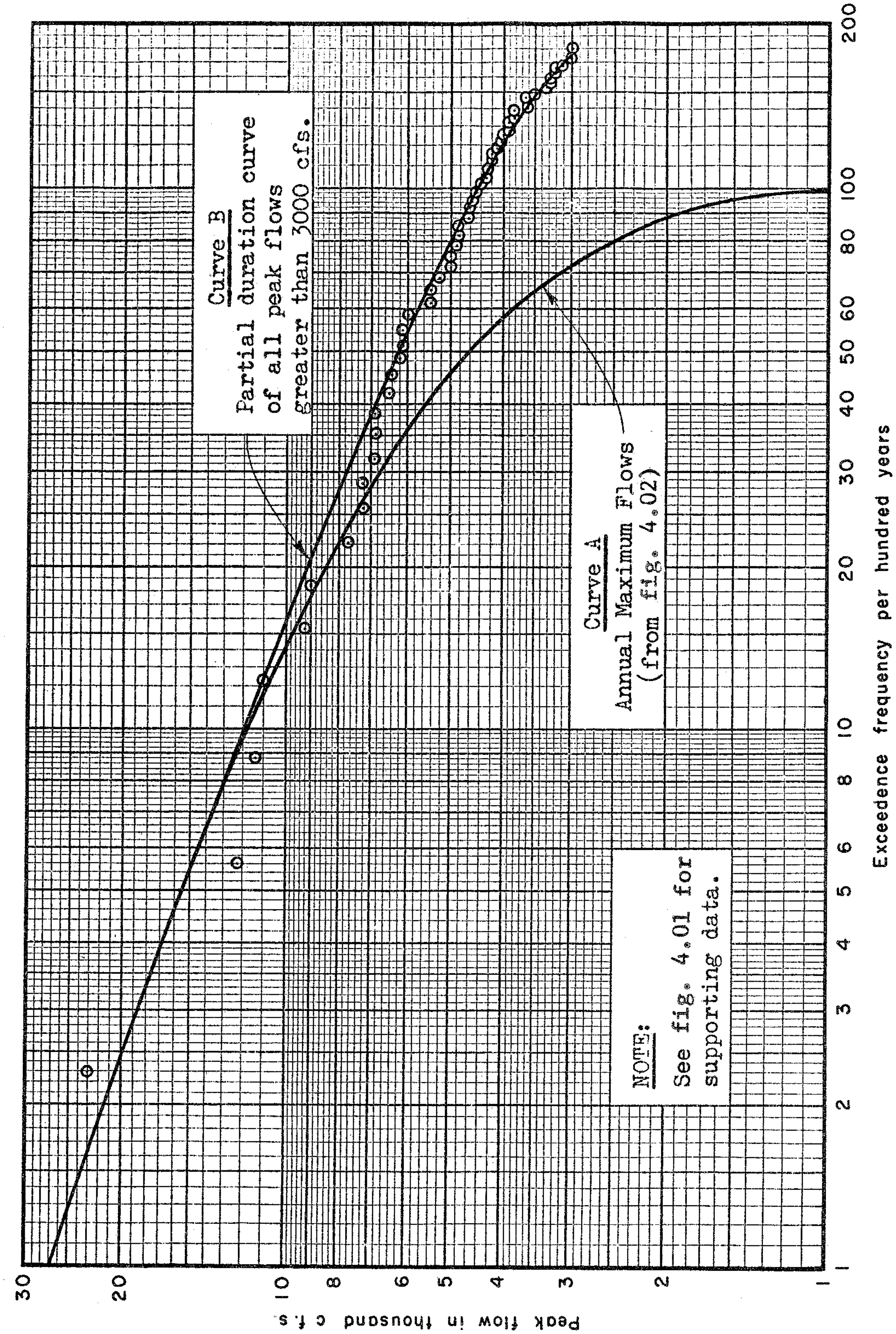
GRAPHICAL FREQUENCY CURVE
MILL CREEK NR LOS MOLINOS



ILLUSTRATIVE EXAMPLE

RIVER STAGE FREQUENCY CURVE, 1935-1956

Figure 4.03



ILLUSTRATIVE EXAMPLE

PARTIAL-DURATION FREQUENCY CURVE

Section 4.02. Computation by Analytical Method

The analytical method of computing a frequency curve is limited almost exclusively to curves of annual maximum stream flows for a specified duration (including peak flows) and annual maximum precipitation amounts for a specified duration. In general, the results obtained by analytical methods where they apply are considerably more reliable than those obtained by graphical procedures. They have the additional advantage that the degree of reliability of frequency estimates can be evaluated.

The most commonly used analytical frequency procedure for hydrologic investigations involving precipitation and natural runoff is the logarithmic Pearson type III distribution. This distribution requires three parameters for complete mathematical specification. The parameters are: the mean, or first moment, (estimated by the sample mean, $\overline{X}$); the variance, or second moment, (estimated by the sample variance, S^2); and the skew, or third moment, (estimated by the sample skew, g). Since the distribution is a logarithmic distribution, all parameters are estimated from logarithms of the observations, rather than from the observations themselves. The Pearson type III distribution is particularly useful for hydrologic investigations because the third parameter, the skew, permits the fitting of non-normal samples to the distribution. When the skew is zero the log-Pearson type III distribution becomes a two-parameter distribution that is identical to the logarithmic normal (often called log-normal) distribution.

The distribution is fitted to runoff and precipitation data by calculating the sample mean, variance, and skew from the following equations:

$$\overline{X} = \frac{\Sigma X}{N} \qquad (4\text{-}2)$$

$$S^2 = \frac{\Sigma x^2}{N-1} = \frac{\Sigma X^2-(\Sigma X)^2/N}{N-1} \qquad (4\text{-}3)$$

$$g = \frac{N\ (\Sigma x^3)}{(N-1)(N-2)S^3} = \frac{N^2(\Sigma X^3)-3N(\Sigma X)(\Sigma X^2)+2(\Sigma X)^3}{N\ (N-1)(N-2)S^3} \qquad (4\text{-}4)$$

in which:

X = logarithm of the magnitude of an event in the sample

$\overline{X}$ = the mean of the logarithms

x = $X-\overline{X}$, the deviation of a single event from the mean

N = the number of events in the sample

S^2 = unbiased estimate of the variance

g = unbiased estimate of the skew coefficient

Σ = perform calculation for all values from 1 to N

In terms of the frequency curve itself, the mean represents the general magnitude or average ordinate of the curve, the square root of the variance (the standard deviation, S) represents the slope of the curve, and the skew represents the degree of curvature of the curve.

The types of cumulative frequency curves fitted in hydrologic studies do not require moments of a higher order than these three, and frequently only the first two will be required. There is no need to tabulate the individual deviation for each logarithm, as the second forms of equations 4-3 and 4-4 can be used to compute the standard deviation and skew coefficient much more rapidly, directly from the original logarithms. In computing these quantities in this manner, however, it is essential that the intermediate quantities in the computation be accurate to at least four decimal places, which is not practicable without an automatic desk calculator or electronic computer. Computation of the unadjusted frequency curve is accomplished by computing values for the logarithms of the streamflow corresponding to selected points (exceedence frequencies) on the frequency scale. The points ordinarily selected are shown as column headings on Exhibit 2. The number of points needed to define the curve depends on the degree of curvature (i.e., the skew). For skew values near zero, fewer points will be needed, while for skew values near one (or greater than one) all of the points shown in Exhibit 2 should be used. The selected exceedence frequency points are labeled $P\infty$ in this volume--the subscript ∞ indicating an infinite sample size. Since the samples used in hydrologic studies are of finite size, an adjustment of the exceedence frequency, $P\infty$, is necessary before plotting the frequency.

The logarithms of the event magnitudes corresponding to each of the selected $P\infty$ points are computed by the following equation:

$$\log Q_{@P\infty} = \overline{X} + k_{@P\infty} S \qquad (4\text{-}5)$$

where $\overline{X}$ and S are defined as in equations 4-2 and 4-3 and where

$\log Q_{@P\infty}$ = the logarithm of the flow (or other variable) corresponding to a specified value of $P\infty$

$k_{@P\infty}$ = the deviation from the mean, $\frac{X-\overline{X}}{S}$, (in standard deviation units) of the event with an exceedence frequency of $P\infty$. Values of k are obtained from Exhibit 2 for each selected value of $P\infty$ required to define the curve

As shown in the following example, the equation 4-5 is solved by using the computed values of $\overline{X}$ and S and obtaining from Exhibit 2 the value of k corresponding to the computed skew, g, and the selected point $P\infty$.

It has been shown in reference 2 that a frequency curve computed in this manner is biased in relation to average future expectation because of the assumption regarding infinite sample size. It has also been shown that the effect of this bias can be eliminated by an adjustment of the $P\infty$ exceedence frequency values to an expected probability value (designated as P_N in this volume) that accounts for the actual sample size. The adjustment is accomplished by substituting a P_N value obtained from Exhibit 3 (based on $P\infty$ and N, the sample size) for the value of $P\infty$ before plotting the frequency curve. The sample size used to obtain P_N from Exhibit 3 should correspond to the equivalent length of record as determined by procedures described in Section 4.03.

Frequency curves of annual maximum or minimum events are computed as follows:

a. Mean Logarithm. After tabulation of the data in chronological order or in the order of magnitude, the logarithm of each discharge is tabulated to at least two decimal places. The mean logarithm ($\overline{X}$) is obtained by dividing the sum of these logarithms by the number of events (equation 4-2). Time can be saved by obtaining the sum of the logarithms on one register of a calculator at the same time that the sum of the squares of the logarithms are obtained on a second register for step b.

b. Standard Deviation. The standard deviation which is the square root of the variance (equation 4-3) is computed as follows:

(1) Obtain the sum of the squares of the logarithms in an automatic calculator. This quantity should not be rounded off, but all figures carried in the computation.

(2) The sum of the logarithms obtained in the same machine operation, which figure is also not to be rounded off, is squared and divided by the number of events. This is a single machine operation, and the quotient should be carried to as many places as is the sum of the squares.

(3) This quotient is subtracted from the sum of the squares to obtain a quantity numerically equal to the sum of the squares of the deviations from the mean.

(4) Divide this quantity by one less than the number of events to obtain the estimated variance (S^2). The square root of the quotient is the standard deviation (S).

When an automatic calculator is not available, steps (1), (2) and (3) can be replaced, as illustrated on fig. 4.05 by the following:

(1') Tabulate to two decimal places the difference between each logarithm and the mean logarithm. This quantity is called the deviation.

(2') Tabulate the square of each deviation to three decimal places.

(3') Add the squares of the deviations.

c. Skew. Statisticians have demonstrated, through numerous sampling experiments that the skew parameter obtained from limited samples is highly unreliable. That is, the sample skew exhibited by small samples is not representative of the skew of the population from which the samples are drawn. Consequently, the value of the skew parameter computed from samples with fewer than, say, 100 events must be viewed with considerable skepticism unless there is physical evidence to support the computed skew value. Because hydrologic records are usually less than 100 years in length, the sample skew is frequently considered to be unrepresentative. A more reliable skew value can be obtained by studying the skew characteristics of all available streamflow records in a fairly large region, weighting the skew values obtained for each streamflow record in accordance with the length of the record (longer records receive greater weight) and computing an average regional skew based on the weighted individual values. In Section 5.03 skew values of annual flood peaks and flood volumes obtained from a regional study of streams throughout the United States are tabulated.

ILLUSTRATIVE EXAMPLE

ANALYTICAL COMPUTATION OF PEAK-FLOW FREQUENCY CURVE
MILL CREEK NEAR LOS MOLINOS, CALIFORNIA

(1) Water Year	(2) Flow (c.f.s.)	(3) Log (X)	(4) Dev. (x)	(5) x^2
1928-29	1,520	3.18	-.49	.240
30	6,000	3.78	.11	.012
31	1,500	3.18	-.49	.240
32	5,440	3.74	.07	.005
33	1,080	3.03	-.64	.410
34	2,630	3.42	-.25	.062
1934-35	4,010	3.60	-.07	.005
36	4,380	3.64	-.03	.001
37	3,310	3.52	-.15	.022
38	23,000	4.36	.69	.476
39	1,260	3.10	-.57	.325
1939-40	11,400	4.06	.39	.152
41	12,200	4.09	.42	.176
42	11,000	4.04	.37	.137
43	6,970	3.84	.17	.029
44	3,220	3.51	-.16	.026
1944-45	3,230	3.51	-.16	.026
46	6,180	3.79	.12	.014
47	4,070	3.61	.06	.004
48	7,320	3.86	.19	.036
49	3,870	3.59	-.08	.006
1949-50	4,430	3.65	-.02	.000
51	3,870	3.59	-.08	.006
52	5,280	3.72	.05	.002
53	7,710	3.89	.22	.048
54	4,910	3.69	.02	.000
1954-55	2,480	3.39	-.28	.078
56	9,180	3.96	.29	.084
57	6,140	3.79	.12	.014
58	6,880	3.84	.17	.029
(6)	N	30		
(7)	ΣX	109.97	-.13	2.665
(8)	$\bar{X}$	3.666		

(9)	ΣX^2	405.7813	$S^2 = 2.6679/29 = .091996$
(10)	$(\Sigma X)^2/N$	403.1134	
(11)	Σx	2.6679	$S = .303$

(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)
P_∞	0.1	1.0	10	50	90	99	99.9
k (Ex 2)	3.09	2.33	1.28	0.0	-1.28	-2.33	-3.09
log $Q_{@p_\infty}$ (Eq 4-5)	4.602	4.372	4.054	3.666	3.278	2.960	2.730
$Q@p_\infty$	40,000	23,550	11,320	4,630	1,900	912	537
P_N (Ex 3)	.25	1.47	10.8	50	89.2	98.53	99.75

Note: Columns 4 and 5 are not required when desk calculator is available, but are shown to illustrate procedure usable without desk calculator. The difference between 2.665 and 2.668 is due to rounding values in column 4.

Figure 4.05

These values are probably satisfactory for use in studies for regions where the general climatogical conditions do not differ significantly from those experienced in the continental U.S., and they may be satisfactory for other regions as well. However, where it is possible to do so it would be desirable to confirm the applicability of the values to the region in question by a comprehensive study of the streamflow data in that region.

If a regional skew value (or the values given in Section 5.03) are used for analytical computation of a frequency curve it is usually desirable to plot the analytically derived curve and the graphically derived curve (as obtained by procedures described in Section 4.01) on a single sheet of probability paper for comparison purposes. If there are radical differences between the two curves and if a sound physical reason exists for use of either the sample or an adjusted skew rather than the regional skew, it may be desirable to recompute the analytical frequency curve using the sample, or adjusted skew instead of the regional skew. The fit in the upper range of the data for annual flood peaks should be noted. Occasionally one or two of the annual peaks at the low end of the curve will result in a computed skew coefficient that causes unreasonable results at the upper end, i.e., the largest recorded value may have an exceedence frequency of .01 percent (10,000-year exceedence interval) based on the computed statistics. It may be necessary to drop the low values and recompute the skew coefficient if a regional value is not practical.

There are physical geographic features that affect the skewness of the annual flood peaks. A location that has a sizeable percent of the area in lakes and ponds or that has an extremely wide flood plain relative to the volume of flow will tend to have negative skew coefficents when computing the log-Pearson type III statistics. A basin that is fan-shaped, or shaped such that the peak from two or more tributaries may arrive at the location in question at the same time, will tend to have a positive skew coefficient.

d. Computation of Curve. Each frequency curve can be computed as follows (see fig. 4.05 for example):

(1) For selected values of P_∞, tabulate values of k obtained from Exhibit 2 corresponding to the adopted skew coefficient (zero value assumed for skew in the example).

(2) Multiply each of these by the computed standard deviation, and add each product in turn to the mean logarithm (equation 4-5).

(3) Tabulate values of P_N from Exhibit 3 corresponding to the selected P_∞ values and the value of N-1.

(4) Plot the antilogarithm of each of the sums obtained in (2) against the corresponding P_N values obtained in (3).

Because the Pearson type III distribution is identical to the normal distribution when the skew is zero it is possible to simplify the procedure described above if a skew value of zero is used. Rather than selecting values of P_∞ initially, obtaining values of k from Exhibit 2 and converting P_∞ to P_N through use of Exhibit 3, one selects values of P_N initially, obtains corresponding values of k from Exhibit 4, computes corresponding values of log Q as in step (2) above, and plots the antilogs of the sums thus obtained against the selected values of P_N.

Section 4.03. Partial-Duration Curves

Once a frequency curve of annual events has been established, an approximation to the corresponding partial-duration curve can be determined analytically by use of average criteria derived by Walter Langbein (reference 7). These criteria are based on the assumption that there are a large number of flood events each year and that these events are mutually independent. The criteria should not be used without checking their applicability unless only very approximate results are desired and time does not permit a more accurate determination. An average relationship developed empirically from many stations in a region would ordinarily be preferred, because experience indicates that the observed relationship is often different from the theoretical relationship. Before adopting a partial-duration curve based on the analytical calculations it is always desirable to plot the partial-duration data using the graphical techniques described in Section 4.01 and to compare the graphically derived curve with the analytical curve. Radical differences between the two curves would ordinarily suggest that there is reason to question the validity of the analytical curve.

A summary of the Langbein criteria is contained in the following tabulation.

Corresponding Exceedence Frequencies per Hundred Years

Annual-event curve (No. of years flow is exceeded per hundred years)	Partial-duration curve (No. of times flow is exceeded per hundred years)
1.00	1.00
2.00	2.02
5.0	5.1
10.0	10.5
20	22.3
30	35.6
40	51.0
50	69.3
60	91.7
63.2	100
70	120
80	161
90	230
95	300

Section 4.04. Analytical Adjustment of Statistics

In most cases of frequency studies of runoff or precipitation there are locations in the region where records have been obtained over a long period. The additional period of record at such a nearby station is useful for extending the record at the site provided there is reasonable correlation between recorded values at the two locations.

It is possible, by correlation or other means, to estimate from concurrent records at nearby locations the magnitude of individual missing events at a given site. Techniques for making these estimates are fully described in Chapter 4 of Volume 2. However, as explained in that chapter, the use of regression analysis produces estimates which do not preserve the full variance exhibited by recorded data. While this may not create serious problems if only one or two events must be estimated to "fill in" or complete an otherwise unbroken record of 40 years or more, it does create serious problems if it becomes necessary to estimate more than several events. Consequently, in frequency studies, missing events should not be freely estimated by regression analysis.

In order to ensure that estimates, when necessary, do not unduly reduce the computed standard deviation, an estimate of a missing event based on concurrent data at a nearby long record (base) station can be made through use of the following equation:

$$X_1-\overline{X}_1 = (X_b-\overline{X}_b)\ S_1/S_b \qquad (4\text{-}6)$$

where:

X_1 = the logarithm of the missing event

$\overline{X}_1$ = the mean logarithm of events at station 1 for the period of continuous, concurrent record with the base station

X_b = the logarithm of the event at the base station corresponding to the time period for which the event at station 1 is to be estimated

$\overline{X}_b$ = the mean logarithm of events at the base station for the period of continuous, concurrent record with station 1.

S_1 = the standard deviation of the logarithms of events at station 1 during the period of concurrent record

S_b = the standard deviation of the logarithms of events at the base station during the period of concurrent record.

In cases where frequency curves are calculated analytically using the techniques described in Section 4.02, it is neither necessary nor desirable to estimate individual missing events. Instead, the statistics of the frequency curve can be adjusted to reflect the degree of correlation between a station and a nearby long-record base station. The procedure for adjusting the statistics involves three steps: (1) computing the degree of correlation between the two stations, (2) using the computed degree of correlation and the statistics of the base station to compute an adjusted set of statistics for the shorter-record station, and (3) computing an equivalent "length of record" that appropriately reflects the adjustment of the statistics of the short-record station.

The degree of correlation is reflected in the adjusted correlation coefficient $\overline{R}$ as computed through use of the following equations:

$$\overline{R}^2 = 1-(1-R^2)\left(\frac{N-1}{N-2}\right) \tag{4-7}$$

where:

$\overline{R}$ = the adjusted correlation coefficient

R^2 = is the determination coefficient as computed in equation 4-8 below; is the square of the unadjusted correlation coefficient

N = the number of years of concurrent record

$$R^2 = \frac{(\Sigma XY-(\Sigma X\Sigma Y)/N)^2}{[\Sigma X^2-(\Sigma X)^2/N][\Sigma Y^2-(\Sigma Y)^2/N]} \qquad (4\text{-}8)$$

where:

R^2 = the determination coefficient

X = the logarithm of streamflow at the short-record station

Y = the logarithm of streamflow at the base station

N = the number of years of concurrent record

Σ = perform calculation for all values from 1 to N

XY = the product of concurrent pairs of logarithms at the two stations

The computations required to solve equations 4-7 and 4-8 can be facilitated by arranging streamflow data pairs in order of chronological sequence (not order of magnitude) as shown on fig. 4.06. The correlation coefficient thus computed may be higher or lower than the "true" correlation coefficient (because it is a sample statistic in the same sense that $\overline{X}$ and S are estimators of μ and σ) depending on chance variations in the data. When many such correlation coefficients have been computed for many pairs of stations in a region, it may be possible to modify the computed correlation coefficient by the use of judgement to produce a more reliable estimate of the degree of correlation. It should be remembered, however, that this correlation coefficient usually has a minor influence on the ultimate frequency curve, and therefore extensive studies designed to improve its reliability might not be warranted.

The correlation coefficient is used in the adjustment of the statistics of the short-record station in the following manner:

$$S_1'-S_1 = (S_b'-S_b)\ R^2\ S_1/S_b \quad \text{(approximate)} \qquad (4\text{-}9)$$

ANALYTICAL FREQUENCY COMPUTATION AND ADJUSTMENT

I Computation of Statistics

Water Year	Annual max flow (Q) Sta. 1 cfs	Sta. 2 cfs	Log Q Sta. 1 (X)	Sta. 2 (Y)	Sta. 2
(1)	(2)	(3)	(4)	(5)	(6)
1911-12		4,570			-3.66
1912-13		7,760			3.89
1913-14		32,400			4.51
1914-15		27,500			4.44
1915-16		19,000			4.28
1916-17		24,000			4.38
1917-18		13,200			4.12
1918-19		15,500			4.19
1919-20		10,200			4.01
1920-21		14,100			4.15
1921-22		14,800			4.17
1922-23		10,500			4.02
1923-24		11,500			4.06
1924-25		27,500			4.44
1925-26		17,800			4.25
1926-27		36,300			4.56
1927-28		67,600			4.83
1928-29	1,520	5,500	3.18	3.74	(Values same as in Column 5)
1929-30	6,000	25,500	3.78	4.41	
1930-31	1,500	5,570	3.18	3.75	
1931-32	5,440	9,980	3.74	4.00	
1932-33	1,080	5,100	3.03	3.71	
1933-34	2,630	11,100	3.42	4.05	
1934-35	4,010	25,500	3.60	4.41	
1935-36	4,380	38,200	3.64	4.58	
1936-37	3,310	7,920	3.52	3.90	
1937-38	23,000	93,000	4.36	4.97	
1938-39	1,260	3,230	3.10	3.51	
1939-40	11,400	60,200	4.06	4.78	
1940-41	12,200	30,300	4.09	4.48	
1941-42	11,000	35,100	4.04	4.55	
1942-43	6,970	54,300	3.84	4.73	
1943-44	3,220	8,460	3.51	3.93	
1944-45	3,230	28,600	3.51	4.46	
1945-46	6,180	22,000	3.79	4.34	
1946-47	4,070	17,800	3.61	4.25	
1947-48	7,320	16,600	3.86	4.22	
1948-49	3,870	6,140	3.59	3.79	
1949-50	4,430	17,900	3.65	4.25	
1950-51	3,870	50,200	3.59	4.70	
1951-52	5,280	21,000	3.72	4.32	
1952-53	7,710	40,000	3.89	4.60	
1953-54	4,910	22,900	3.69	4.36	
1954-55	2,480	5,900	3.39	3.77	
1955-56	9,180	104,000	3.96	5.02	
1956-57	6,140	32,700	3.79	4.51	
1957-58	6,880	39,300	3.84	4.59	

		(10)	(11)	(12)
	N	30	30	47
	ΣX	109.97	128.68	200.64
(Eq. 4-2)	$\bar{X}$	3.666	4.289	4.269
		(7)	(8)	(9)
	ΣX^2	405.7813	556.5276	862.3764
	$(\Sigma X)^2/N$	403.1134	551.9514	856.5194
	Σx^2	2.6679	4.5762	5.8570
	$\Sigma x^2/N - 1$	.0920	.1578	.1273
(Eq. 4-3)	S	.303	.397	.357
	ΣXY		474.5925	
	$\Sigma X \Sigma Y/N$		471.6980	
	Σxy		2.8945	

II Adjustment of Statistics

(7) $S_1 = .303$ (Eq. 4-3)

(8) $S_b = .397$ (Eq. 4-3)

(9) $S_b' = .357$ (Eq. 4-3)

(10) $\bar{X}_1 = 3.666$ (Eq. 4-2)

(11) $\bar{X}_b = 4.289$ (Eq. 4-2)

(12) $\bar{X}_b' = 4.269$ (Eq. 4-2)

(13) $R^2 = \dfrac{(2.8945)^2}{(2.6679)(4.5762)} = .685$ (Eq. 4-8)

(14) $\bar{R}^2 = 1 - (1 - .685)\dfrac{29}{28} = .67$ (Eq. 4-7)

(15) $S_1' = .303 + (.357 - .397)(.685)(.303 / .397)$

$= .282$ (Eq. 4-9)

(16) $\bar{X}_1' = 3.666 + (4.269 - 4.289)(.685)^{1/2}$

$(.303 / .397) = 3.653$ (Eq. 4-10)

(17) $g = .00$ (assumed)

(18) $N_1' = \dfrac{30}{1 - 17/47(.67)} = 39.6$ years (Eq. 4-11)

III Computation of frequency curve based on $\bar{X}_1'$, S_1' and g

P_N'	k* (Ex. 4)	Log Q ($\bar{X}_1' + kS_1'$)	Q
(19)	(20)	(21)	(22)
.25	3.01	4.501	31,700
1.0	2.45	4.344	22,100
5.0	1.70	4.132	13,600
10.0	1.32	4.025	10,600
30.0	0.54	3.805	6,380
50.0	0.00	3.653	4,500
70.0	-0.54	3.501	3,170
90.0	-1.32	3.281	1,910
95.0	-1.70	3.174	1,490
99.0	-2.45	2.962	916
99.75	-3.01	2.804	637

*Based on zero skew and 41 years of record.

Figure 4.06

where:

S_1' = the adjusted standard deviation at the short-record station

S_1 = the standard deviation for the period of concurrent record at the short-record station

S_b' = the standard deviation for the entire period of record at the base station

S_b = the standard deviation for the period of concurrent record at the base station

R^2 = the determination coefficient

$$\overline{X}_1' - \overline{X}_1 = (\overline{X}_b' - \overline{X}_b) \; R \;\; S_1 \; / \; S_b \qquad (4\text{-}10)$$

where:

$\overline{X}_1'$ = the adjusted mean logarithm at the short-record station

$\overline{X}_1$ = the mean logarithm for the period of concurrent record at the short-record station

$\overline{X}_b'$ = the mean logarithm for the period of concurrent record at the base station.

$\overline{X}_b$ = the mean logarithm for the period of concurrent record at the base station

R = the unadjusted correlation coefficient

S_1 = the standard deviation at the short-record station

S_b = the standard deviation for the concurrent period of record at the base station.

An example which illustrates the computation of $\overline{R}^2$, $\overline{X}_1'$, and S_1' from the above equation is shown on fig. 4.06.

The final step in adjusting the statistics is the computation of the "equivalent length of record" which is defined as the period of time which would be required to establish unadjusted statistics that are as reliable (in a statistical sense) as the adjusted values. Thus, the equivalent length of record is an indirect indication of the reliability of the adjusted values of $\overline{X}_1'$ and S_1'. The equivalent length of record is computed from the following equation:

$$N_1' = \frac{N_1}{1 - \dfrac{N_1' - N_1}{N_b'} \overline{R}^2} \qquad (4\text{-}11)$$

where:

N' = the equivalent length of record at the short-record station

N_1 = the number of years of concurrent record at the two stations

N_b' = the number of years of record at the base station

$\overline{R}$ = the adjusted correlation coefficient

Thus, on fig. 4.06, it can be seen that the adjustment of the frequency statistics provides an increased reliability equivalent to having an additional 11 years of record at the short-record station.

The procedure for computing and adjusting frequency statistics using a base station can be summarized as follows:

a. Arrange the streamflow data by pairs in order of chronological sequency.

b. Compute $\overline{X}$ and S for the entire record at the short-record station.

c. Compute $\overline{X}$ and S for the entire record at the base station.

d. Compute $\overline{X}$ and S for the portion of the base station record which is concurrent with the period of record at the short-record station.

e. Compute the correlation coefficient using equations 4-7 and 4-8.

f. Adjust $\overline{X}$ and S for the short-record station using equations 4-9 and 4-10.

g. Calculate the equivalent length of record for the short-record station using equation 4-11.

h. Compute the frequency curve using adjusted values of $\overline{X}'$ and S' in equation 4-5.

Section 4.05. Use of Estimated Data

Where the magnitude of one or more large historical floods has been estimated, the estimated magnitudes can be used as a guide in drawing the frequency curve, particularly in the range of higher flows. Such adjustment would ordinarily be done graphically, and the historical flows would be plotted in the order of magnitude in which they are known to have occurred, using the entire period of history (not only that period dating from the earliest measured flood) for the computation of plotting positions. While it is important that all known information be used in the construction of a frequency curve, it should be recognized that historical estimates are not as valuable as comparable recorded flows, and that the largest known floods are not always representative for the period. If historical flows are particularly outstanding relative to recorded data, procedures illustrated on figs. 4.07 and 4.08 can be used to compute a composite frequency curve.

The procedure requires the computation of a best-fit line from the following equations:

$$b = \left[\frac{\Sigma X_1^2 - (\Sigma X_1)^2/N}{\Sigma X_2^2 - (\Sigma X_2)^2/N} \right]^{1/2} \tag{4-12}$$

ILLUSTRATIVE EXAMPLE

ANALYTICAL FREQUENCY COMPUTATION
USING PRE-RECORD DATA

Location: Willamette R at Albany, Oregon
Period of record: 1893-1958 (65 years)
Years of historical flood estimates: 1861, 1887, 1890 (3 largest known)
Period covered by history and record: 1858-1958)(100 years)

Event No.	k + 4.00 N=100	k + 4.00 N=65	k + 4.00 Smaller	log Q
(1)	(2)	(3)	(4)	(5)
	(Ex 5)	(Ex 5)		
1	6.58		6.58	2.53
2	6.17		6.17	2.46
3	5.96		5.96	2.42
4	5.81	6.43	5.81	2.36
5	5.70	6.00	5.70	2.36
6	5.60	5.77	5.60	2.36
7	5.51	5.61	5.51	2.33
8	5.44	5.48	5.44	2.32
9	5.37	5.38	5.37	2.31
10	5.31	5.29	5.29	2.31
11	5.25	5.20	5.20	2.29
12	5.20	5.13	5.13	2.29
13	5.15	5.05	5.05	2.28
14	5.10	4.99	4.99	2.26
15	5.06	4.93	4.93	2.25
16		4.87	4.87	2.23
17		4.82	4.82	2.23
18		4.76	4.76	2.22
19		4.71	4.71	2.22
20		4.66	4.66	2.18
21		4.62	4.62	2.16
22		4.57	4.57	2.14
23		4.53	4.53	2.14
24		4.48	4.48	2.14
25		4.44	4.44	2.13
26		4.40	4.40	2.13
27		4.35	4.35	2.11
28		4.31	4.31	2.11
29		4.27	4.27	2.11
30		4.23	4.23	2.10
31		4.19	4.19	2.10
32		4.16	4.16	2.10
33		4.12	4.12	2.09
34		4.08	4.08	2.08
35		4.04	4.04	2.07
36		4.00	4.00	2.06
37		3.96	3.96	2.06
38		3.92	3.92	2.04
39		3.88	3.88	1.99

Event No.	k + 4.00 N=100	k + 4.00 N=65	k + 4.00 Smaller	log Q
(6)	(7)	(8)	(9)	(10)
		(Ex 5)	(Ex 5)	
40		3.84	3.84	1.98
41		3.81	3.81	1.97
42		3.77	3.77	1.97
43		3.73	3.73	1.96
44		3.69	3.69	1.96
45		3.65	3.65	1.96
46		3.60	3.60	1.94
47		3.56	3.56	1.94
48		3.52	3.52	1.91
49		3.47	3.47	1.91
50		3.43	3.43	1.90
51		3.38	3.38	1.89
52		3.34	3.34	1.88
53		3.29	3.29	1.88
54		3.24	3.24	1.88
55		3.18	3.18	1.88
56		3.13	3.13	1.86
57		3.07	3.07	1.85
58		3.01	3.01	1.85
59		2.95	2.95	1.84
60		2.87	2.87	1.78
61		2.80	2.80	1.77
62		2.71	2.71	1.76
63		2.62	2.62	1.73
64		2.52	2.52	1.72
65		2.39	2.39	1.72
66		2.23	2.23	1.69
67		2.00	2.00	1.67
68		1.57	1.57	1.61
N"			68	68
ΣX			277.47	139.73
$\bar{X}'$ (Eq 4-2)			4.080	2.055
ΣX^2			1206.2039	290.1517
$(\Sigma X)^2/N''$			1132.2000	287.1246
diff.			74.0039	3.0271

$\underline{S} = b = (3.0271/74.0039)^{1/2} = .202$ (Eq 4-12)
$\bar{X} = a = 2.055 - .202(4.080-4.000) = 2.039$ (Eq 4-13)

Computation of Frequency Curve (N = 65, g = 0)

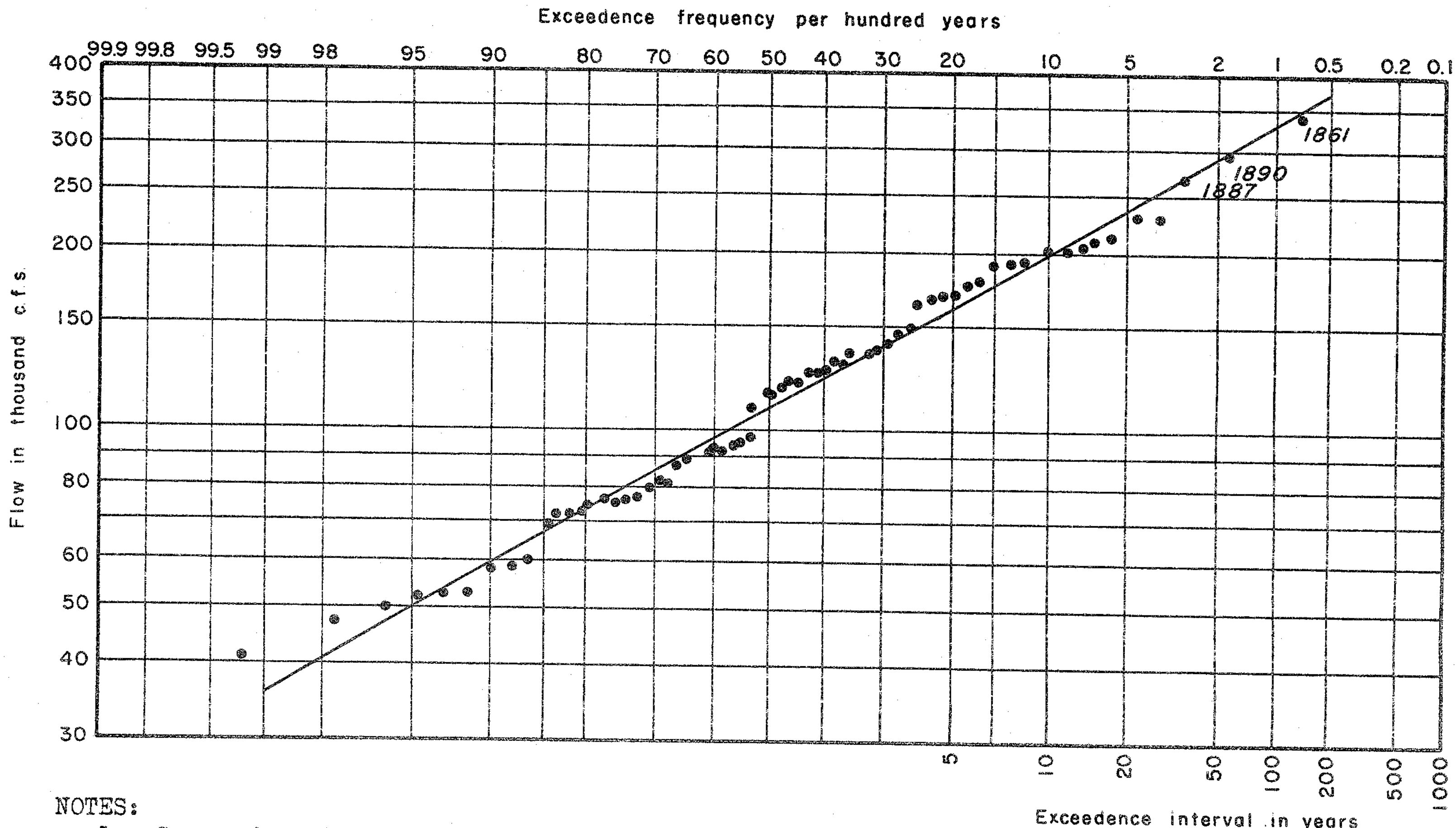
(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
P_N	0.25	1.0	10	50	90	99	99.75
k (Ex 4)	2.94	2.41	1.31	0.00	-1.31	-2.41	-2.94
Log Q (Eq 4-5)	2.633	2.526	2.304	2.039	1.774	1.552	1.445
Q, cfs	430,000	336,000	201,000	109,000	59,400	35,600	27,900

NOTES:

Events numbered 4 to 9 could not have occupied higher positions if the entire 100-year record were available, (but might occupy lower positions). Accordingly, k values were selected as shown above.

4.00 was added to all k values in order that all numbers are positive and have 3 digits for simplicity of machine operation.

Figure 4.07



NOTES:

1. Curve shown hereon is for illustrative purposes only and was computed from data recorded from 1893 to 1958 and on estimates of three pre-record peak flows. Flows that were partly regulated (since 1941) were adjusted to natural conditions.

2. See fig. 4.06 for computations.

ILLUSTRATIVE EXAMPLE

ANALYTICAL FREQUENCY CURVE USING PRE-RECORD DATA
WILLAMETTE RIVER AT ALBANY, OREGON

where:

b = the slope of the best-fit line

X_1 = the logarithms of the historical and recorded events

X_2 = the corresponding k values for the historical and recorded events

Σ = perform calculation for all values from 1 to N

N = the total number of events (historical and recorded)

$$a = \overline{X}_1 - b\overline{X}_2 \qquad (4\text{-}13)$$

where:

a = the intercept of the best-fit line

$\overline{X}_1$ = the mean logarithm of the historical and recorded events

$\overline{X}_2$ = the mean k value of the historical and recorded events

b = the slope of the best-fit line

In general the procedure assuming zero skew can be summarized as follows:

a. Tabulate all floods (historical and recorded) in order of magnitude.

b. Select (from Exhibit 5) k values for each event (historical and recorded) based on N equal to the total period of time from the earliest historical flood to the latest recorded flood.

c. Select (from Exhibit 5) k values for each recorded event based on N equal to the period of record only.

d. Tabulate the selected k value for the historical floods, and the smaller of the two k values for each of the recorded floods.

e. Tabulate the logarithms of all historical and recorded events.

f. Compute the mean k value and the mean logarithm for all events.

g. Compute the slope and intercept of the best-fit line using equations 4-12 and 4-13.

h. Compute the frequency curve using equation 4-5 and assuming that the slope of the best-fit line equals the standard deviation and the intercept equals the mean.

Although it is possible to estimate flows based on records at nearby locations, use of a large number of such estimates might lead to erroneous frequency estimates. However, there are some cases where one or a few estimates are essential. In the case where a record is not obtained because the gage was washed out, it is imperative that some estimate of that value be used. The fact that the estimate may be in error by 25 percent is minor compared to the error introduced by omitting the value from the record. Also, where comprehensive flood volume-duration series is being studied (see Chapter 5), and a few of the items are missing, it is ordinarily advantageous to estimate these items rather than to have a different number of items for each duration studied.

The shape of the frequency curve can be seriously distorted by the dominance of minor runoff factors during dry years. This is particularly true where (a) floods are normally caused by rainstorms, yet high base flows from underground sources or occasional snowmelt prevent true measures of rainflood runoff during very dry years, or (b) floods are normally caused by rainstorms, but excessive diversion or channel losses deplete flows during dry years so that the frequency curve drops sharply at the lower end. In such cases, it may be advantageous to fit only the upper half of the annual floods with a theoretical frequency curve. A suggested procedure, illustrated on figs. 4.09 and 4.10, simply converts plotting positions to a linear distance as measured on probability grid (k values on Exhibit 5) and solves for a best-fit line (standard deviation) by use of equations 4-12 and 4-13. This procedure produces a substantial reduction in reliability and should only be used when the frequency curve is grossly distorted by conditions such as those described above.

ILLUSTRATIVE EXAMPLE

ANALYTICAL FREQUENCY COMPUTATION OMITTING LOWER FLOWS

Chronological order			Order of magnitude				
Water Year	Date	Peak flow (cfs)	No.	Plotting position (%)	Peak flow (cfs)	Log of peak	k
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1921	16 May	1250	1	1.9	1890	3.28	2.20
1922	6 May	1380	2	4.7	1780	3.25	1.74
1923	10 May	1450	3	7.4	1480	3.17	1.48
1924	4 May	618	4	10.2	1450	3.16	1.30
1925	8 May	523	5	12.9	1430	3.16	1.15
1926	21 Apr	508	6	15.7	1380	3.14	1.03
1927	30 Apr	1220	7	18.4	1300	3.11	0.91
1928	28 Apr	1180	8	21.1	1280	3.11	0.81
1929	14 May	1060	9	23.9	1250	3.10	0.72
1930	25 Apr	412	10	26.6	1220	3.09	0.63
1931	5 May	170	11	29.4	1180	3.07	0.55
1932	14 May	1480	12	32.1	1090	3.04	0.47
1933	21 May	876	13	34.9	1060	3.02	0.39
1934	21 Jul	113	14	37.6	1020	3.01	0.32
1935	10 May	516	15	40.4	1000	3.00	0.25
1936	4 May	1780	16	43.1	995	3.00	0.17
1937	8 May	1090	17	45.9	985	2.99	0.11
1938	22 Apr	760	18	48.6	876	2.94	0.04
1939	30 Apr	397	19	51.3	788		
1940	21 Apr	282	20	54.1	788		
1941	2 May	353	21	56.9	760		
1942	13 Apr	597	22	59.6	678		
1943	23 Apr	995	23	62.3	618		
1944	14 May	611	24	65.1	611		
1945	4 May	985	25	67.9	611		
1946	18 Apr	1430	26	70.6	597		
1947	3 May	788	27	73.4	523		
1948	17 May	1280	28	76.1	516		
1949	24 Apr	1020	29	78.9	508		
1950	18 May	1300	30	81.6	412		
1951	12 May	1000	31	84.3	409		
1952	3 May	1890	32	87.1	397		
1953	29 May	611	33	89.8	353		
1954	25 Apr	409	34	92.6	282		
1955	6 May	788	35	95.3	170		
1956	24 Dec	678	36	98.1	113		

	(7)	(8)
N"	18	18
ΣX	55.64	14.27
$\bar{X}$"	3.091	0.793
ΣX^2	172.1336	17.4139
$\Sigma X^2/N$"	171.9894	11.3129
Σx^2	.1442	6.1010

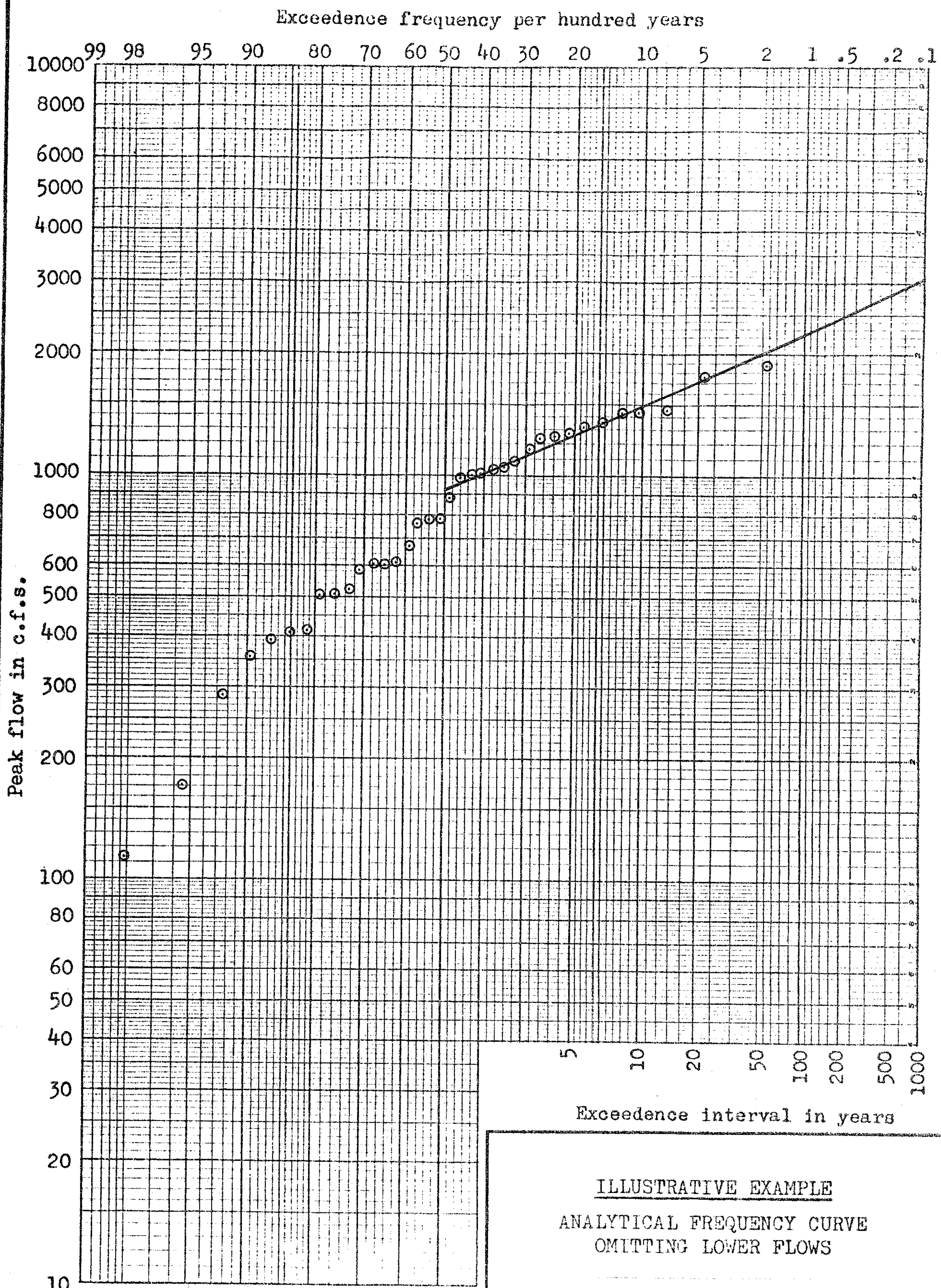
Computation of curve (N=36)

(9)	(10)	(11)	(12)
P_N	k	log Q	Q
.0025	3.04	3.437	2,740
.01	2.47	3.349	2,230
.1	1.32	3.172	1,490
.5	0	2.969	931

$S = b = (.144/6.101)^{1/2} = 0.154$ (Eq 4-12)

$\bar{X} = a = 3.091 - .154(.793) = 2.969$ (Eq 4-13)

Figure 4.09



NOTE: See fig. 4.08 for computations.

ILLUSTRATIVE EXAMPLE

ANALYTICAL FREQUENCY CURVE
OMITTING LOWER FLOWS

Figure 4.10

Frequency estimates at best are not fully reliable. Although theoretical devices are employed to extrapolate frequency curves beyond experienced values, such extrapolation is highly dangerous. It is sometimes possible to use synthetic floods for constructing a frequency curve or especially for extrapolating one more accurately.

Where runoff records are not available, an attempt is sometimes made to compute floods that would result from various rainfall amounts, and then to construct a synthetic flood frequency curve, using the rainfall frequency curve as a guide. This method is considered satisfactory for airport drainage design and for urban storm drain design, but is ordinarily not satisfactory where infiltration losses are a considerable percentage of precipitation. In these latter cases, it is best to construct synthetic frequency curves from regional correlation studies as discussed in Chapter 7.

A large hypothetical flood can sometimes be used as a guide in extrapolating the frequency curve. Where it appears that major storms may have accidentally missed the basin considered, a major observed storm might be transposed to the basin, and the flood resulting from this storm on wet ground conditions could be assigned a reasonable frequency and used as an "anchor point" for extrapolating the frequency curve.

While the probable maximum flood is defined as the largest flood that is reasonably possible at a location, it would not be said that a larger flood could absolutely not occur. The science of meteorology and hydrology has not yet advanced to the stage where an absolute maximum estimate can be made. Consequently, it is not considered necessary that a frequency curve be limited to values smaller than the probable maximum flood. However, it is considered that such a flood would have an exceedence interval considerably in excess of 1,000 years.

Section 4.06. Special Geographic Considerations

Frequency studies of peak flows made in some areas have indicated that the frequency curves have a strong positive skew. In studying this problem, two peculiar conditions were observed.

a. At sites where the period of record was short with respect to other records in the region, and where a large proportion of the major floods experienced in the region had occurred in the recent past, the positive skew tended to be greatest. Thus, it might be concluded that the positive skew was caused by a highly unrepresentative period of flood activity. The period of the last years, which includes the period of record at most of the stream gaging stations, has an abnormally large number of major floods in comparison with the 100-year period covered by the longer records and in comparison with the 300-year period of history.

b. Major floods in these regions occur from at least two independent causes, tropical hurricane storms and extratropical cyclones. Hurricane floods are comparatively rare, but produce extreme flows, and therefore cause an upward curvature of the frequency curve of annual maximum flows. Some improvement in frequency estimates is attained by segregating hurricane and non-hurricane floods (see reference 21). However, this apparently does not solve entirely the problem of upward curvature of the frequency curves.

In many desert areas, streamflow from general storms is ordinarily moderate each year, but an occasional intense thunderstorm may center over a basin and cause an exceptionally large flood. As cited in the second instance above, floods from these two different causes tend to produce a sharp upward curvature of the frequency curve. While a satisfactory solution to this problem has not been attained, a reasonable approach appears to be to estimate the frequency of thunderstorms on a regional basis and if desired, to combine the estimated frequency of thunderstorm flows with the frequency of flows from general storms. In some basins, an entire year may pass with zero runoff. When analytical methods using flow logarithms are used, this presents a particular difficulty, since the logarithm of zero is minus infinity. It may be best to omit such years from the record, compute a tentative frequency curve based on the remaining years, and then adjust the exceedence frequency by the ratio of the number of years of record to the number of years with runoff. However, factors resulting in zero flow usually also affect small flows to the extent that the shape of the frequency curve is distorted in the range of lower flows. It is possible to compute only the upper half of the frequency curves, in such cases, using procedures described in Section 4.04.

In several regions in the United States, runoff frequency curves are very steep. This occurs particularly in areas where the combination of extreme meteorological events (such as infrequent severe thunderstorms) and extreme topographic and geologic features (steep slopes or high imperviousness) produce extreme variations in runoff. In extrapolating these frequency curves beyond the range of experienced floods by analytical means, unreasonably high flood estimates may result. In these cases, extra care must be used in extrapolation, and hypothetical computed floods might well be used as a general guide.

Where frequency curves are constructed for two locations on the same stream, care should be exercised to assure that the curves are mutually consistent throughout their length. This can be accomplished by first establishing a curve for the location where the record is longest and then deriving the second curve from this in conjunction with an empirical relationship of flows at the two locations in a manner described in Section 6.04 and illustrated on fig. 6.04.

In estimating flood peak frequencies for large drainage basins such as the Ohio River and Missouri River basins in the United States, regional studies of skew coefficients based on smaller basins may not be applicable. However, studies of several records from very large basins have indicated that the frequency curves of annual peak flows in all cases correspond at least approximately to the logarithmic normal (zero skew) curve. Accordingly, use of a skew coefficient of zero should be tested, and used if the data do not depart radically from the frequency curve.

A study of Mississippi River flows at several locations has been conducted to determine which frequency method is most applicable to these flows. The conclusion drawn is that both the logarithmic normal and the extreme-value method (reference 30) give satisfactory results, because the standard deviation of flows is small in comparison to average flows for Mississippi River. The logarithmic normal procedure is recommended herein, because it has been extensively tested and because fitting procedures are generally more efficient, especially in cases where the standard deviation is large in comparison to average flows.

Section 4.07. Special Problems in Peak Discharge Frequency Studies

While instantaneous peak flows are the most common items studied in a hydrologic frequency analysis, reliable records of peak flows are not ordinarily obtainable unless a continuous recording stream gage is used. In the early years, most of the stream flow records were obtained from daily readings of staff gages, and consequently peak flows were usually not recorded. The ordinary procedure used in such a case is to construct a frequency curve of daily flows for the entire period of record, then to construct a curve relating peak flows to daily flows for the period during which the recording gage existed, and then to derive a frequency curve of peak flows from these two curves. An alternative method that can be used when analytical methods are employed consists of computing frequency statistics of peak and mean daily flows separately for the period of the recording gage, then computing a second set of frequency statistics of mean daily flows for the entire period of record, and adjusting the frequency statistics of peak flows to the longer period by equations 4-9 and 4-10 (using the daily flows as a base station would be used). It is generally best to arbitrarily use a determination coefficient of 1.0 for this computation.

Where a substantial portion of the drainage basin is regulated by one or more reservoirs, reconstitution of the unregulated flows is difficult because of reservoir and channel routing effects. In cases where there is a substantial record prior to construction of the reservoir, it might be possible to relate peak flows to maximum 3-day flows for that early period, to construct a frequency curve of 3-day flows adjusted for storage effects for the entire period of record, and then construct a frequency curve of peak flows for the entire period, using 3-day flows as a base and procedures described in Section 4.04. The advantage of such a procedure is due to the fact that 3-day flows are ordinarily not appreciably affected by channel storage and can be easily adjusted for the effect of reservoir storage.

Chapter 5

Flood Volume-Duration Frequencies

CHAPTER 5. FLOOD VOLUME-DURATION FREQUENCIES

Section 5.01. Nature and Purpose

A comprehensive flood volume-duration frequency series consists of a set of frequency curves as follows:

a. Maximum rate of flow for each water year.
b. Maximum 1-day average flow for each water year.
c. Maximum 3-day average flow for each water year.
d. Maximum 7 or 10-day average flow for each water year.
e. Maximum 30-day average flow for each water year.
f. Maximum 90-day average flow for each water year.
g. Average flow for each water year.

Runoff volumes are expressed as average flows in order that peak flows and volumes can be readily compared and coordinated. Whenever it is necessary to consider flows for only a portion of the water year, such as the rain season or snowmelt season, the same items (up to the 30-day or 90-day values) are selected from flows during that season only. A comprehensive flood volume-duration series is used primarily for reservoir design and operation studies for reservoirs having flood control as a major function. When reservoir analysis involve runoff durations greater than one year, frequency studies might well include multi-annual runoff volumes and consideration of seasonal effects, as discussed in Section 5.05.

Section 5.02. Data for Comprehensive Series

Data to be used for a comprehensive volume-frequency study should be selected from complete water-year records. Unless overriding reasons exist, durations specified in Section 5.01 should be used, in order to assure consistency among various studies for comparison purposes.

Peak flows should be selected only for those years when recorder gages existed or when peak flows were measured by other means. Where a minor portion of an annual record is missing, the longer-duration flood volumes for that year can often be estimated adequately. Where upstream regulation or diversion exists, care should be exercised to assure that each period selected is that when flows would have been maximum under the specified (usually natural) conditions. The dates and amounts of each selected average flow should be tabulated in chronologic order. A typical tabulation is illustrated on fig. 5.01.

Section 5.03. Statistics for Comprehensive Series

The analytical method used for volume-frequency computations is based on fitting the Pearson type III function by use of moments of flow logarithms. In practice, only the first two moments, computed by use of equations 4-2 and 4-3, are based on station data. As discussed in Section 4.02, the skew coefficient should not be based on a single record, but should be derived from regional studies. The following coefficients are based on studies of streams throughout the United States and are considered to be generally applicable for annual maximum flood volume-frequency computations in areas of similar climatological conditions if regional values have not been derived.

Duration	Skew Coefficient
Instantaneous	0
1 day	-.04
3 days	-.12
10 days	-.23
30 days	-.32
90 days	-.37
1 year	-.40

MAXIMUM-RUNOFF VOLUME FREQUENCY DATA--CHRONOLOGICAL ORDER

Average Flows in c.f.s. at: Mill Creek nr. Los Molinos, California

Drainage area: 134 sq. mi. Runoff type: All-season Basin regulation: Natural

Water Year (1)	Peak date (2)	Peak flow (3)	1-day date (4)	1-day flow (5)	3-day first day (6)	3-day flow (7)	10-day first day (8)	10-day flow (9)	30-day first day (10)	30-day flow (11)	90-day first day (12)	90-day flow (13)	Water-yr. flow (14)
1928-29	3 Feb.	1520	3 Feb.	965	2 Feb.	719	1 Feb.	377	22 Apr.	292	9 Mar.	225	150
29-30	15 Dec.	6000	15 Dec.	4080	14 Dec.	2400	10 Dec.	1270	10 Dec.	590	10 Dec.	475	255
30-31	23 Jan.	1500	23 Jan.	1200	22 Jan.	650	11 Mar.	315	11 Mar.	235	27 Feb.	186	121
31-32	24 Dec.	5440	27 Dec.	2160	26 Dec.	1380	23 Dec.	1040	21 Dec.	480	18 Mar.	380	214
32-33	16 Mar.	1080	28 Mar.	662	28 Mar.	472	25 May.	386	20 May.	330	27 Mar.	280	150
33-34	29 Dec.	2630	29 Dec.	1730	29 Dec.	1300	29 Dec.	745	27 Mar.	335	7 Feb.	290	166
34-35	4 Jan.	4010	4 Jan.	2300	7 Apr.	1470	7 Apr.	958	3 Apr.	760	28 Feb.	565	290
35-36	21 Feb.	4380	21 Feb.	2660	21 Feb.	1920	14 Feb.	1220	12 Feb.	716	9 Jan.	513	270
36-37	14 Feb.	3310	4 Feb.	1540	4 Feb.	902	12 Mar.	584	1 May	520	11 Mar.	478	219
37-38	11 Dec.	23000	11 Dec.	12300	10 Dec.	7340	10 Dec.	2750	10 Dec.	1100	12 Mar.	890	566
38-39	8 Mar.	1260	3 Dec.	594	8 Mar.	453	19 Mar.	350	8 Mar.	329	7 Mar.	260	156
39-40	28 Feb.	11400	27 Feb.	7640	27 Feb.	6150	26 Feb.	2580	3 Feb.	1200	1 Jan.	915	383
40-41	10 Feb.	12200	11 Feb.	5980	10 Feb.	4730	9 Feb.	1990	8 Feb.	1310	8 Feb.	605	467
41-42	6 Feb.	11000	6 Feb.	5690	5 Feb.	3250	1 Feb.	1810	22 Jan.	1160	30 Nov.	825	443
42-43	8 Mar.	6970	23 Jan.	3770	21 Jan.	3060	21 Jan.	1660	21 Mar.	820	5 Mar.	654	373
43-44	4 Mar.	3220	4 Mar.	1720	3 Mar.	890	29 Feb.	515	29 Apr.	373	29 Feb.	330	198
44-45	5 Feb.	3230	5 Feb.	1580	1 Feb.	1260	1 Feb.	890	31 Jan.	540	1 Feb.	392	170
45-46	21 Dec.	6180	27 Dec.	3100	27 Dec.	2320	21 Dec.	1760	21 Dec.	890	30 Oct.	470	294
46-47	12 Feb.	4070	12 Feb.	2590	11 Feb.	1380	11 Feb.	630	11 Feb.	400	11 Feb.	350	192
47-48	23 Mar.	7320	23 Mar.	3650	23 Mar.	2040	23 Mar.	865	23 Mar.	770	23 Mar.	682	310
48-49	11 Mar.	3870	23 Mar.	1810	10 Mar.	1300	10 Mar.	700	2 Mar.	410	2 Mar.	407	195
49-50	4 Feb.	4430	4 Feb.	3210	4 Feb.	2360	4 Feb.	990	17 Jan.	590	18 Mar.	452	251
50-51	16 Nov.	3870	22 Jan.	2120	19 Nov.	1540	3 Dec.	1060	16 Nov.	840	1 Dec.	565	343
51-52	26 Dec.	5280	1 Dec.	3040	1 Feb.	2330	31 Jan.	1250	24 Jan.	860	14 Mar.	726	445
52-53	9 Jan.	7710	9 Jan.	5240	8 Jan.	3200	7 Jan.	1900	26 Dec.	1080	1 Dec.	570	348
53-54	17 Jan.	4910	5 Apr.	2290	4 Apr.	1870	3 Apr.	1000	9 Apr.	740	12 Feb.	633	318
54-55	11 Nov.	2480	15 Nov.	1060	14 Nov.	565	2 Dec.	480	2 May	402	25 Mar.	325	202
55-56	22 Dec.	9180	22 Dec.	6770	21 Dec.	5060	18 Dec.	3260	19 Dec.	1880	18 Dec.	1080	493
56-57	24 Feb.	6140	24 Feb.	3840	24 Feb.	2630	24 Feb.	1340	23 Feb.	806	23 Feb.	528	257
57-58	24 Feb.	6880	24 Feb.	3580	24 Feb.	2530	16 Feb.	1560	29 Jan.	1290	24 Jan.	918	488

Comp. by: R.P.L. Date: May 1959 Sheet: Station: USGS 11-3810

SPK FORM 6 MAY 59 237

Figure 5.01

A sample computation of frequency statistics for a comprehensive series (using a nearby long-record station to adjust the statistics) is given on figs. 5.02 through 5.07. This can be accomplished in steps as follows:

a. Tabulate annual-maximum average flows for each duration in chronological order as shown on fig. 5.02.

b. Tabulate logarithms in chronological order as shown on fig. 5.03. If records from a long-record station nearby are available, tabulate corresponding logarithms for that base station. The work thus far can be checked approximately by ascertaining that the logarithm for each succeeding duration decreases, but by not more than about 0.5. Much inconvenience can be eliminated by using data for the same years for each duration insofar as is feasible. If a year's record is incomplete, the remainder can usually be estimated satisfactorily.

c. By use of a statistical calculator, the sums of each column of logarithms, the sums of their squares and the sum of their cross products (doubled) can be obtained for each duration in a single cumulative operation. These should be entered on fig. 5.04. In cases of peak flows, advantage can sometimes be gained by using 1-day values at the same station as a base instead of peak values at the base station, particularly if the records of peak flows are incomplete. The extended 1-day statistics would then be used as long-term values.

d. Compute the means, standard deviations and determination coefficients for each duration. If peak-flow statistics are extended by use of 1-day flows at the same station, possible inconsistencies can be avoided by using a determination coefficient of 1.00 instead of calculating the coefficient. Enter the long-term mean and standard deviation for each duration at the base station and compute the long-term mean and standard deviation for each duration for the station concerned. These operations are shown on fig. 5.04.

MAXIMUM-RUNOFF VOLUME FREQUENCY DATA--ORDER OF MAGNITUDE

Average Flows in c.f.s. at: Mill Creek nr Los Molinos, Calif.

Annual-event curve data								Partial-duration curve data				
								Additional flows				
plotting position (1)	peak flow (2)	1-day flow (3)	3-day flow (4)	10-day flow (5)	30-day flow (6)	90-day flow (7)	water-year flow (8)	water-yr. (9)	date (10)	flow (11)	plotting position (12)	flow (13)
2.3	23,000	12,300	7,340	3,260	1,880	1,080	566	1934-35	28 Feb	3,190	2.3	23,000
5.6	12,200	7,640	6,150	2,750	1,310	918	493	1934-35	8 Apr	3,040	5.6	12,200
8.9	11,400	6,770	5,060	2,580	1,290	915	488	1935-36	11 Jan	3,930	8.9	11,400
12.2	11,000	5,980	4,730	1,990	1,200	890	467	1937-38	20 Nov	4,700	12.2	11,000
15.4	9,180	5,690	3,250	1,990	1,160	825	445	1937-38	2 Feb	5,050	15.4	9,360
18.7	7,710	5,240	3,200	1,810	1,100	726	443	1937-38	23 Mar	4,950	18.7	9,180
22.0	7,320	4,080	3,060	1,760	1,080	682	383	1939-40	2 Jan	4,600	22.0	7,710
25.3	6,970	3,840	2,630	1,660	890	654	373	1939-40	30 Mar	9,360	25.3	7,320
28.6	6,880	3,770	2,530	1,560	860	633	348	1940-41	24 Dec	6,240	28.6	7,260
31.9	6,180	3,650	2,400	1,340	840	605	343	1940-41	1 Mar	4,250	31.9	6,970
35.2	6,140	3,580	2,360	1,270	820	570	318	1940-41	4 Apr	7,260	35.2	6,910
38.5	6,000	3,210	2,330	1,250	806	565	310	1941-42	3 Dec	4,130	38.5	6,880
41.8	5,440	3,100	2,320	1,220	770	565	294	1941-42	16 Dec	6,910	41.8	6,480
45.1	5,280	3,040	2,040	1,060	760	528	290	1941-42	27 Jan	5,450	45.1	6,450
48.4	4,910	2,660	1,920	1,040	740	513	270	1942-43	21 Jan	6,450	48.4	6,240
51.6	4,430	2,590	1,870	1,000	716	478	257	1945-46	4 Dec	3,660	51.7	6,180-
54.9	4,380	2,300	1,540	990	590	475	255	1947-48	28 Apr	3,380	55.0	6,140
58.2	4,070	2,290	1,470	958	590	470	251	1950-51	22 Jan	3,510	58.3	6,000
61.5	4,010	2,160	1,380	890	540	452	219	1950-51	11 Feb	3,660	61.7	5,450
64.8	3,870	2,120	1,380	865	520	407	214	1951-52	1 Dec	4,930	65.0	5,440
68.1	3,870	1,810	1,300	745	480	392	202	1951-52	1 Feb	4,650	68.3	5,280
71.4	3,310	1,730	1,300	700	410	380	198	1952-53	27 Apr	3,070	71.7	5,050
74.7	3,230	1,720	1,260	630	402	350	195	1953-54	13 Feb	3,300	75.0	5,020
78.0	3,220	1,580	902	584	400	330	192	1953-54	4 Apr	4,240	78.3	4,950
81.3	2,630	1,540	890	515	373	325	170	1955-56	7 Jan	5,020	81.7	4,930
84.6	2,480	1,200	719	480	335	290	166	1955-56	22 Feb	6,480	85.0	4,910
87.8	1,520	1,060	650	386	330	280	156	1957-58	26 Jan	3,060	88.3	4,700
91.1	1,500	965	565	377	329	260	150	1957-58	12 Feb	4,330	91.7	4,650
94.4	1,260	662	472	350	292	225	150	1957-58	21 Mar	4,540	95.0	4,600
97.7	1,080	594	453	315	235	186	121	1957-58	2 Apr	3,970	98.3	4,540

Comp. by: R.P.L. Date: May 1959 Sheet: Station: USGS 11-3810

SPK FORM 6 MAY 59 237a

Figure 5.02

MAXIMUM-RUNOFF VOLUME FREQUENCY DATA - LOGARITHMS

Logarithms of Average Flow in c.f.s.

Station 1: Mill Creek nr. Los Molinos, Calif. Station 2 (base sta.): Feather River at Bidwell Bar, Calif.

Water year (1)	Peak Sta. 1 (2)	Peak Sta. 2 (3)	1-day Sta. 1 (4)	1-day Sta. 2 (5)	3-day Sta. 1 (6)	3-day Sta. 2 (7)	10-day Sta. 1 (8)	10-day Sta. 2 (9)	30-day Sta. 1 (10)	30-day Sta. 2 (11)	90-day Sta. 1 (12)	90-day Sta. 2 (13)	Water-year Sta. 1 (14)	Water-year Sta. 2 (15)
1928-29	3.18		2.98	3.67	2.86	3.50	2.58	3.45	2.47	3.39	2.35	3.26	2.18	2.87
29-30	3.78		3.61	4.33	3.38	4.20	3.10	4.00	2.77	3.68	2.68	3.60	2.41	3.27
30-31	3.18		3.08	3.61	2.81	3.56	2.50	3.38	2.37	3.23	2.27	3.07	2.08	2.72
31-32	3.74		3.33	3.94	3.14	3.89	3.02	3.76	2.68	3.67	2.58	3.64	2.33	3.22
32-33	3.03		2.82	3.64	2.67	3.63	2.59	3.56	2.52	3.44	2.45	3.34	2.18	2.90
33-34	3.42		3.24	3.88	3.11	3.76	2.87	3.49	2.53	3.36	2.46	3.28	2.22	2.92
34-35	3.60		3.36	4.33	3.17	4.21	2.98	4.06	2.88	3.96	2.75	3.75	2.46	3.29
35-36	3.64		3.42	4.33	3.28	4.26	3.09	4.04	2.86	3.67	2.71	3.71	2.43	3.30
36-37	3.52		3.19	3.89	2.96	3.83	2.77	3.74	2.72	3.68	2.68	3.60	2.34	3.14
37-38	4.36		4.09	4.86	3.87	4.64	3.44	4.23	3.04	4.05	2.95	3.99	2.75	3.61
38-39	3.10		2.77	3.48	2.66	3.47	2.54	3.45	2.52	3.41	2.41	3.22	2.19	2.84
39-40	4.06		3.88	4.72	3.79	4.57	3.41	4.33	3.08	3.98	2.96	3.89	2.58	3.42
40-41	4.09		3.78	4.36	3.68	4.26	3.30	4.02	3.12	3.90	2.78	3.77	2.67	3.45
41-42	4.04		3.76	4.44	3.51	4.30	3.26	4.03	3.06	3.96	2.92	3.75	2.65	3.47
42-43	3.84		3.58	4.43	3.49	4.40	3.22	4.17	2.91	3.89	2.82	3.82	2.57	3.42
43-44	3.51		3.24	3.77	2.95	3.73	2.71	3.69	2.57	3.57	2.52	3.45	2.30	3.04
44-45	3.51		3.20	4.33	3.10	4.19	2.95	3.97	2.73	3.72	2.59	3.57	2.23	3.24
45-46	3.79		3.49	4.28	3.37	4.21	3.25	4.10	2.95	3.83	2.67	3.54	2.47	3.28
46-47	3.61		3.41	4.08	3.14	3.96	2.80	3.66	2.60	3.55	2.54	3.43	2.28	3.02
47-48	3.86		3.56	4.18	3.31	4.05	2.94	3.93	2.89	3.81	2.83	3.66	2.49	3.22
48-49	3.59		3.26	3.77	3.11	3.76	2.84	3.74	2.61	3.66	2.61	3.52	2.29	3.05
49-50	3.65		3.51	4.18	3.37	4.08	3.00	3.84	2.77	3.74	2.66	3.64	2.40	3.26
50-51	3.59		3.33	4.51	3.19	4.40	3.02	4.12	2.92	4.02	2.75	3.81	2.54	3.44
51-52	3.72		3.48	4.23	3.37	4.20	3.10	4.15	2.93	4.11	2.86	3.97	2.65	3.58
52-53	3.89		3.72	4.52	3.50	4.29	3.28	4.08	3.03	3.87	2.76	3.61	2.54	3.35
53-54	3.69		3.36	4.29	3.27	4.19	3.00	3.91	2.87	3.77	2.80	3.67	2.50	3.23
54-55	3.39		3.03	3.72	2.75	3.70	2.68	3.64	2.60	3.51	2.51	3.36	2.31	2.99
55-56	3.96		3.83	4.91	3.70	4.78	3.51	4.52	3.27	4.23	3.03	3.95	2.69	3.57
56-57	3.79		3.58	4.37	3.42	4.33	3.13	4.08	2.91	3.90	2.72	3.67	2.41	3.20
57-58	3.84		3.55	4.43	3.40	4.30	3.19	4.13	3.11	3.98	2.96	3.88	2.69	3.50

Comp. by: RPL Date: May 1959 Sheet: Station: USGS 11-3810

SPK FORM 6 MAY 59 237b

Figure 5.03

MAXIMUM-RUNOFF VOLUME FREQUENCY COMPUTATIONS

Station 1: Mill Creek nr Los Molinos, Calif. Station 2 (base sta.) Feather River at Bidwell Bar, Calif.

(32) $R^2=(\Sigma x_1 x_2)^2/\Sigma x_1^2 \Sigma x_2^2$ $(1-\bar{R}^2)=(1-R^2)(N-1)/(N-2)$ $S_1'-S_1=(S_2'-S_2)R^2(S_1/S_2)$ $\bar{X}_1'-\bar{X}_1=(\bar{X}_2'-\bar{X}_2)R(S_1'/S_2')$ $N_1'=N_1/[1-(N_2'-N_1)\bar{R}^2/N_2']$

		Peak		1-day		3-day		10-day		30-day		90-day		Water-year	
		Sta. 1	Sta. 2	Sta. 1	Sta. 2	Sta. 1	Sta. 2	Sta. 1	Sta. 2	Sta. 1	Sta. 2	Sta. 1	Sta. 2	Sta. 1	Sta. 2
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
(33)	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30
(34)	ΣX	109.97	Use 1-day values at Sta. 1 as base (assume $\bar{R}$ = 1)	102.44	125.48	97.33	122.65	90.07	117.27	84.29	112.54	80.58	108.42	72.83	96.81
(35)	$\bar{X}$	3.666		3.415	4.183	3.244	4.088	3.002	3.909	2.810	3.751	2.686	3.614	2.428	3.227
(36)	ΣX^2	405.781		352.561	528.884	318.688	504.802	272.605	460.721	238.321	423.907	217.513	393.424	177.758	314.038
(37)	$(\Sigma X)^2/N$	403.113		349.798	524.841	315.771	501.434	270.420	458.408	236.827	422.175	216.438	391.830	176.807	312.406
(38)	Σx^2	2.668		2.763	4.043	2.917	3.368	2.185	2.313	1.494	1.732	1.075	1.594	.951	1.632
(39)	$\Sigma x^2/N-1$	.0920		.0953	.1394	.1006	.1161	.0753	.0798	.0515	.0597	.0371	.0550	.0328	.0563
(40)	S	.303		.309	.373	.317	.341	.274	.283	.227	.244	.193	.234	.181	.237
(41)	$\Sigma X_1 X_2$			431.454		400.758		354.161		317.701		292.456		236.206	
(42)	$\Sigma X_1 \Sigma X_2/N$			428.472		397.918		352.084		316.200		291.216		235.022	
(43)	$\Sigma x_1 x_2$			2.982		2.840		2.077		1.501		1.240		1.184	
(44)	R^2			.796		.821		.854		.871		.897		.903	
(45)	$\bar{R}^2$	1.00		.789		.815		.849		.866		.893		.900	
(46)	S'−S	−.022	−.022	−.022	−.034	−.020	−.027	−.015	−.019	−.008	−.010	−.002	−.003	−.003	−.004
(47)	S'	.281		.287	.339	.297	.314	.259	.264	.219	.234	.191	.231	.178	.233
(48)	S'(Adj.)	.288		.288		.288		.254		.277		.210		.172	
(49)	$\bar{X}'-\bar{X}$	−.005	−.005	−.005	−.007	−.001	−.002	0	0	.003	+.004	−.001	−.001	−.005	−.007
(50)	$\bar{X}$	3.661	3.410	3.410	4.176	3.243	4.086	3.002	3.909	2.813	3.755	2.685	3.613	2.423	3.220
(51)	N'	43	43	43	47	44	47	44	47	45	47	45	47	45	47
(52)	g	.00		−.04		−.12		−.23		−.32		−.37		−.40	

	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)	(29)	(30)	(31)
(53)	P_∞	$P_{N'}$	Log Q	Q	Log Q	Q	Log Q	Q	Log Q	Q	Log Q	Q	Log Q	Q	Log Q	Q
(54)	.01	.031	4.735	54,300	4.461	28,900	4.245	17,600	3.830	6,760	3.512	3,250	3.311	2,050	2.925	841
(55)	.1	.19	4.551	35,600	4.283	19,200	4.084	12,100	3.706	5,080	3.421	2,640	3.227	1,690	2.860	724
(56)	1.0	1.32	4.332	21,500	4.072	11,800	3.888	7,730	3.551	3,560	3.287	1,940	3.116	1,310	2.772	592
(57)	10.0	10.6	4.030	10,700	3.776	5,970	3.606	4,040	3.320	2,090	3.092	1,240	2.941	873	2.633	430
(58)	50.0	50.0	3.661	4,580	3.413	2,590	3.249	1,770	3.010	1,020	2.824	667	2.698	499	2.433	271
(59)	90.0	89.0	3.292	1,960	3.041	1,100	2.871	743	2.672	470	2.516	328	2.408	256	2.196	157
(60)	99.0	98.68	2.990	977	2.730	537	2.546	352	2.367	233	2.232	171	2.139	138	1.972	94
(61)	99.9	99.81	2.771	590	2.503	318	2.304	201	2.133	136	2.007	102	1.923	84	1.792	62
(62)	99.99	99.969	2.587	386	2.310	204	2.094	124	1.928	85	1.810	65	1.734	54	1.632	43

Period of record: 1928-58 Base station period: 1911-58

By: R.P.L. Date: May 1959 Station No. USGS 11-3810

SPK FORM 237c 5 FEB 60 (Replaces SPK Form 237c dtd 6 May 59)

COMPUTATION PROCEDURE

1. Annual-event Data. For each water year or seasonal portion of a water year at a given site, tabulate to three significant figures the maximum unregulated runoff volume in average-flow units for each duration with date as indicated on fig. 5.01. Long-duration volumes need not cover the time period of the corresponding short-duration volumes. It is usually best to restrict the tabulation to the same years for all durations.

2. Partial-duration Data. If a partial-duration curve of peak or daily flows is also desired, tabulate all additional peak (or daily) flows exceeding a selected non-damage base value and separated from larger flows by a minimum specified time period, as indicated on fig. 5.02. Unless doing so would require discarding damaging flows, the base should be selected or adjusted to provide a total of N (number of years) events, including annual-event flows exceeding that base. A value exceeded by two-thirds of the annual-event flows is almost always low enough for a first approximation of the base.

3. Plotting of Points. Arrange each set of annual-event flows and the set of partial-duration flows (including the corresponding annual-event flows) in the order of magnitude and tabulate plotting positions from Exhibit 1 as shown on fig. 5.02. Plot all the above data on logarithmic probability grid as shown on fig. 5.06.

4. Logarithms. Tabulate common logarithms to two decimal places for all tabulated annual-event flows at the site (Station 1) and for corresponding annual maximum flows at a selected long-record base station (Station 2) as shown on fig. 5.03.

5. Sums, Squares and Cross Products. Logarithms are designated X_1 for Station 1 and X_2 for Station 2. Compute the sum (ΣX, line 34) and the sum of squares (ΣX^2, line 36) of logarithms in each column. Compute the sum of cross products of Station 1 and Station 2 logarithms for each duration (ΣX_1X_2, line 41). Tabulate number of items (N, line 33) of data used.

6. Mean ($\overline{X}$) and Standard Deviation (S). For each column, compute the mean ($\overline{X}$, line 35) from $\overline{X} = \Sigma X/N$ and the standard deviation (S, line 40) from

$$S = \left[\frac{\Sigma X^2 - (\Sigma X)^2/N}{N - 1}\right]^{1/2}$$

. To compute S: square ΣX, divide by N, and enter on line 37. Subtract this value from ΣX^2 to obtain Σx^2 and enter on line 38. Divide Σx^2 by (N-1) and enter on line 39. Take the square root to obtain S and enter on line 40.

7. Determination Coefficient ($\overline{R}^2$). For each duration, compute the determination coefficient between corresponding site and base-station logarithms by use of the first two equations on line 32 as follows: Multiply ΣX_1 by ΣX_2, line 34, divide by N and enter on line 42. Subtract this from ΣX_1X_2 to obtain Σx_1x_2 and enter on line 43. Square Σx_1x_2 and divide by the product of Σx_1^2 and Σx_2^2, line 38, to obtain R^2 and enter on line 44. Subtract R^2 from 1.0 and multiply by the ratio (N-1)/(N-2). Subtract the result from 1.0 to obtain $\overline{R}^2$ and enter on line 45.

8. Extended Statistics $\overline{X}'$, S', and N'. For each duration, compute the long-period statistics by use of the last three equations in line 32 as follows. In the Station 2 column, enter the long-period mean, $\overline{X}_2'$, on line 50, the long-period standard deviation, S_2', on line 47, and the long-period number of events, N_2', on line 51.

To compute S_1': subtract S_2 from S_2' and enter on line 46. Multiply by R^2 and by the ratio S_1/S_2, line 40, and enter the result, $S_1' - S_1$, in the Station 1 column on line 46. Add algebraically to S_1 to obtain S_1' and enter on line 47.

To compute $\overline{X}_1'$: subtract $\overline{X}_2$ from $\overline{X}_2'$ and enter on line 49. Multiply this difference by R and by the ratio S_1'/S_2', line 47, and enter the result, $\overline{X}_1' - \overline{X}_1$, on line 49. Add algebraically to $\overline{X}_1$ and enter result, $\overline{X}_1'$, on line 50.

To compute N_1': Multiply the difference between N_2' and N_1 by $\overline{R}^2$ and divide by N_2. Subtract results from 1.0 and divide N_1 by this value.

9. Adjusted Standard Deviation (S' Adj). Plot the extended standard deviation S_1' against the extended mean $\overline{X}_1'$ for each duration as illustrated on fig. 5.05 and draw a straight or broken line to smooth the S' values. This will assure consistency among frequency curves for the various durations. Do not exceed the maximum indicated slope. Tabulate adjusted S_1' values on line 48.

10. Skew Coefficients (g). Select a skew coefficient for each duration from Section 5.03 or from a regional study. Tabulate in the Station 1 columns on line 52.

11. Computation of Curves. For each duration, compute the logarithm of flow for selected values of P∞ by the equation, $\log Q = \overline{X} + kS$. The values of P∞ shown in column 16 are ordinarily satisfactory. Obtain corresponding values of k from Exhibit 2 for each value of g. Multiply each k value in turn by S_1' (adj), line 48, and add to $\overline{X}_1'$ to obtain log Q. Tabulate in log-Q column. Tabulate corresponding antilogarithms (Q) in the column headed Q. For each value of P∞, tabulate the corresponding value of P_N from Exhibit 3, using the average value of N_1' for all durations.

12. Plotting of Curves. Plot flows, Q, against corresponding values of P_N for each duration and draw smooth curves as shown on fig. 5.06. Check curves against points plotted in step 3 for possible arithmetic mistakes. If there is a radical departure of points from the curves, examine the hydrologic features of the basin for possible special treatment as described in Section 4.05.

Figure 5.04

e. Plot the extended standard deviation against the extended mean, and adopt a smooth relationship, as shown on fig. 5.05. Tabulate these values on fig. 5.04.

f. When many peak flows are missing from the record, it is best to add the average difference between corresponding peak and 1-day logarithms to the 1-day mean in order to obtain the peak mean logarithm. The peak standard deviation can then be obtained by extrapolation of the curve similar to that shown on fig. 5.05.

g. Select a skew coefficient for each duration from the above tabulation or from special regional studies. Tabulate on fig. 5.04.

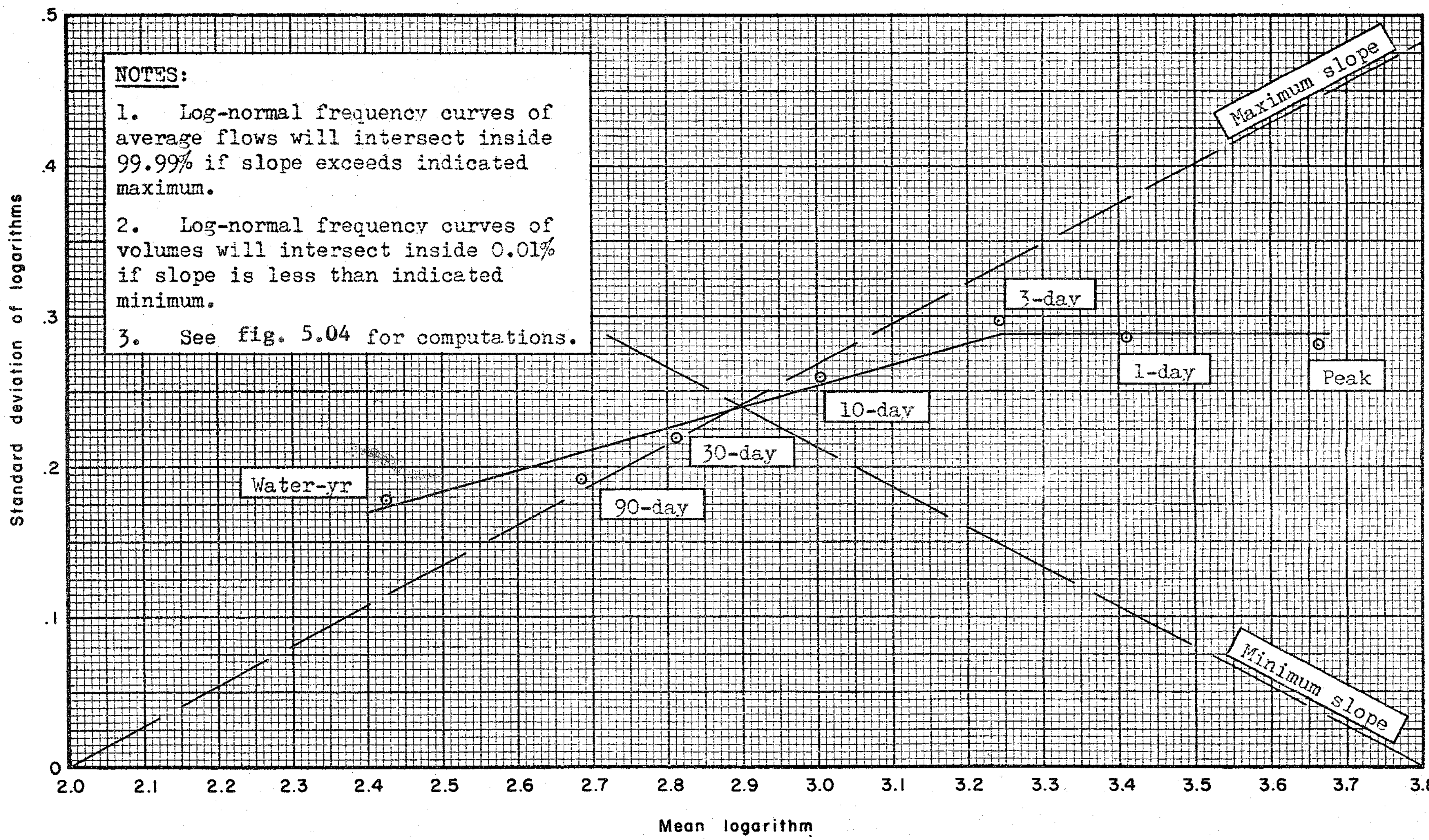
Section 5.04. Frequency Curves for Comprehensive Series

Frequency curves of flood volumes are computed analytically using general principles and methods of Chapter 4. They should also be shown graphically and compared with the data on which they were based. This is a general check on the analytic work and will ordinarily reveal any inconsistency in data and methodology. Data are plotted on a single sheet for comparison purposes, using procedures described in Section 4.01.

Frequency curves are obtained from the frequency statistics and compared with observed frequencies for each of the seven basic durations as follows:

a. Tabulate the average flows for each duration in the order of magnitude and obtain the plotting position for each event from Exhibit 1. This operation is shown on fig. 5.02.

b. Compute flows corresponding to related values of P_∞ for each frequency curve from the adjusted means and smoothed standard deviations, using equation 4-5 and coefficients obtained from Exhibit 2, as shown on fig. 5.04.



STANDARD DEVIATION ADJUSTMENT

Figure 5.05

c. Tabulate values of P_N from Exhibit 3 for each selected value of P_∞, using the average value of N' for all durations.

d. Plot the points obtained in step a and the curves from coordinates obtained in steps b and c as shown on fig. 5.06.

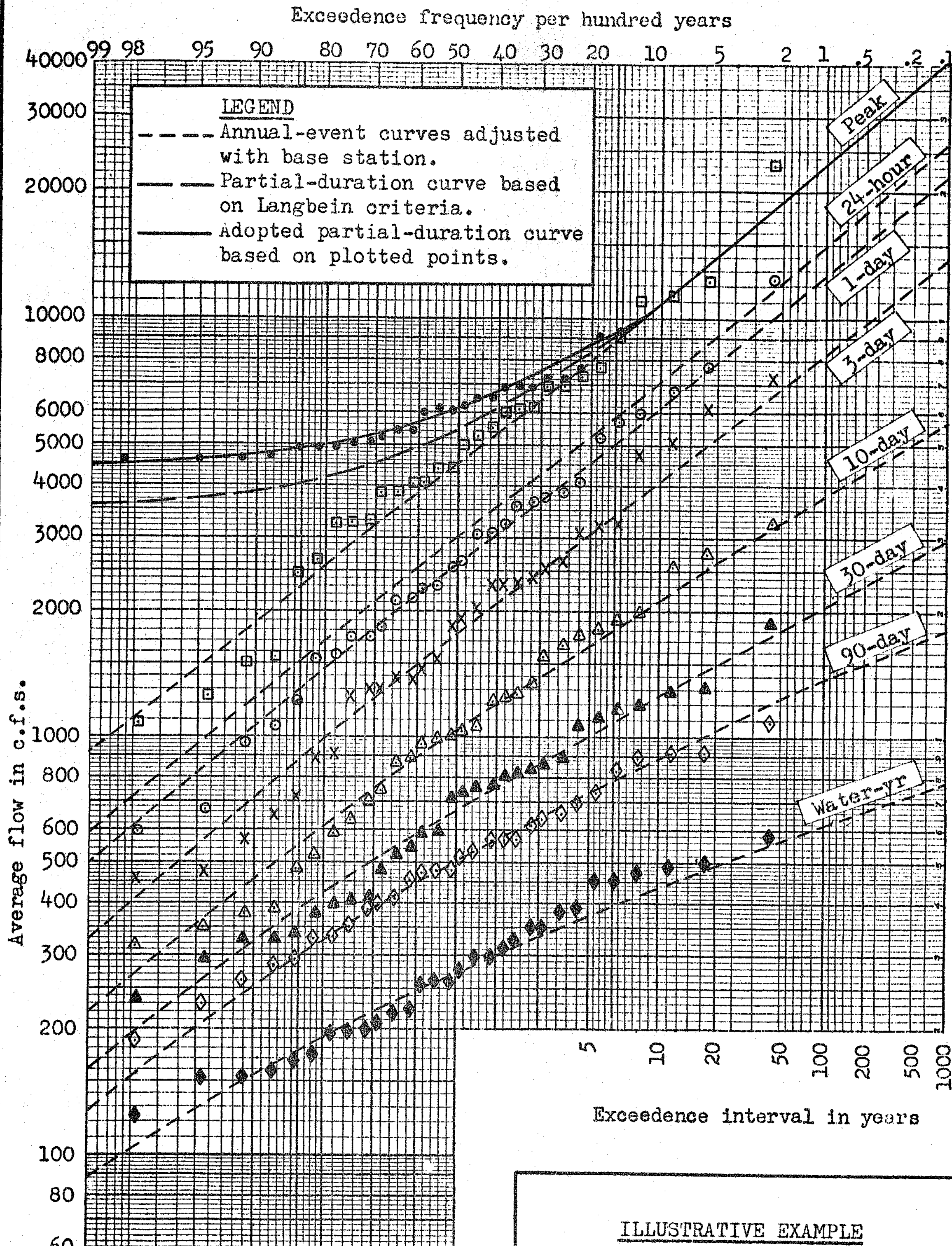
The runoff volume for any specified frequency can be determined for any duration between 24 hours and 1 year by drawing a curve on logarithmic paper as illustrated on fig. 5.07, relating volume to duration for that specified frequency, using maximum 24-hour criteria summarized in terms of average flows as follows:

$$\log Q_{24\text{-hr}} = 0.77 \log Q_{1\text{-day}} + 0.23 \log Q_{peak} \qquad (5\text{-}1)$$

When runoff volumes for durations shorter than 24 hours are very important, special frequency studies should be made. These could be done in the same manner as for the longer durations, using skew coefficients interpolated in some reasonable manner between those used for peak and 1-day flows.

Section 5.05. Durations Exceeding 1 Year

In the design of reservoirs for conservation purposes (and occasionally for flood-control purposes), the volumes of runoff that can be expected to occur during the lifetime of the structure for durations exceeding 1 year can be of primary concern. In general, the design of such reservoirs has been based on the lowest (or highest) volume of runoff observed for the critical duration during the period of record, which generally encompasses 40 to 100 years. Frequently, however, it is felt that the observed minimum or maximum is much more extreme than should normally be anticipated in a similar future period. In order to determine expected volumes with a greater degree of reliability,



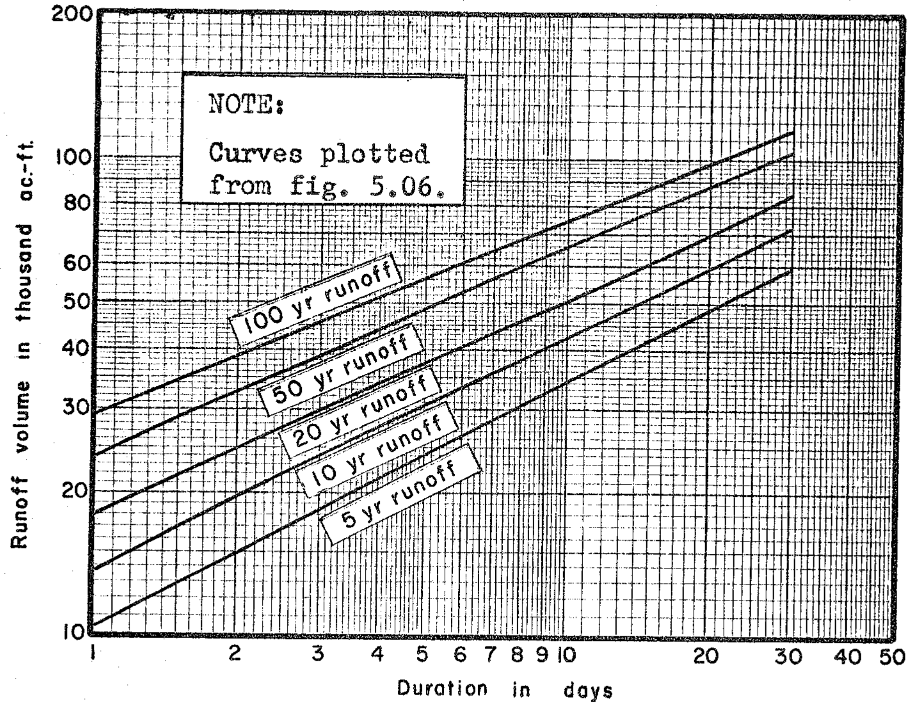
NOTES: 1. Annual-event curves adjusted using record of Feather River at Bidwell Bar.

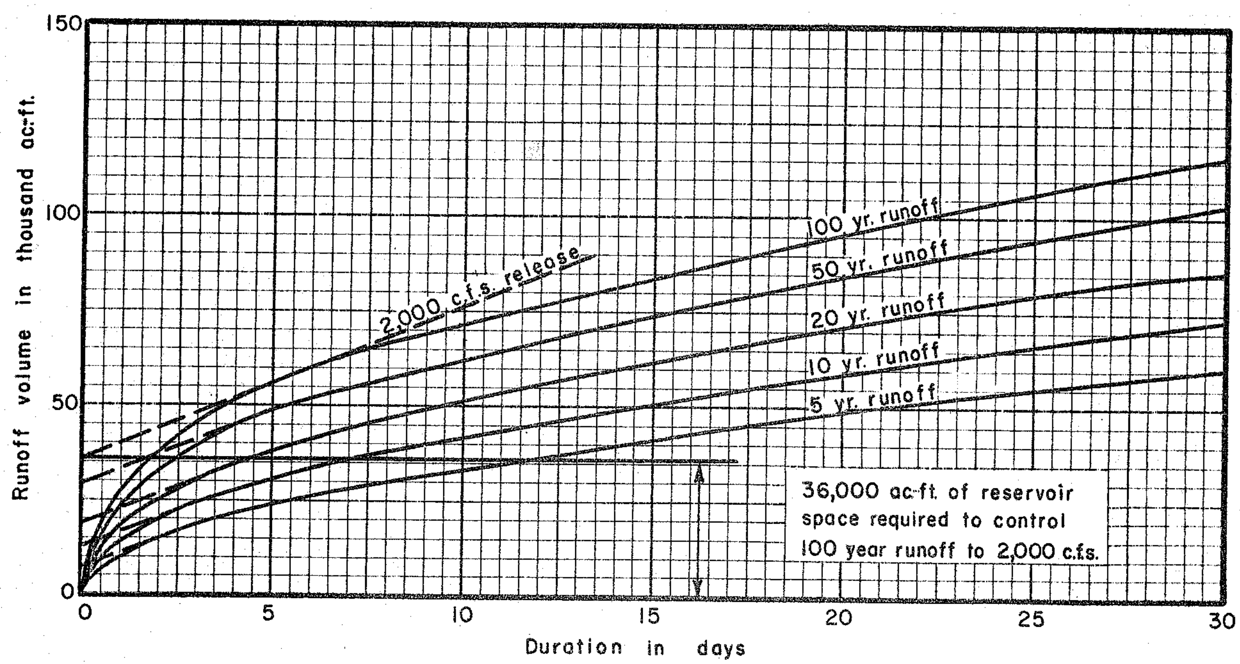
2. See fig. 5.04 for supporting computations.

ILLUSTRATIVE EXAMPLE

MAX-RUNOFF VOLUME FREQUENCY CURVES
MILL CREEK NR LOS MOLINOS

Figure 5.06





ILLUSTRATIVE EXAMPLE

VOLUME - DURATION CURVES
MILL CR NEAR LOS MOLINOS, CALIFORNIA

Figure 5.07

frequency studies of long-duration volumes can be made. However, direct frequency analysis is ordinarily not practicable, because a relatively small number of independent long-duration volumes is contained in a single record. (Only 20 independent 5-year volumes are contained in a 100-year record, for example.) In order to overcome this limitation, a study has been made to relate long-duration volumes to annual volumes. Using criteria developed in this study, long-duration volumes can be derived from a frequency curve of annual volumes and the correlation coefficient between successive annual flows. Where this correlation coefficient is considered to be zero, the criteria should be fairly dependable, but where there appears to be substantial correlation between successive annual runoff values, the reliability decreases, principally because of the uncertainty as to the true correlation coefficient. An alternative procedure is to generate several long sequences of streamflows using techniques described in Chapter 5 of Volume 2 of this report. The statistics needed to compute frequency curves for long duration volumes may then be calculated from the generated streamflow.

Criteria for determining multi-annual runoff are based on the logarithmic mean and standard deviation of annual runoff. These should be computed as described in Section 5.03.

In many river basins, surface or subsurface storage effects cause the flows in 1 year to reflect conditions in the preceding year to some extent. This will result in a positive correlation between the runoff in successive years. A measure of this persistence effect is the correlation coefficient, or preferably its square, the determination coefficient, between successive years' runoff logarithms. This is determined by pairing each year's runoff logarithm (except the first) with that of the preceding year, and computing the determination coefficient by use of equations 4-7 and 4-8. Because of the important effect of the determination coefficient on long-duration volume estimates, its degree of unreliability as discussed in Section 4.04 should be given special consideration.

It must be recognized that multi-annual runoff based on water-year volumes is not representative of critical conditions. A drier or wetter period of the same duration can usually be found by starting the period some days or months earlier or later. Also, an integral number of years does not represent a critical duration, because adding one more dry season (or wet season) will ordinarily worsen the condition.

Section 5.06. Applications of Flood Volume-Duration Frequencies

The use of flood volume-duration frequencies in solving reservoir planning, design, and operation problems usually requires the construction of volume-duration curves for specified frequencies. These are drawn first on logarithmic paper for interpolation purposes as illustrated on fig. 5.07, and are then replotted on arithmetic grid as shown on the same figure. A straight line on this grid represents a constant rate of flow (so many acre-feet per day). The straight lines on fig. 5.07 represent a uniform flow of 2,000 c.f.s., and placement on the 100-year volume-duration curve demonstrates that a reservoir capacity of 36,000 acre-feet is required to control the indicated runoff volumes to a project maximum release of 2,000 c.f.s. The curve also indicates that durations of 4 to 7 days are critical for this project release and flood-control space.

In the case of a single flood-control reservoir located immediately upstream of the only important damage center concerned, the volume frequency problems are relatively simple. A series of volume-duration curves similar to that shown on fig. 5.07 corresponding to selected frequencies should first be drawn. The project release rate should be determined, giving due consideration to possible channel deterioration, encroachment into the flood plain, and operational contingencies. Lines representing this flow rate are then drawn tangent to each volume-duration curve, and the intercept in each case determines the reservoir space used to control the flood of that selected frequency. The point of tangency represents the critical duration of runoff. This procedure can be used not only as an approximate aid in selecting a reservoir capacity, but as an aid in preparing filling-frequency curves.

Where reservoir operation schedules are variable or where many reservoirs are operated jointly, it may be necessary to route historical flows month by month or day by day in order to demonstrate the adequacy of a design or operation procedure. However, the techniques described in the preceding paragraph may be applicable approximately, and may shed considerable light on the problem. In applying such techniques, the following guides should be used.

a. Volume-duration curves are needed for unregulated flows at each important damage center.

b. The straight line corresponding to the average non-damaging flow, allowing for operational contingencies, when drawn tangent to the volume-duration curve corresponding to the selected design frequency will indicate the storage required in the system if it is located and operated so as to be fully effective.

c. The same straight line will indicate the range of critical durations for design and operation studies. A system of reservoirs should be "tuned" to this range of durations, insofar as is feasible, because a reservoir that fills and empties in 5 days may be of no value if the critical duration at a downstream damage center is 15 days. Likewise, a reservoir that is only half full in 15 days would not have provided its best control at the damage center.

In solving complex reservoir problems, representative hydrographs at all locations can be patterned after one or more past floods. The ordinates of these hydrographs can be adjusted so that their volumes for the critical durations will equal corresponding magnitudes at each location for the selected frequency. A design or operation scheme based on regulation of such a set of hydrographs would be reasonably well balanced. Computation techniques for developing representative hydrographs are described in Section 6.08.

Section 5.07. Use of Computer Programs

A computer program has been developed in The Hydrologic Engineering Center for the purpose of computing frequency statistics for one or more stations in a region. The program, Regional Frequency Computation, is designed to accept annual events for one or more durations at one or more locations, fill in missing data, compute and adjust statistics, arrange data for plotting and compute frequency curves in accordance with procedures described herein. The program is described in detail in reference 33. The advantage of estimating missing values is that, for an incomplete record, the observed values can be plotted and will be in agreement with the computed and adjusted frequency curve. In presenting the data in reports, care should be exercised to identify estimated values for missing data. Estimated values need not be plotted, but the observed values can be plotted so as to leave vacant spaces for estimated values and thus agree with the frequency curve based on all data. It is important to recognize that values estimated using this program contain a random component, which is needed in order to preserve the variance and correlation in the observed data, thus overcoming the objection of using standard regression techniques cited earlier.

Chapter 6

Development Effects On Flood Frequencies

CHAPTER 6. DEVELOPMENT EFFECTS ON FLOOD FREQUENCIES

Section 6.01. Nature of the Problem

Flood control works are designed to substantially alter the frequency of flood flows or flood stages or both at various locations. Many developments such as urbanization or cultivation of large areas can also have significant effects on flood flows. Channel improvements intended to reduce stages, and levee improvements intended to confine flows, can substantially affect downstream flows by eliminating natural storage effects. The degree to which flows are modified depends on the magnitude, time, and areal distribution of rainfall (and snowmelt, if pertinent) causing the flood. Accordingly, evaluations should include studies of effects of proposed works on different flood magnitudes.

Section 6.02. Terminology

Natural Conditions. Natural conditions in the drainage basin are defined as hydrologic conditions that would prevail if no regulatory works or other works of man were constructed. Natural conditions, however, include such effects as those of natural lakes and swamp areas.

Present Conditions. Present conditions are defined as the conditions that exist as of the date of the study report.

Without-project Conditions. Without-project conditions are defined as the conditions that would prevail if the project under consideration were not constructed but with all existing projects and future projects having a higher priority of construction assumed to exist.

With-project Conditions. With-project conditions are defined as the conditions that will exist after the project is completed and after completion of all projects having an equal or higher priority of construction.

Section 6.03. Reservoir Level Frequency Computation

The factors affecting the frequency of reservoir stage include anticipated future inflow rates, the plan of operation of reservoir outlets and spillway, storage level assignments, rate of sedimentation, and the nature of the storage elevation curves.

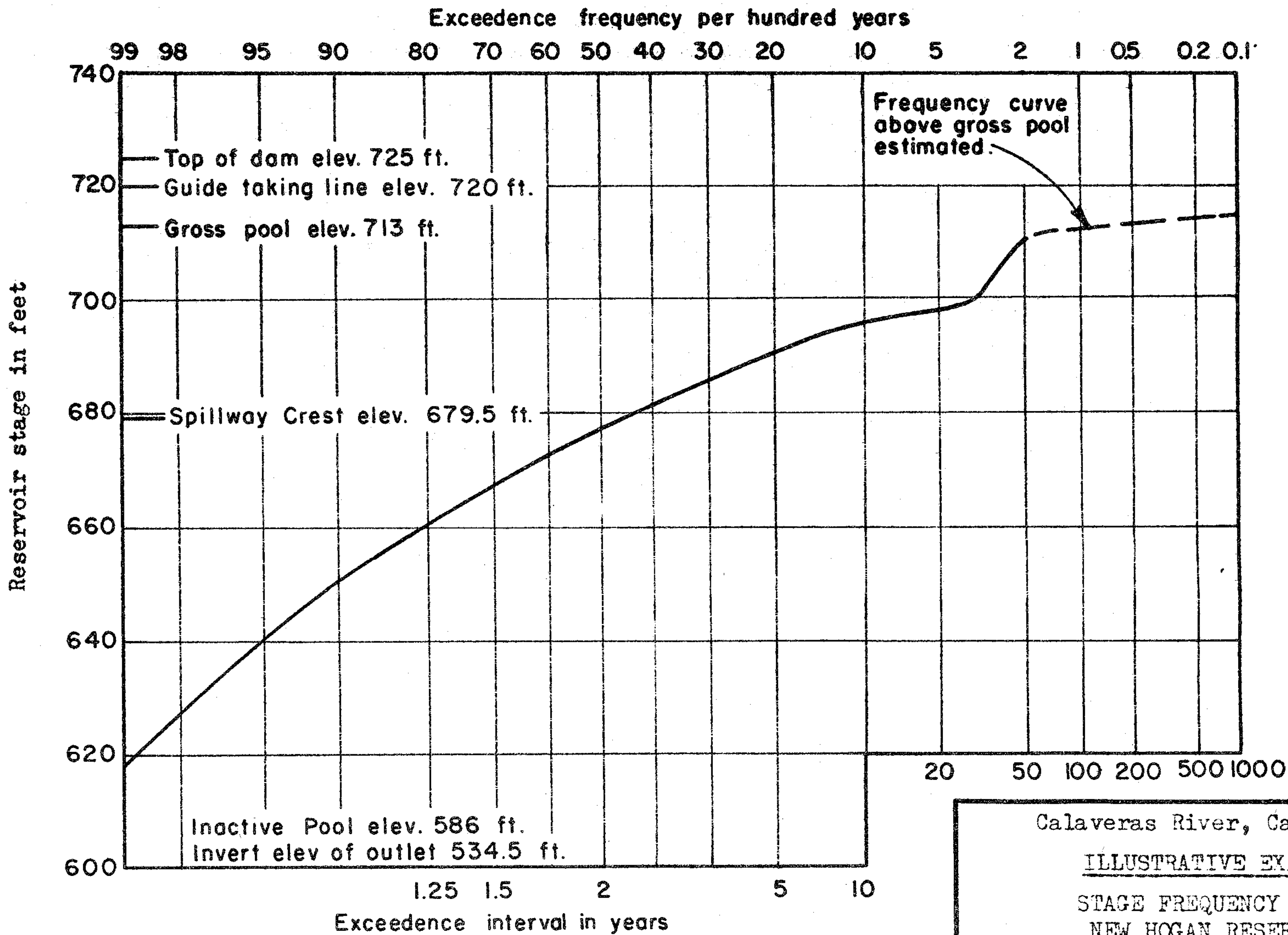
A reservoir frequency curve of annual maximum storages is ordinarily constructed graphically, using procedures outlined in Section 4.01, by the use of observed storages to the extent that these are available (but only if the reservoir has been operated in the past in accordance with future plans) and by the use of routings of historical runoff and of large hypothetical floods under future operating plans. This storage curve is then converted to a stage curve. A typical frequency curve is illustrated on fig. 6.01. Stage-duration curves are constructed from historical operation data or from routings (usually monthly) of historical runoff. Such curves may be constructed for the entire period of record or for a selected wet period or dry period as illustrated on fig. 6.02. For many purposes, particularly recreation uses, the seasonal variation of reservoir stages is of importance, and a set of frequency or of duration curves for each month of the year may be valuable. One suggested type of presentation is illustrated on fig. 6.03.

Section 6.04. Effects of Reservoirs on Flows at Downstream Points

The frequency of reservoir outflow or of flows at a remote downstream location can be obtained from hypothetical routings of the entire runoff record in three general ways as follows:

a. Estimating maximum flow for each year at the location concerned and constructing a frequency curve of regulated flows graphically, using procedures discussed in Section 4.01.

b. Routing all historical flood of record through system and plotting up frequency curves using procedures discussed in Section 4.01.



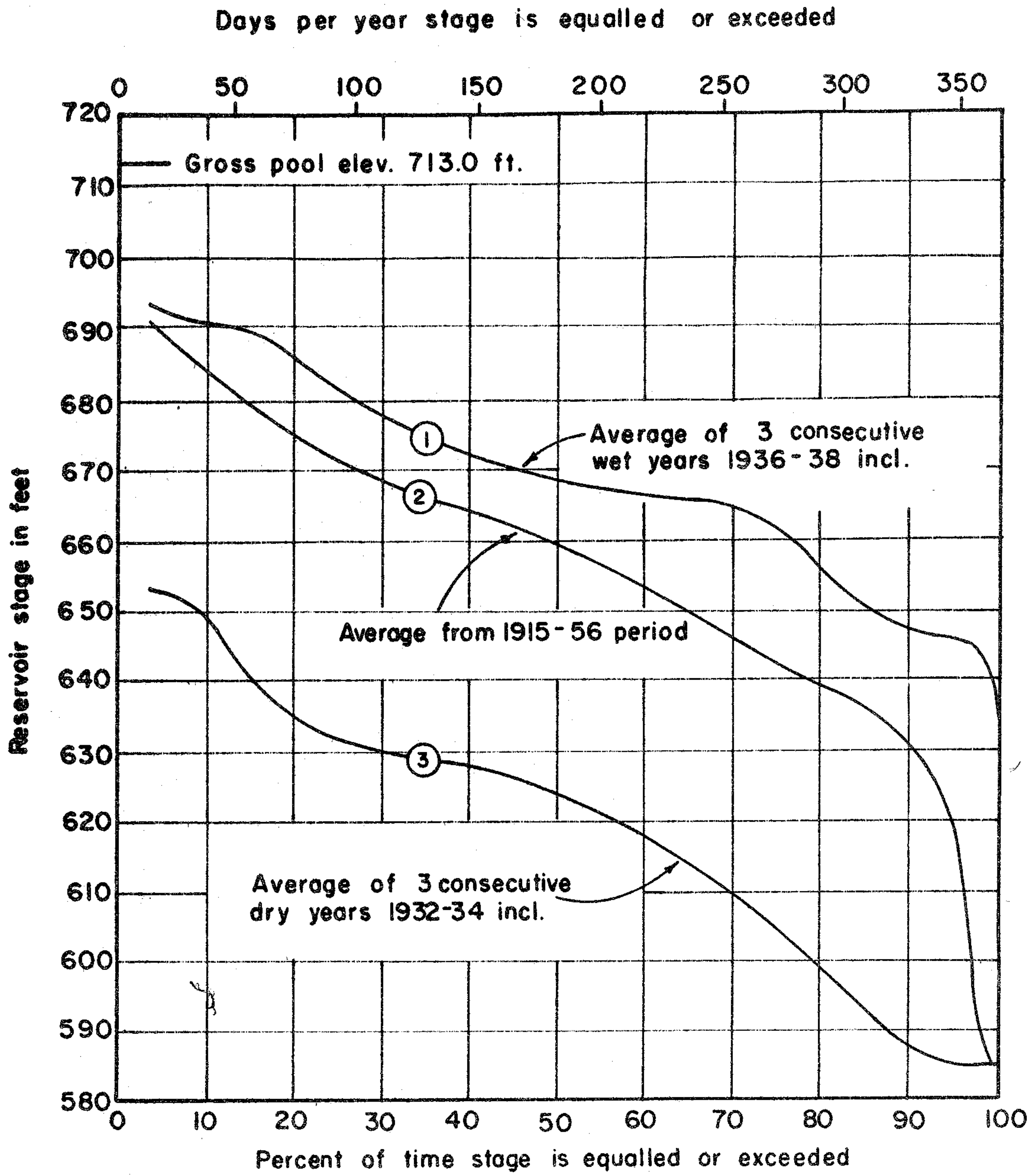
Note: Curve below 700 feet is controlled by conservation operation, and the remainder is controlled by operations during major rain floods.

Calaveras River, California

ILLUSTRATIVE EXAMPLE

STAGE FREQUENCY CURVE
NEW HOGAN RESERVOIR

Figure 6.01

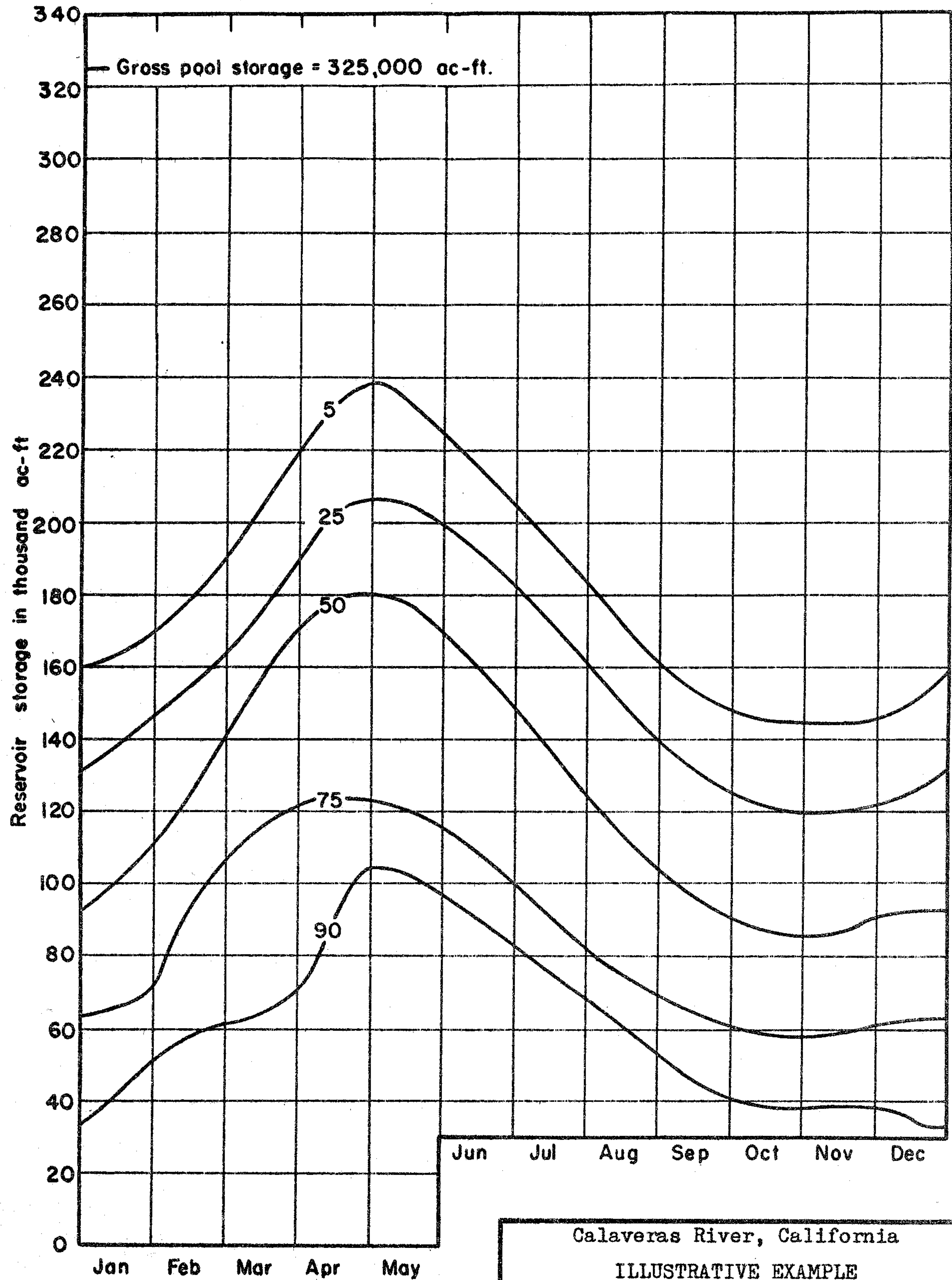


Calaveras River, California

ILLUSTRATIVE EXAMPLE

STAGE-DURATION CURVES
NEW HOGAN RESERVOIR

Figure 6.02



Note:
Indicated value is percentage of years that storage is exceeded on given date. Values computed on monthly basis.

Calaveras River, California

ILLUSTRATIVE EXAMPLE

SEASONAL VARIATION OF
RESERVOIR STORAGE FREQUENCY
NEW HOGAN RESERVOIR

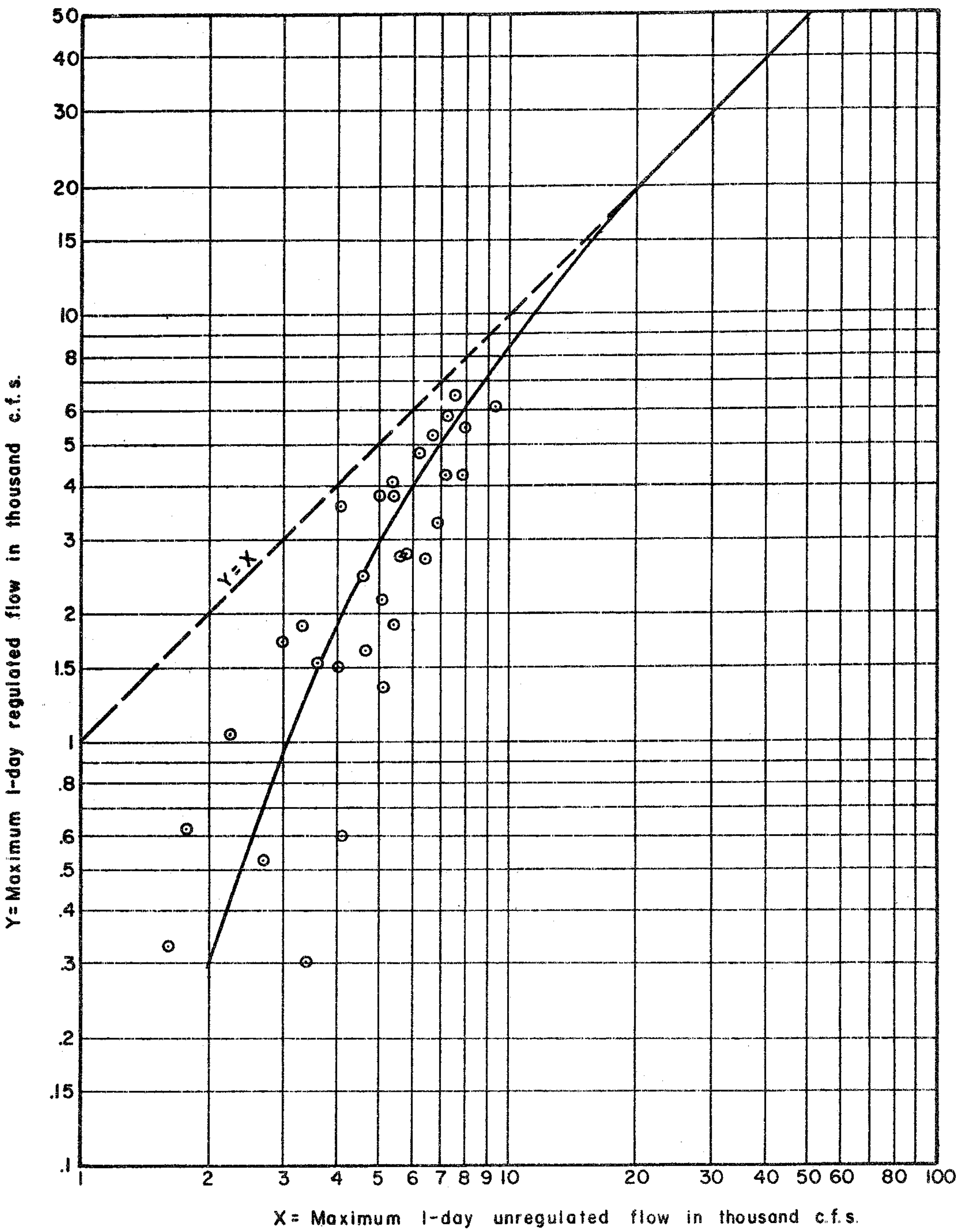
Figure 6.03

c. Constructing a graph of with-project vs. without-project flows at a specified point from the routed data and drawing a curve relating the two quantities as illustrated on fig. 6.04. The points should be balanced in the direction transverse to the curve, but factors such as representativeness of plotted points and reliability of regulation may be considered in drawing the curve. This curve can be used in conjunction with a frequency curve of without-project flows to construct a frequency curve of with-project flows as illustrated on fig. 6.05. This latter procedure assures consistency in the analysis and gives a graphical demonstration of the degree of reliability of control.

Usually recorded values of flows are not large enough to define the frequency of regulated flows at the upper end of the frequency curve. In such cases, it is ordinarily possible to use one or more large hypothetical floods whose frequency can be estimated from the frequency curve of unregulated flows, to establish the corresponding magnitude of regulated flows. These floods can be multiples of the largest observed floods or of floods computed from rainfall, but it is best not to multiply any one flood by a factor greater than two or three. The floods are best selected or adjusted to represent about equal severity, in terms of runoff frequency, of peak and volumes for various durations. The routings should be made under reasonably conservative assumptions as to initial reservoir stages.

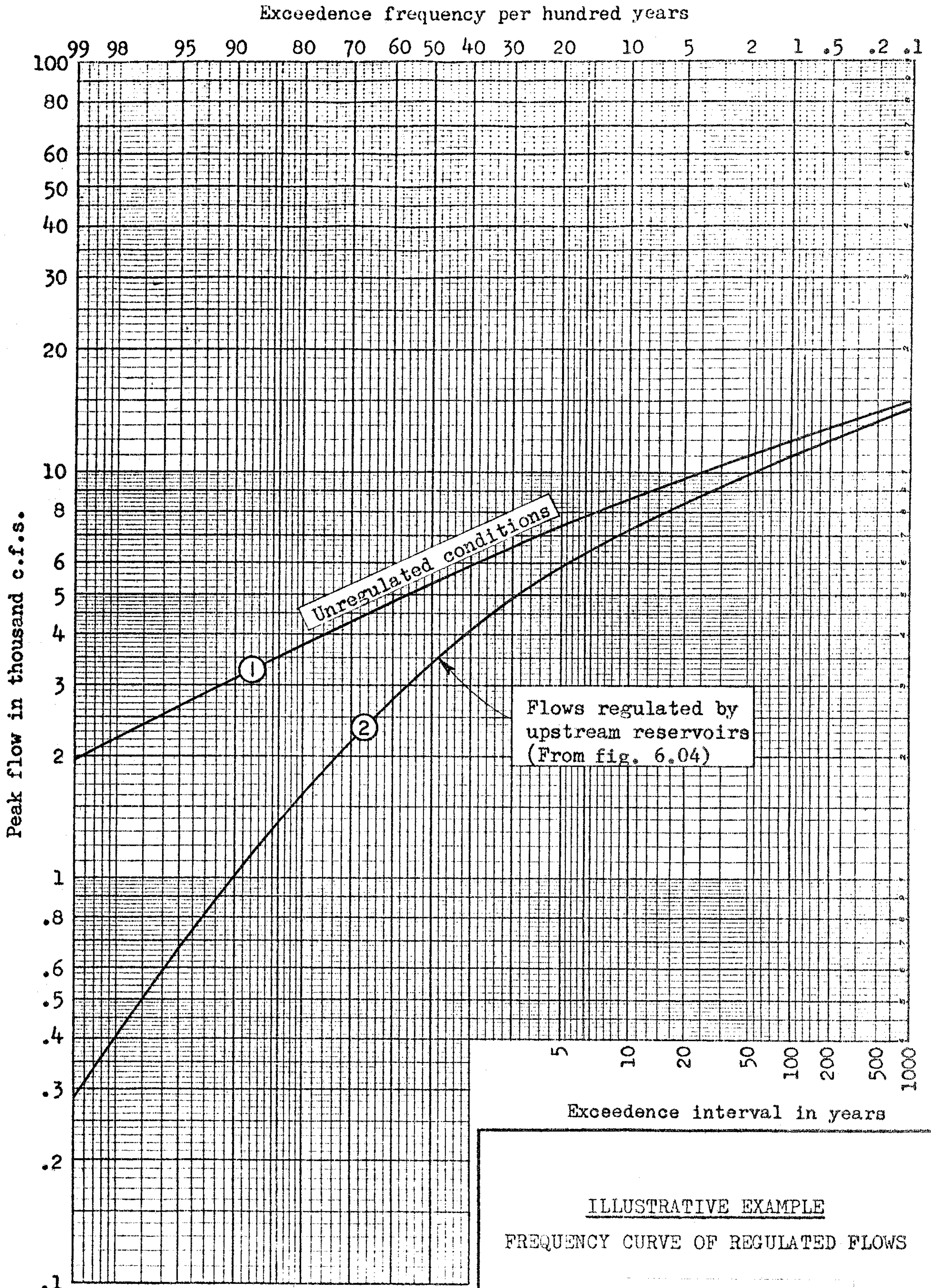
In constructing frequency curves of regulated flows, it must be recognized that reservoir operation for purposes other than flood control will frequently provide incidental regulation of floods. However, usually the availability of such space cannot be depended upon, and its value is considerably diminished for this reason. Consequently, the effects of such space on the reduction of floods should be estimated from careful study.

In constructing frequency curves of regulated flows it should be recognized that actual operation is rarely perfect and that releases



ILLUSTRATIVE EXAMPLE

RELATION OF REGULATED TO UNREGULATED FLOWS



ILLUSTRATIVE EXAMPLE

FREQUENCY CURVE OF REGULATED FLOWS

Figure 6.05

will frequently be curtailed or diminished because of unforeseen operation contingencies. Also, where flood forecasts are involved in the reservoir operation, it must be recognized that these are subject to considerable error and that some allowance for error will be made during operation. In accounting for these factors, it will be found that the actual control of floods is somewhat less than could be expected if full release capacities and downstream channel capacities were utilized efficiently and if all forecasts were exact.

In estimating the frequency of runoff at a damage center considerably downstream from a reservoir project, it must be recognized that none of the runoff from the intermediate areas between the reservoir and the damage center will be regulated. To account for this factor, a frequency curve of runoff for the intermediate area can be used as a lower limit for the curve of regulated flows.

Section 6.05. Effects of Channel, Levee and Floodway Modifications

The effect of channel, levee and floodway modifications on river stages at the project location and on river discharges downstream from the project location can generally be evaluated by routing several typical floods through the reaches of the modification and the affected upstream reaches. The stages or discharges thus derived can be plotted against corresponding without-project values, and a smooth curve drawn. This curve could be used in conjunction with a frequency curve of without-project values to construct a frequency curve of with-project values as discussed in Section 6.04. Corresponding stages upstream from the selected control point can be estimated from water surface profile computations.

It is recognized that levee failures may and do occur occasionally when river stages are considerably below project design level. On the other hand, there are many cases of levees remaining intact when stages exceed project design level or even levee crest grade. In the absence of comprehensive statistics on the failure of various types of levees at various stages, it may be satisfactory to assume for the purpose of estimating flood frequencies, that a levee will always fail at stages above the design water level and that failure will not occur at lower stages. Detailed studies for levees constructed under adverse conditions or areas requiring very long levees, may suggest more conservative assumptions to be appropriate.

Section 6.06. Changes in Stage-Discharge Relationships

Changes in stage-discharge relations due to channel modifications, levee construction or flow obstructions can best be evaluated by computing water surface profiles for each of a number of discharges. Procedures for computing water surface profiles are contained in Volume 6 of this report. The resulting relationships for modified conditions can be used in the storage routing procedures for evaluating the effects of these changes on flows downstream (due to changes in channel storage).

Section 6.07. Effects of Multiple-Reservoir Systems

When more than one reservoir exists above a damage center, the problem of evaluating reservoir stages and downstream flows under with-project conditions becomes increasingly complex. The most reliable method of estimating the effects is to make complete routings of historical runoff throughout the system and to supplement these routings as necessary with routings of large hypothetical floods. The resulting data can be used as discussed in Sections 6.03 and 6.04 for constructing frequency and duration curves. As the number of reservoirs increases, the amount of work involved can become prohibitive, and some sort of representative hydrograph procedure may be needed.

Section 6.08. Representative Hydrograph Procedures

Because of the large amount of computation that would be involved in evaluating project effects on all sizes, and time and area patterns of floods, it is usually necessary to limit analysis to a relatively few typical floods. This procedure must be used with caution, however, because certain characteristics of atypical floods may be responsible for critical flooding conditions. Accordingly, such studies should be supplemented by a careful examination of the potential effects of atypical floods.

Typical floods for use in a study of regulated flood frequencies can be obtained by adjusting the ordinates of observed flood hydrographs to conform to runoff volumes for various durations and locations that have a uniform exceedence frequency. They could also be computed from rainfall having a uniform exceedence frequency for all durations and locations and having a typical time pattern. In the latter case, the runoff frequency does not necessarily correspond to the rainfall frequency, and must ordinarily be evaluated from observing the runoff frequency at locations where frequency curves have been derived from observed runoff data.

As long as the typical flood hydrographs used for this purpose are balanced with respect to the severity of runoff in various tributaries and for various durations, those floods that are larger under without-project conditions will be larger under with-project conditions for the same initial conditions and the same project operation. Accordingly, the exceedence frequency of the flood under project conditions will be the same as that under without-project conditions.

The general procedure for use of representative hydrographs in deriving rainflood frequencies for modified conditions using balanced floods described in Volume 4 is as follows:

a. Develop a hypothetical storm having uniform severity of rainfall throughout the stream system and uniformly severe for all pertinent durations of rainfall.

b. Develop generalized unit hydrograph and loss functions using procedures discussed in Volume 4.

c. Subdivide the drainage basin into areas as necessary and establish unit hydrograph and loss functions for each area and condition of development.

d. Establish routing coefficients for each river reach and for each condition of development and routing data for each reservoir.

e. Develop runoff frequency curves at locations where data are available for a specified condition of development, such as without-project conditions.

f. For each of several multiples of the rainstorm, compute flood runoff for each condition. A frequency curve for modified conditions at any location can then be constructed by plotting the modified flow of each storm for that condition against the known exceedence frequency of corresponding flows for the without-project condition. A check should be made to assure that regulated frequencies for any location are at least as high as the frequency of runoff from the unregulated portion of the tributary area.

Chapter 7

Regional Frequency Analysis

CHAPTER 7. REGIONAL FREQUENCY ANALYSIS

Section 7.01. General Principles

As discussed in previous chapters of this volume, regionalized frequency studies are often desirable. Procedures recommended herein consist of correlating the mean and standard deviation of annual maximum flow values with pertinent drainage basin characteristics by use of multiple linear correlation procedures. The same principles can be followed using graphical frequency and correlation techniques where these are more appropriate. Where regional studies of the skew coefficient are necessary (where the applicability of values given in Section 5.03 is suspect), procedures suggested in Section 7.08 can be used.

Section 7.02. Frequency Statistics

A regional frequency correlation study is based on the two principal frequency statistics - the mean and standard deviation of annual maximum flow logarithms. Prior to relating these frequency statistics to drainage basin characteristics, it is essential that the best possible estimate of each frequency statistic be made. This is done by adjusting short-record values by the use of longer records at nearby locations. When many stations are involved, it is best first to select long-record base stations for each portion of the region. It might be desirable to adjust the base station statistics by use of the one or two longest-record stations in the region, and then adjust the short-record station values by use of the nearest or most appropriate base station. Methods of adjusting statistics are discussed in Section 4.04.

Section 7.03. Drainage-basin Characteristics

A regional analysis involves the determination of the main factors responsible for differences in precipitation or runoff regimes between different locations. This would be done by correlating important factors with the long-record mean and with the long-record standard deviation of the frequency curve for each station (the long-record values are those based on extension of the records as discussed in Section 4.04). Statistics based on rainfall measurements might be correlated in mountainous terrain with the following factors:

- a. Elevation of station
- b. General slope of surrounding terrain
- c. Orientation of that slope
- d. Elevation of windward barrier
- e. Exposure of gage
- f. Distance to leeward controlling ridge

Statistics based on runoff measurements might be correlated with the following factors:

- a. Drainage area (contributing)
- b. Slope of drainage area or of main channel
- c. Surface storage (lakes and swamps)
- d. Mean annual rainfall
- e. Number of rainy days per year
- f. Infiltration characteristics
- g. Stream length

Correlation methods and their application are discussed in Chapter 4 of Volume 2.

Section 7.04. Selection of Appropriate Relationships

In order to obtain satisfactory results using multiple linear correlation techniques, all variables must be expressed so that the relation between the independent and any dependent variable can be expected to be linear, and so that the interaction between two independent variables is reasonable. An illustration of the first condition is the relation between rainfall and runoff. If the runoff coefficient is sensibly constant, as in the case of parking lot or airport drainage, then runoff can be expected to bear a linear relation to rainfall. However, in many cases initial losses and infiltration losses cause a marked curvature in the relationship. Ordinarily, it will be found that the logarithm of runoff is very nearly a linear function of rainfall, regardless of loss rates, and in such cases, linear correlation of logarithms would be most suitable. An illustration of the second condition is the relation between rainfall, D, drainage area, A, and runoff, Q. If the relation used for correlation is as follows:

$$Q = aD + bA + c \qquad (7\text{-}1)$$

then it can be seen that one inch change in precipitation would add the same amount of flow, regardless of the size of drainage area. This is not reasonable, but a transformation to logarithms would yield a reasonable relation:

$$\text{Log } Q = d \log D + e \log A + \log f \qquad (7\text{-}2)$$

or transformed:

$$Q = fD^d A^e \qquad (7\text{-}3)$$

Thus, if logarithms of certain variables are used, doubling one independent quantity will multiply the dependent variable by a fixed ratio, regardless of what fixed values the other independent variables have. This particular relationship is reasonable and can be easily visualized after a little study. There is no simple rule for deciding when to use logarithmic transformation. It is only feasible, however, when the variable has a fixed lower limit of zero.

Section 7.05. Example of Regional Correlation

An illustrative example of a regional correlation analysis of standard deviation of annual peak flows with drainage area and number of rainy days per year is given on fig. 7.01. Since many important variables are neglected, the analysis is not of the scope necessary for a complete study, but is useful for illustrating various techniques and problems involved in such a study. In the example, X_1 is one plus the logarithm of the adjusted standard deviation (one is added to eliminate negative values), X_2 is the logarithm of the drainage area size, and X_3 is the logarithm of the average number of rainy days per year for the drainage area. The regression equation is derived as shown, and the calculated coefficient of determination is 0.31, which means 31 percent of the variance of X_1 is explained by the regression equation.

Section 7.06. Selection of Useful Variables

In the regression equation derived on fig. 7.01, the coefficient of X_2 is very small, which appears to indicate that this factor has very little effect. To determine the usefulness of this factor, it is, however, necessary to make an additional analysis using all variables except this one. In this case, the problem would resolve into a simple correlation analysis using only the variables X_3 and X_1. Solution of the simple correlation equations produces the relationship:

ILLUSTRATIVE EXAMPLE

REGIONAL FREQUENCY CORRELATION

$X_1 = 1 + \log. S$ $\quad X_2 = \log. D.A.$ $\quad X_3 = \log.$ No. rainy days per year

Sta. No. (1)	X_2 (2)	X_3 (3)	X_1 (4)	Sta. No. (5)	X_2 (6)	X_3 (7)	X_1 (8)
1	1.61	2.11	0.29	33	1.94	1.87	0.20
2	2.89	2.12	0.18	34	2.73	1.36	0.58
3	4.38	2.11	0.17	35	3.63	1.81	0.64
4	3.20	2.04	0.44	36	1.91	1.58	0.37
5	3.92	2.07	0.38	37	2.26	1.48	0.27
6	1.61	2.04	0.37	38	2.97	1.89	0.54
7	3.21	2.09	0.30	39	0.70	1.32	0.63
8	3.65	1.99	0.35	40	0.30	1.54	0.78
9	3.23	2.15	0.16	41	3.38	1.62	0.46
10	4.33	2.08	0.11	42	2.87	2.03	0.44
11	1.60	2.09	0.32	43	2.42	2.26	0.24
12	2.82	2.00	0.34	44	4.53	1.93	-0.03
13	2.40	2.00	0.25	45	3.04	1.78	0.30
14	3.69	2.09	0.43	46	4.13	2.00	0.17
15	2.18	2.19	0.27	47	1.49	2.01	0.14
16	2.09	2.17	0.25	48	5.37	1.95	0.10
17	4.48	1.91	0.52	49	1.36	2.11	0.27
18	4.95	1.95	0.18	50	2.31	2.23	0.18
19	2.21	1.97	0.39				
20	3.41	2.08	0.40	ΣX	147.55	96.24	17.89
21	4.82	1.88	0.25	$\bar{X}$	2.951	1.925	0.358
22	1.78	1.93	0.23	ΣXX_2	503.7779	285.5627	51.1527
23	4.39	1.74	0.54	$\Sigma X \Sigma X_2/N$	435.4200	284.0042	52.7934
24	3.23	2.01	0.51	Σxx_2	68.3579	1.5585	-1.6407
25	3.58	2.04	0.45				
26	1.64	1.78	0.63	ΣXX_3		187.5912	33.2598
27	4.58	1.76	0.45	$\Sigma X \Sigma X_3/N$		185.2428	34.4347
28	3.26	1.93	0.59	Σxx_3	1.5585	2.3484	-1.1749
29	4.29	1.81	0.46				
30	1.23	1.89	0.32	ΣXX_1			8.1635
31	3.44	1.48	0.96	$\Sigma X \Sigma X_1/N$			6.4010
32	2.11	1.97	0.12	Σxx_1	-1.6407	-1.1749	1.7625

$$\left.\begin{array}{ll} 68.4\, b_2 + 1.56\, b_3 = -1.64 & \text{(Eq. 4-7)} \\ 1.56\, b_2 + 2.35\, b_3 = -1.17 & \text{(Eq. 4-9)} \end{array}\right\} \quad \begin{array}{l} b_2 = -.013 \\ b_3 = -.49 \end{array}$$

NOTE: See Chapter 4, Volume 2, for equations used in this example.

$$a = 0.36 + 0.13\,(2.95) + .49\,(1.92) = 1.34 \quad \text{(Eq. 4-12)}$$

$$X_1 = 1.34 - .013X_2 - .49X_3$$

$$R^2 = \frac{-.013\,(-1.64) - .49\,(-1.17)}{1.76} = .338 \quad \text{(Eq. 4-15)}$$

$$\bar{R}^2 = 1 - (.662)\ 49/47 = .310 \text{ (Eq. 4-16), and } \bar{R} = .56$$

Figure 7.01

$$X_1 = 1.32 - 0.50X_3 \tag{7-4}$$

A solution for $\bar{R}^2$ would yield 0.32. Thus, a better correlation is obtained neglecting drainage area as a factor. If additional factors were considered in the analysis, the effect of drainage area should be reconsidered, as it is possible that its effect is obscured in the example by neglecting some other important variable. A test of importance of a particular factor is a comparison of the correlation coefficient using all factors and that omitting only the factor whose influence is being tested. Even in the case of a slight increase in correlation obtained by adding a variable, consideration of the increased unreliability of $\bar{R}$ might indicate that such a factor should be eliminated in cases of small samples.

Judgment on the worth of including an additional factor should also be based on a comparison of the standard error of estimates. The units of this parameter are that of the dependent variable and, therefore, provides a better indication of the magnitudes of the unexplained errors.

Section 7.07. Use of Map

Many hydrologic factors cannot be expressed numerically. Examples are soil characteristics, vegetal cover, and geologic characteristics. For this reason, numerical regional analysis will explain only a portion of the regional variation of runoff frequencies. The remaining unexplained variance is contained in the regression constant, which can be considered to vary from station to station. These regression constants can be computed by inserting the drainage basin characteristics and frequency statistic for each station in the regression equation and solving for the regression constant. These constants can be plotted on a regional map, and lines of equal values drawn (perhaps using soils or vegetation maps as a guide). Use of such a map for selecting a regression constant

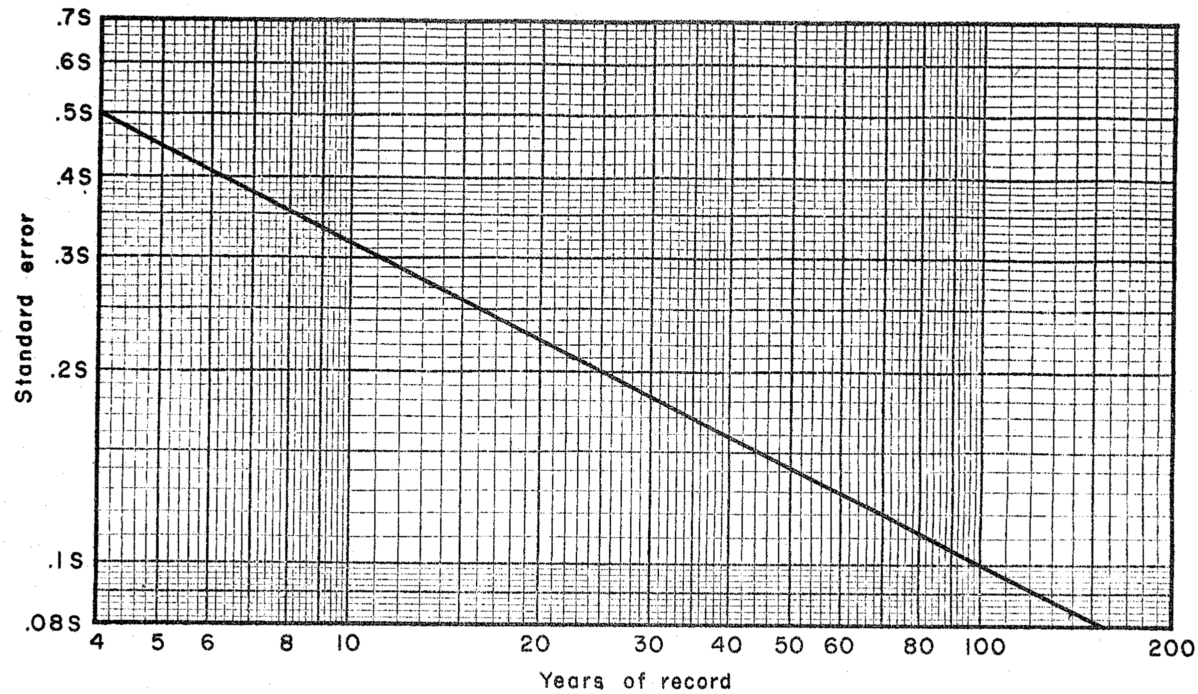


Fig. a. Standard error of a calculated mean.

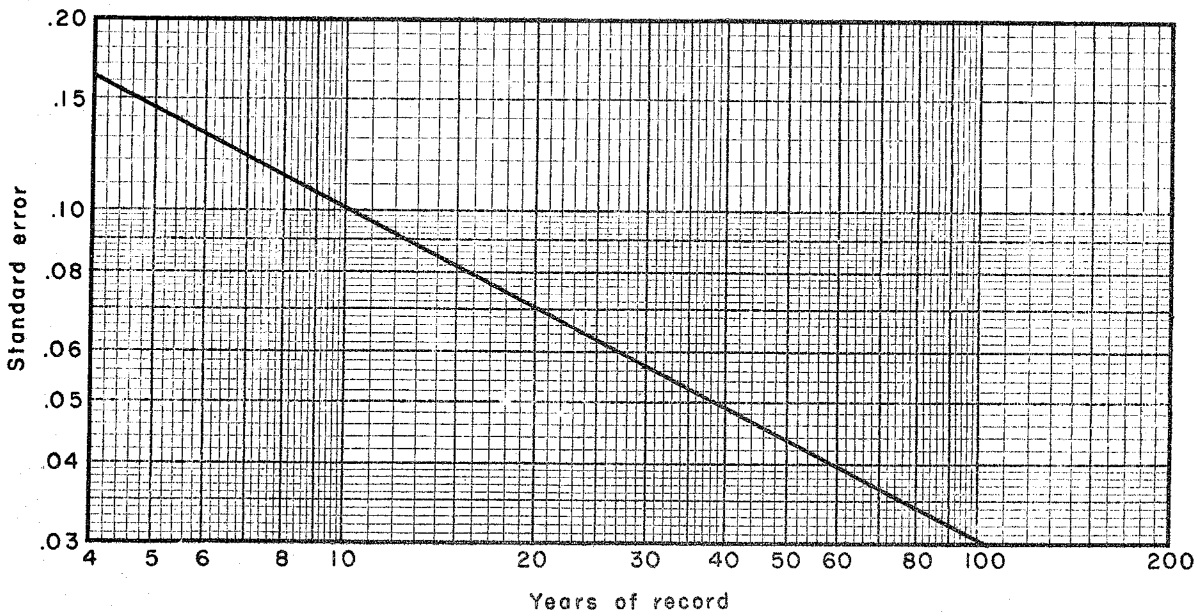


Fig. b. Approximate standard error of the logarithm of a calculated standard deviation

STANDARD ERRORS OF FREQUENCY STATISTICS

should be much better than using a single constant for the entire region. In smoothing lines on such a map, consideration should be given to the reliability of computed statistics. Figure 7.02 shows standard errors of estimating means and standard deviations. As an example, if a computed standard deviation based on 30 years of record is 0.300, there is about one chance in 20 that the mean is in error by more than 0.114 (twice the standard error of 0.057 as indicated by curve (b) for 30 years of record) or that the standard deviation is in error by a factor of 1.3 (antilog of 0.114).

Section 7.08. Summary of Procedure

A regional analysis of precipitation or flood flows is accomplished in the following steps:

a. Select long-record base stations within the region as required for extension of records at each of the short-record stations.

b. Tabulate maximum events of each station, corresponding logarithms, and logarithms of base-station values for the corresponding years. Logarithms should be rounded to 2 decimal places.

c. Calculate $\overline{X}$ and S (equations 4-2 and 4-3) for each base station.

d. Calculate $\overline{X}$ and S for each other station and for the concurrent period at the base station, and calculate the correlation coefficient (equations 4-15 and 4-16 from Volume 2).

e. Adjust all values of $\overline{X}$ and S by use of the base station, (equations 4-9 and 4-10). (If any base station is first adjusted by use of a longer-record base station, the longer-record statistics should be used for all subsequent adjustments.)

f. Select meteorological and drainage basin parameters that are expected to correlate linearly with $\overline{X}$ and log S, and tabulate estimated values of these for each area. (The physical significance of log S is not important, as the transformation simply eliminates a lower limit of zero from the regression variables.)

g. Calculate the regression equations relating $\overline{X}$ and log S in turn to these statistics, using procedures explained in Chapter 4 of Volume 2, and compute the corresponding determination coefficients.

h. Eliminate variables in turn that contribute the least to the determination coefficient, recomputing the determination coefficient each time, and select a regression equation with a minimum number of variables that explains a large portion of the variation in $\overline{X}$ or S (consider determination coefficient and standard error of estimate).

i. Compute the regression constants for each station, plot on a suitable map, usually at the centroid of the areas, and draw isopleths of the regression constant for the regression equations of $\overline{X}$ and S (two maps), considering that the regression constant for a station represents a basin-mean value.

j. A frequency curve can be computed from constants obtained for any basin on the map, using the computed regression equations to obtain $\overline{X}$ and S, and using procedures discussed in Section 4.02 for computing a frequency curve therefrom.

Section 7.09. Regional Skew Determinations

Skew coefficients for use in hydrologic studies should be based on regional studies, since values based on individual records in the order of 100 years or less are highly unreliable. For small areas this can be done by computing skew coefficients for available records and using the average, weighted in accordance with record length. For large areas the computed values, along with the standard error, can be plotted at the centroid of the respective basins and isopleths of equal skew drawn. The standard error of the skew coefficient should be used as criterion for departing from the computed skew value.

Section 7.10. Sample Regional Criteria

The regional frequency analysis described in reference 20 is considered to be a moderately elaborate type of analysis. The criteria derived are suitable for selecting frequency statistics for deriving frequency curves of runoff peaks and volumes for durations up to 30 days. Standard deviations for frequency curves for the various durations of runoff are obtained directly from maps constructed for each duration. The mean logarithms of runoff for each duration, however, have been related to drainage area size, normal annual precipitation, elevation, and, by means of maps, to geographical location. To illustrate the relative simplicity of this scheme, criteria for determining frequency curves of peak flows are included herein as figs. 7.03 to 7.06. Data and computations required for synthetic frequency curves for the drainage basin used in illustrating graphical and analytical methods on figs. 4.02 and 4.03 (Mill Creek at latitude 40^{o} 03') are as follows:

Drainage basin characteristics (location shown on fig. 7.03):

Drainage area	134 sq. mi.
Normal annual precipitation	47 in.
Average elevation	2900 ft. m.s.l.
Average latitude	40^{o} - 03'

Frequency constants:

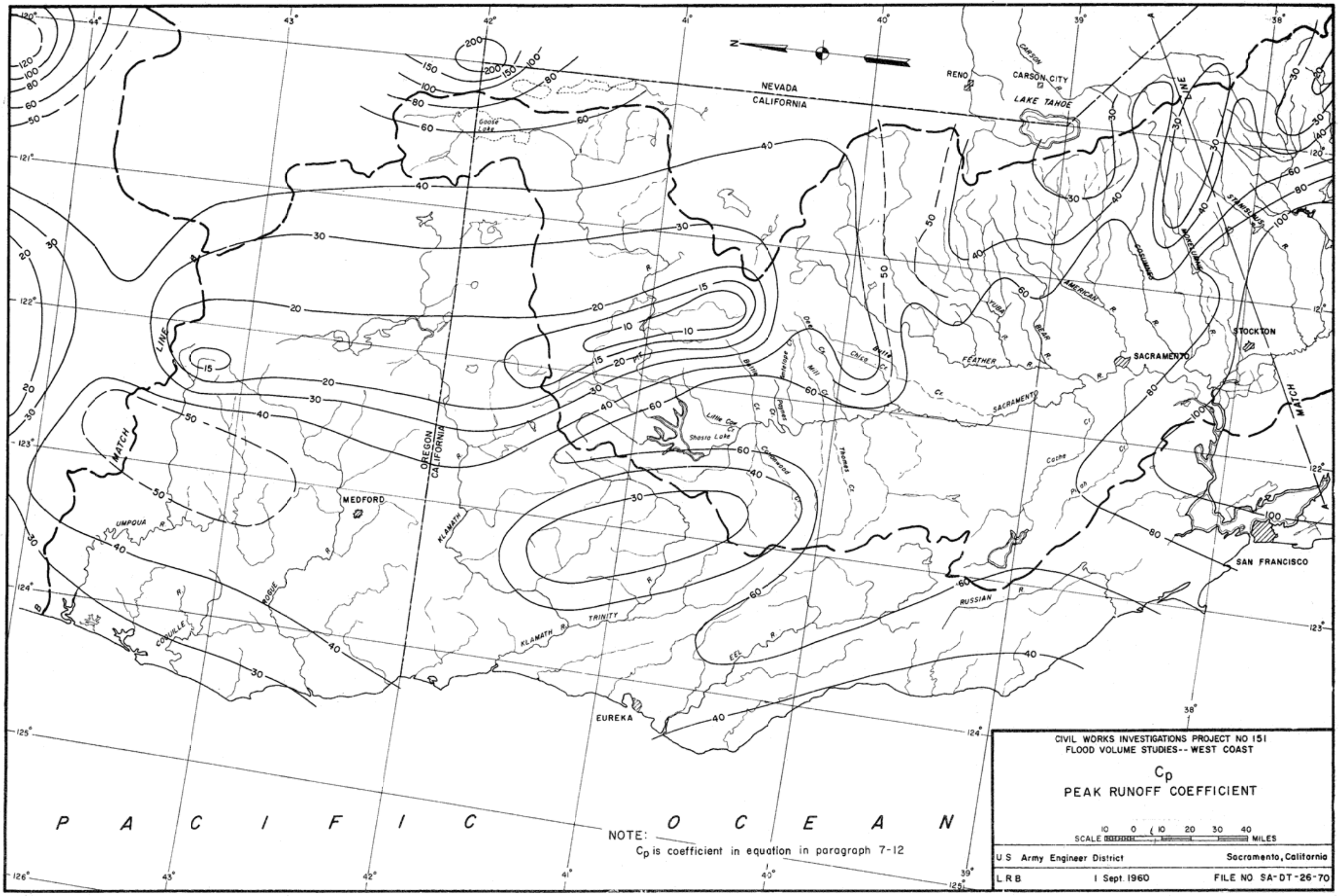
Standard deviation (from fig. 7.04)	.31
C_p (from fig. 7.03)	42
K (from fig. 7.05)	.71

Computation of Q_p

$$Q_p = .001\ C_p A^{.85}\ P^2\ K \text{ (equation 11, reference 20)}$$

$$= .001\ (42)\ (64)\ (47)^2\ (.71)$$

$$= 4220 \text{ c.f.s.}$$



NOTE:
C_p is coefficient in equation in paragraph 7-12

CIVIL WORKS INVESTIGATIONS PROJECT NO 151
FLOOD VOLUME STUDIES-- WEST COAST

C_p
PEAK RUNOFF COEFFICIENT

SCALE 10 0 10 20 30 40 MILES

U S Army Engineer District Sacramento, California
L R B 1 Sept. 1960 FILE NO SA-DT-26-70

7-11

Figure 7.03

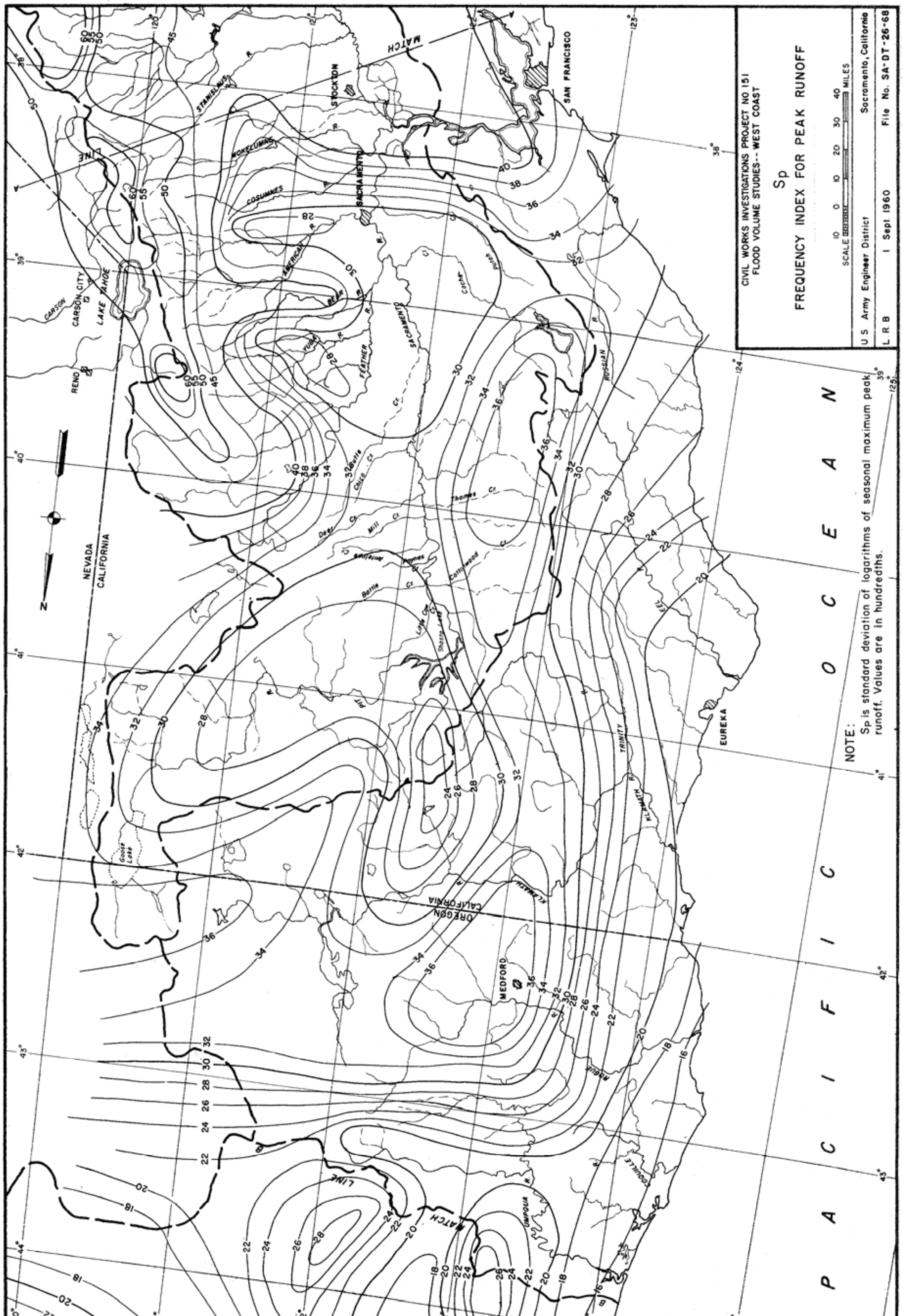
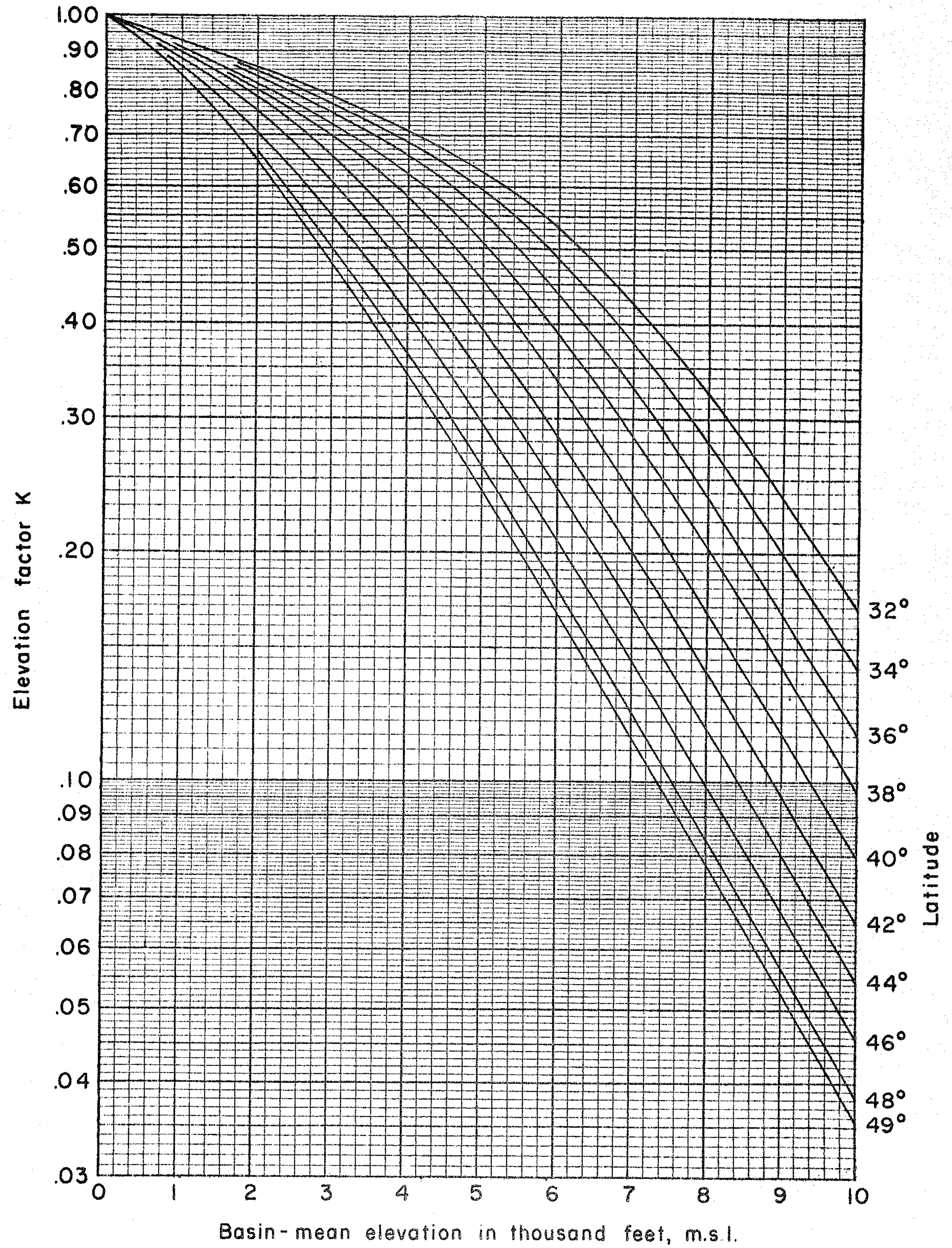


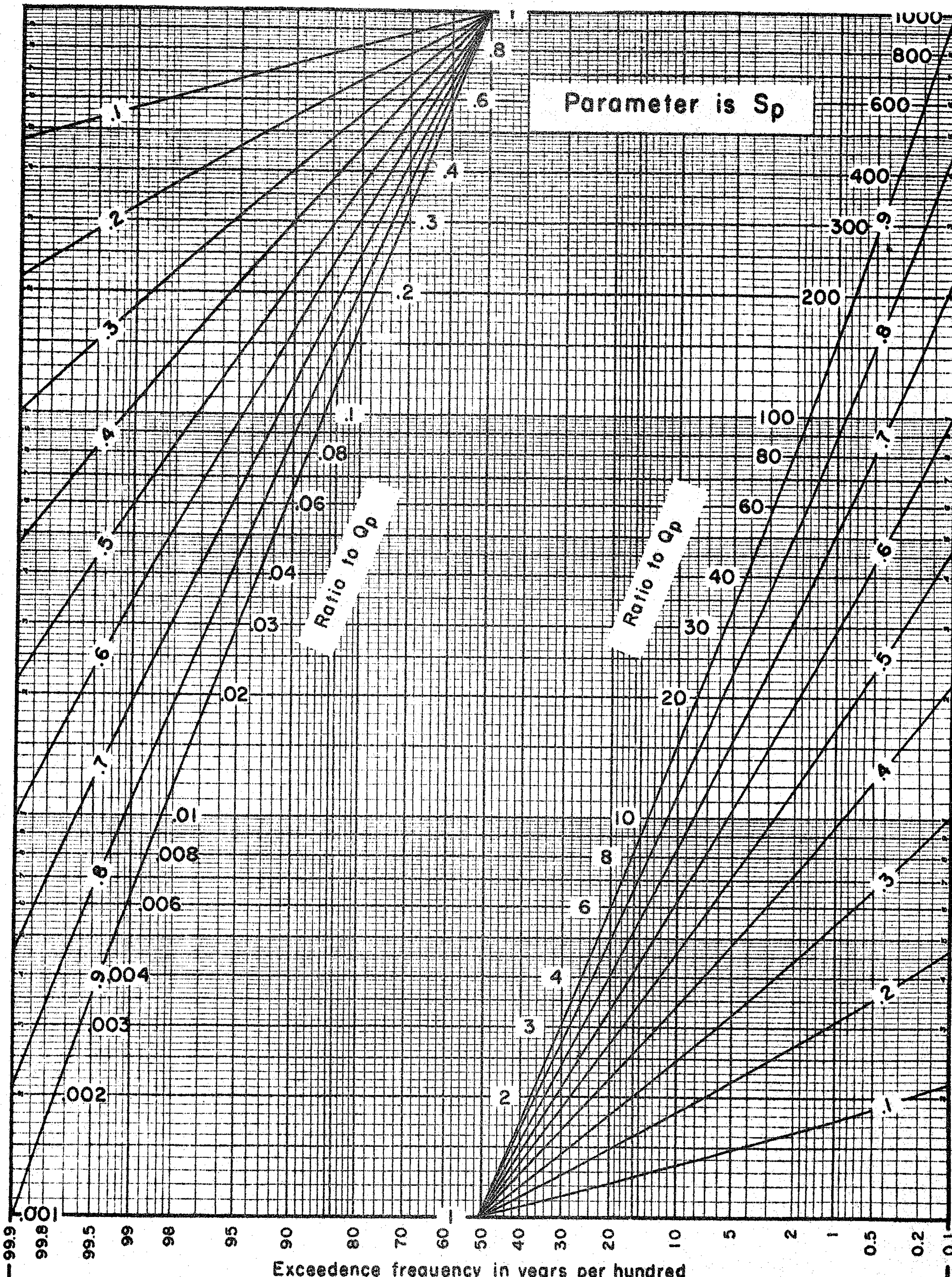
Figure 7.04

7-12



BASIN ELEVATION FACTOR

Figure 7.05



INDEX FREQUENCY CURVES FOR PEAK RUNOFF

Figure 7.06

Flow of specified frequency (say once per 100 years)

$Q_{.01}$ = 5.7 (4,220) = 24,000 c.f.s. (value of 5.7 read from fig. 7.06 at S = .31 and frequency of 1 per 100 years)

Chapter 8

Frequencies of Other Hydrologic Factors

CHAPTER 8. FREQUENCIES OF OTHER HYDROLOGIC FACTORS

Section 8.01. Introduction

The frequency methods described in Chapter 4 have application to runoff, and can also be used in estimating frequencies of various other hydrologic factors. Some of the more common applications are described in the following sections.

Section 8.02. Rainfall Frequencies

Graphical and analytical procedures for the computation of frequency curves of station precipitation are generally identical to those for streamflow analysis. In precipitation studies, however, instantaneous peak intensities are ordinarily not analyzed, since they are virtually impossible to measure and of little practical value. Precipitation amounts for specified durations are commonly analyzed, mostly for durations of less than 3 or 4 days. The few studies made thus far have indicated that the logarithmic normal function (with zero skew coefficient) is fitted fairly well with annual maximum station precipitation data, regardless of the duration used. Station precipitation alone is not adequate for most hydrologic studies, and some means of evaluating the frequency of precipitation occurring simultaneously or near simultaneously over the area is necessary.

Section 8.03. Low Flow Frequencies

The design of hydroelectric powerplants and the design of reservoirs for supplementing low river flows for water quality and other purposes requires the evaluation of the frequencies of low flows for various durations. The method of frequency analysis previously discussed is usually applicable, except that minimum instead of maximum runoff for each period is selected from the basic data. In studying low flows, it will be found that the effects of basin development are relatively great. For example, a relatively moderate diversion can be neglected when studying floodflows, but might greatly modify or even eliminate low flows. Accordingly, one of the most important aspects of low flows concerns the evaluation of past and future effects of basin developments. Analytical frequency techniques are not always applicable to low-flow analysis because no single frequency distribution has been found that satisfactorily fits recorded data. Graphical procedures are more generally applicable.

In regions of water scarcity and where a high degree of development has been attained, projects that entail carryover of water for several years are often planned. In such projects it is desirable to analyze low-flow volume frequencies for periods ranging from 1-1/2 to 8-1/2 years or more. Because the number of independent low-flow periods of these lengths in even the longest historical records is very small and because the concept of multi-annual periods is somewhat inconsistent with the basic concept of an "annual event" as implied by the language, "number of occurrences per hundred years," there is no truly satisfactory way for computing the frequency of low-flow periods more than 1 year in length. One procedure described in reference 10 has been used with long sequences of generated streamflows to derive estimates of drought frequency. Although the results obtained through the use of the procedure (described in Volume 8) seem reasonable, it is impossible to verify the accuracy of the frequency estimates.

Section 8.04. Hurricane Frequencies

In studies of hurricane wind velocities by the Corps of Engineers and U.S. Weather Bureau, the central pressure index (estimated minimum sea level pressure for an individual hurricane) has been used in conjunction with pressure versus wind relationships to determine wind frequencies. The index frequency for each of three large geographic zones was determined as illustrated on fig. 8.01, and subsequently divided into frequencies for various subdivisions of each zone. The minimum hurricane pressure apparently plots close to a straight line on arithmetic probability paper.

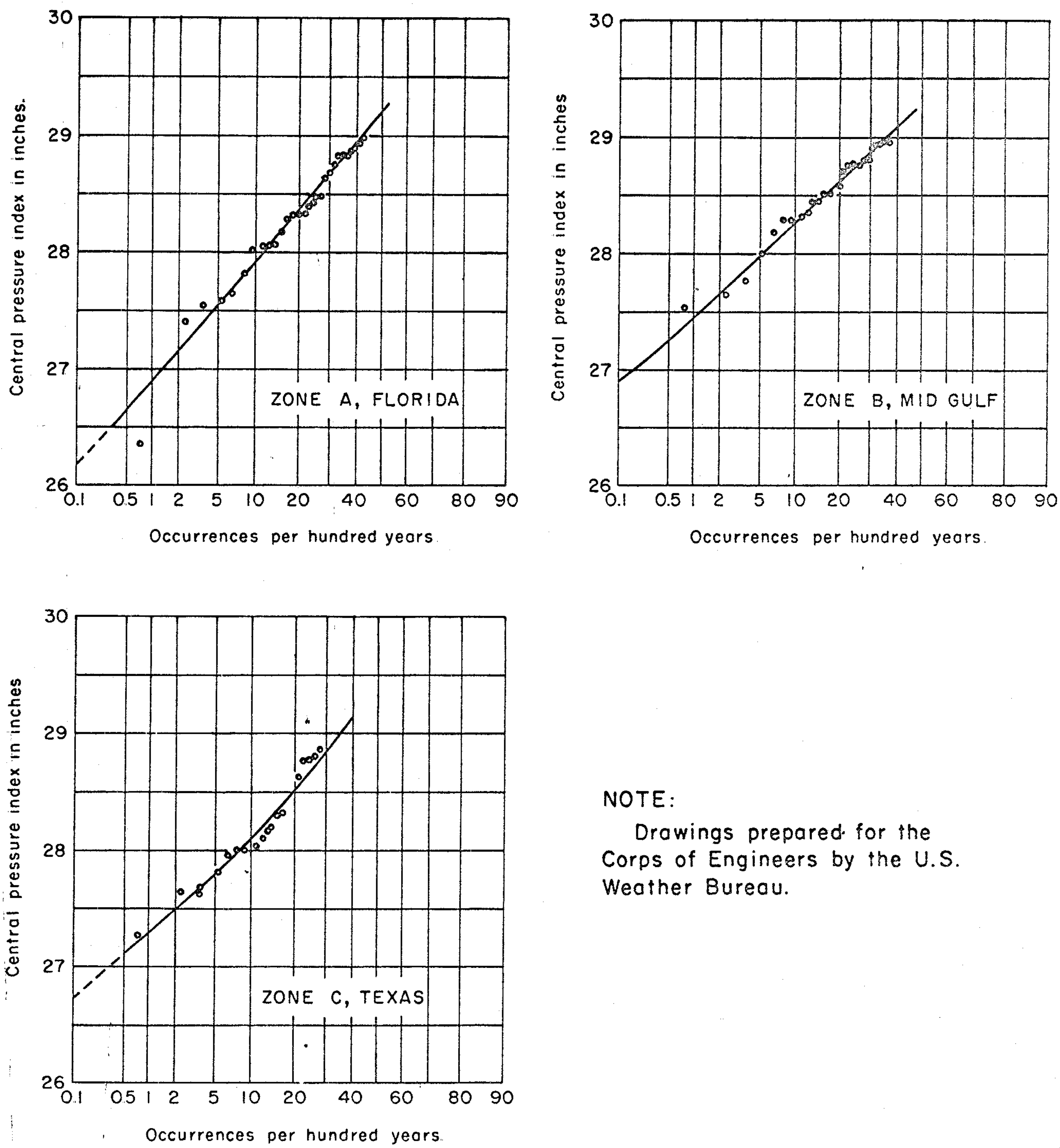
Section 8.05. Sediment Frequencies

Another illustration of use of frequency techniques discussed herein is shown on fig. 8.02, where the frequency of annual sediment load of the Colorado River is shown to approximate a linear relationship on log probability paper.

Section 8.06. Coincident Frequencies

In many cases of hydrologic design, it is necessary to consider only those events which occur coincidentally with other events. For example, a pumping station is usually required to pump water only when interior drainage occurs at a time that the main river stage is above the gravity outlet. In constructing a frequency curve of interior drainage flows that occur only at such times, data selected for direct use should be limited to that recorded during high river stages. In some cases, such data might be adequate, but it is usually possible in cases where the two types of events do not correlate to make indirect use of noncoincident data in order to establish a more reliable frequency curve of coincident events. The general procedure used is as follows:

a. Select the least variable (more stable) of the two variables whose coincidental frequency is to be determined. This will be designated as variable B, and the other as variable A.

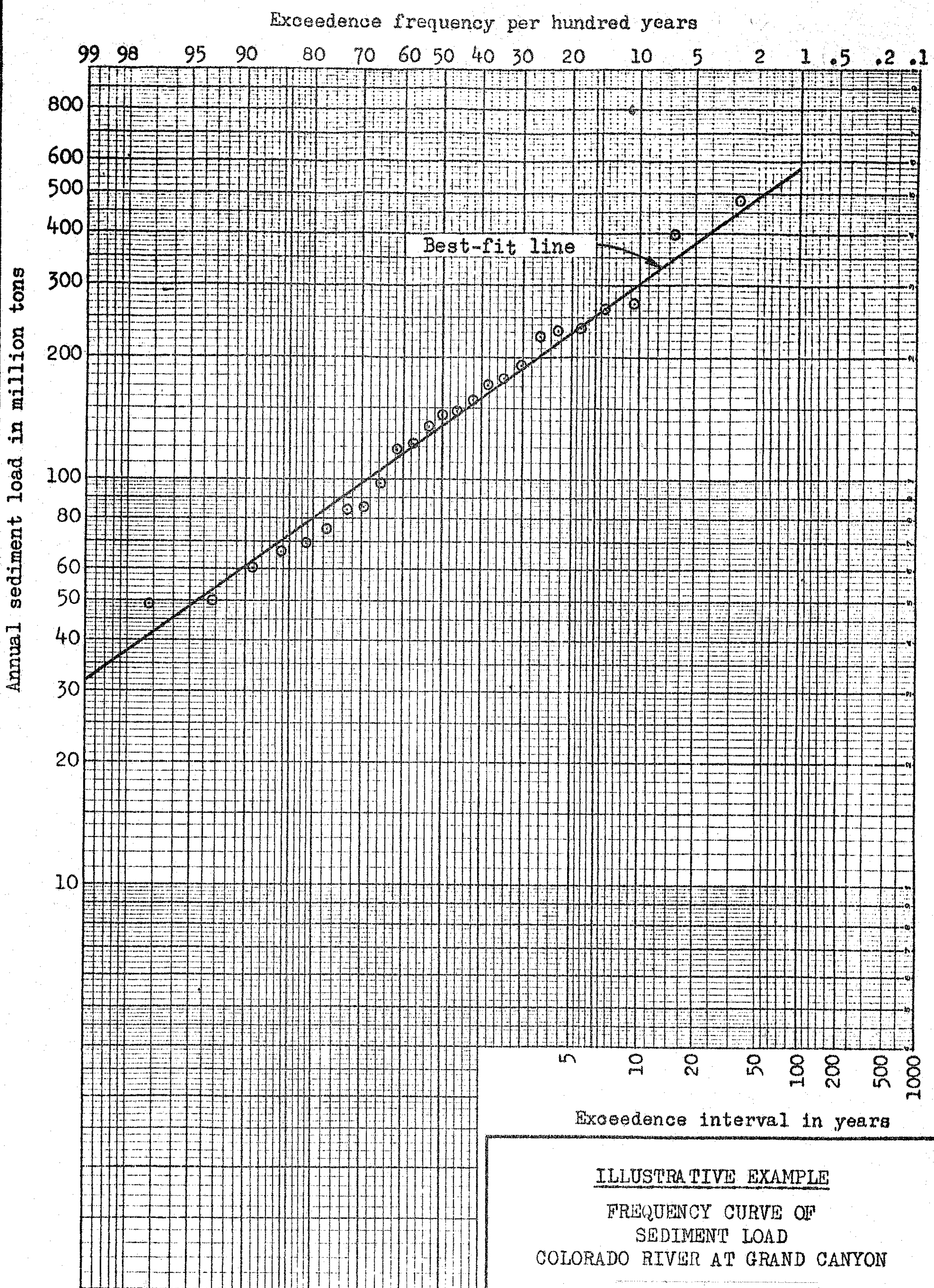


NOTE:

Drawings prepared for the Corps of Engineers by the U.S. Weather Bureau.

ILLUSTRATIVE EXAMPLE

ACCUMULATED FREQUENCY OF HURRICANE CENTRAL PRESSURES (PLOTTED AS FREQUENCY PER 100 YEARS BASED ON 1900-1967)



ILLUSTRATIVE EXAMPLE

FREQUENCY CURVE OF
SEDIMENT LOAD
COLORADO RIVER AT GRAND CANYON

Figure 8.02

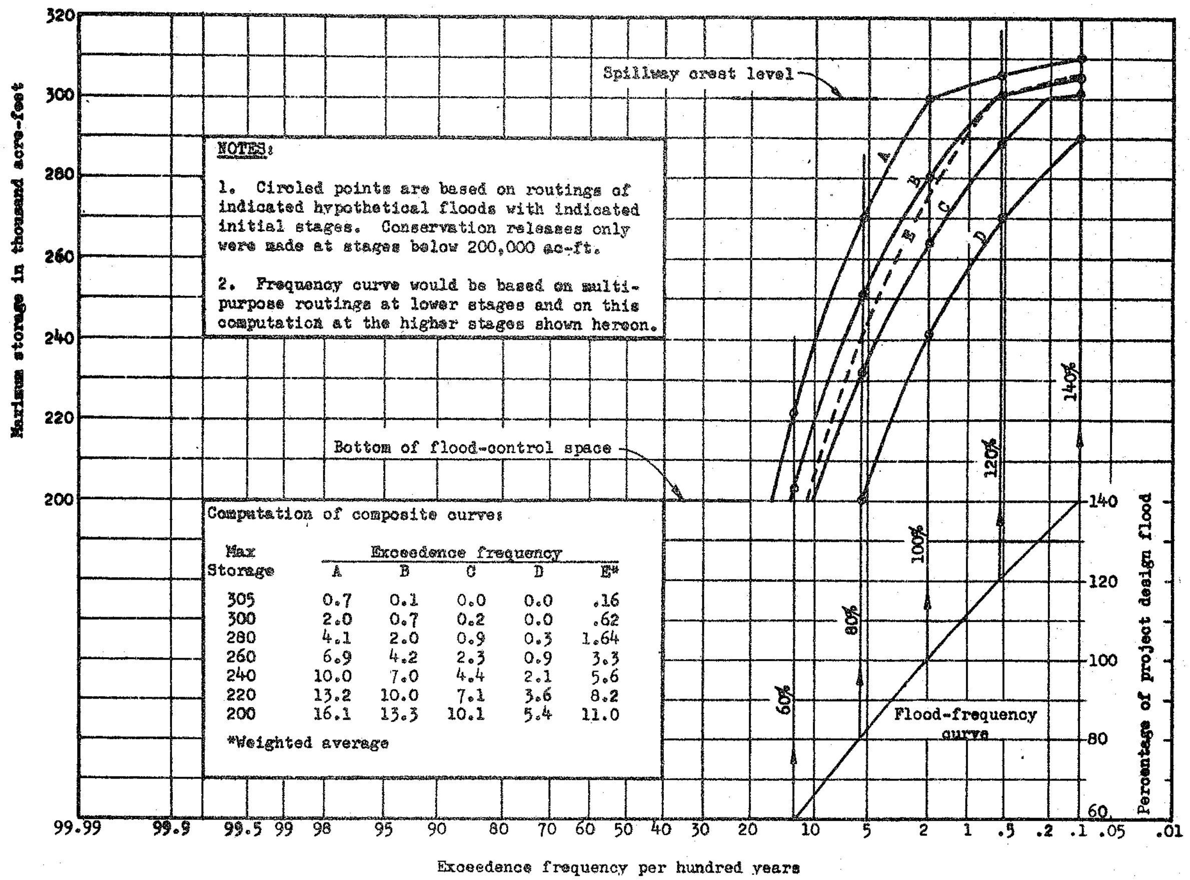
b. Determine the time limits of the seasons during which high stages of variable B are about uniformly likely. If all important stages of variable B can be limited to one season, the analysis will be simplified. Otherwise the following steps must be duplicated for each season.

c. Compute a frequency curve of stages (or flows, rainfall amounts, etc.) for variable A, using data obtained only during the selected season.

d. Construct a duration curve of stages (or flows, etc.) for variable B, using data obtained only during the selected season.

e. The exceedence frequency of any selected magnitude of variable A that is coincidental with any specified range of stage for variable B is equal to the product of the exceedence frequency indicated by the curve derived in c and the proportion of time flows at B are within the selected range of stage, as indicated by the curve derived in d.

Figure 8.03 illustrates a computation of reservoir stage or storage frequency curve from consideration of coincidental frequencies. In this hypothetical case of a 300,000 acre-foot reservoir, the top 100,000 acre-feet are reserved for flood control, the next 50,000 for seasonal irrigation requirements, and the remainder for carryover irrigation storage and power head. Because flood control space would rarely be used, routings of recorded floods do not adequately define the frequency of storage in the flood control space. The example shows a flood frequency curve in terms of the project design (50-year) flood, and a storage duration curve determined from monthly routings of recorded runoff. This latter curve represents only those months when major storms are likely to occur, and does not reflect reservoir rises that would occur during floods. Therefore, it represents conditions that can exist at the beginning of a flood. The duration curve was divided into four ranges, and average storages determined for each range. Routings (not shown)



NOTES:

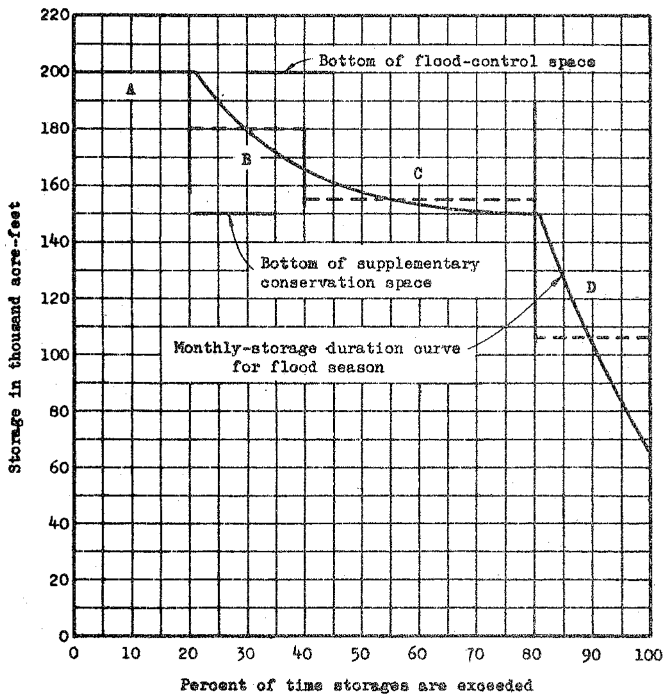
1. Circled points are based on routings of indicated hypothetical floods with indicated initial stages. Conservation releases only were made at stages below 200,000 ac-ft.

2. Frequency curve would be based on multi-purpose routings at lower stages and on this computation at the higher stages shown hereon.

Computation of composite curves:

Max Storage	Exceedence frequency A	B	C	D	E*
305	0.7	0.1	0.0	0.0	.16
300	2.0	0.7	0.2	0.0	.62
280	4.1	2.0	0.9	0.3	1.64
260	6.9	4.2	2.3	0.9	3.3
240	10.0	7.0	4.4	2.1	5.6
220	13.2	10.0	7.1	3.6	8.2
200	16.1	13.3	10.1	5.4	11.0

*Weighted average



A - Initial storage, 200,000 ac-ft. (20% of time)
B - Initial storage, 180,000 ac-ft. (20% of time)
C - Initial storage, 155,000 ac-ft. (40% of time)
D - Initial storage, 106,000 ac-ft. (20% of time)

ILLUSTRATIVE EXAMPLE

STORAGE FREQUENCY COMPUTATION BASED ON COINCIDENT FREQUENCIES of various percentages of the project design flood with these four initial storages were made, and four stage-frequency curves drawn as shown. A composite frequency curve was then drawn as illustrated in the inset table.

Chapter 9

Statistical Reliability Criteria

CHAPTER 9. STATISTICAL RELIABILITY CRITERIA

Section 9.01. Function

One of the principal advantages of the use of statistical procedures is that they provide means for evaluating the reliability of the estimates. This permits a broader understanding of the subject and provides criteria for decision-making. The common statistical index of reliability is the standard error of estimate, which is defined as the root-mean-square error. In general, it is considered that the standard error is exceeded on the positive side one time out of six estimates, and equally frequently on the negative side, for a total of one time in three estimates. An error twice as large as the standard error of estimate is considered to be exceeded one time in 40 in either direction, for a total of one time in 20. These statements are based on an assumed normal distribution of error so they are only approximate frequencies of errors, and exact statements as to error probability must be based on examination of the frequency curve of errors or the distribution of the errors. Both the standard error of estimate and the distribution of errors are discussed in this section.

Section 9.02. Reliability of Frequency Statistics

The standard errors of estimate of the mean, standard deviation, and skew coefficient, which are the principal statistics used in frequency analysis, are given in the following equations:

$$S_{\bar{X}} = \sqrt{S^2/N} \qquad (9\text{-}1)$$

$$S_S = \sqrt{S^2/2N} \qquad (9\text{-}2)$$

$$S_g = \sqrt{6N(N-1)/(N-2)(N+1)(N+3)} \qquad (9\text{-}3)$$

where:

$S_{\bar{x}}$ = the standard of estimate for the mean

S_s = the standard error of estimate for the standard deviation

S_g = the standard error of estimate for the skew coefficient

and S and N are as defined in Section 4.02.

These have been used to considerable advantages as discussed in Chapter 7 in drawing maps of mean standard deviation, and skew coefficient for a regional frequency study.

Section 9.03. Reliability of Frequency Estimates

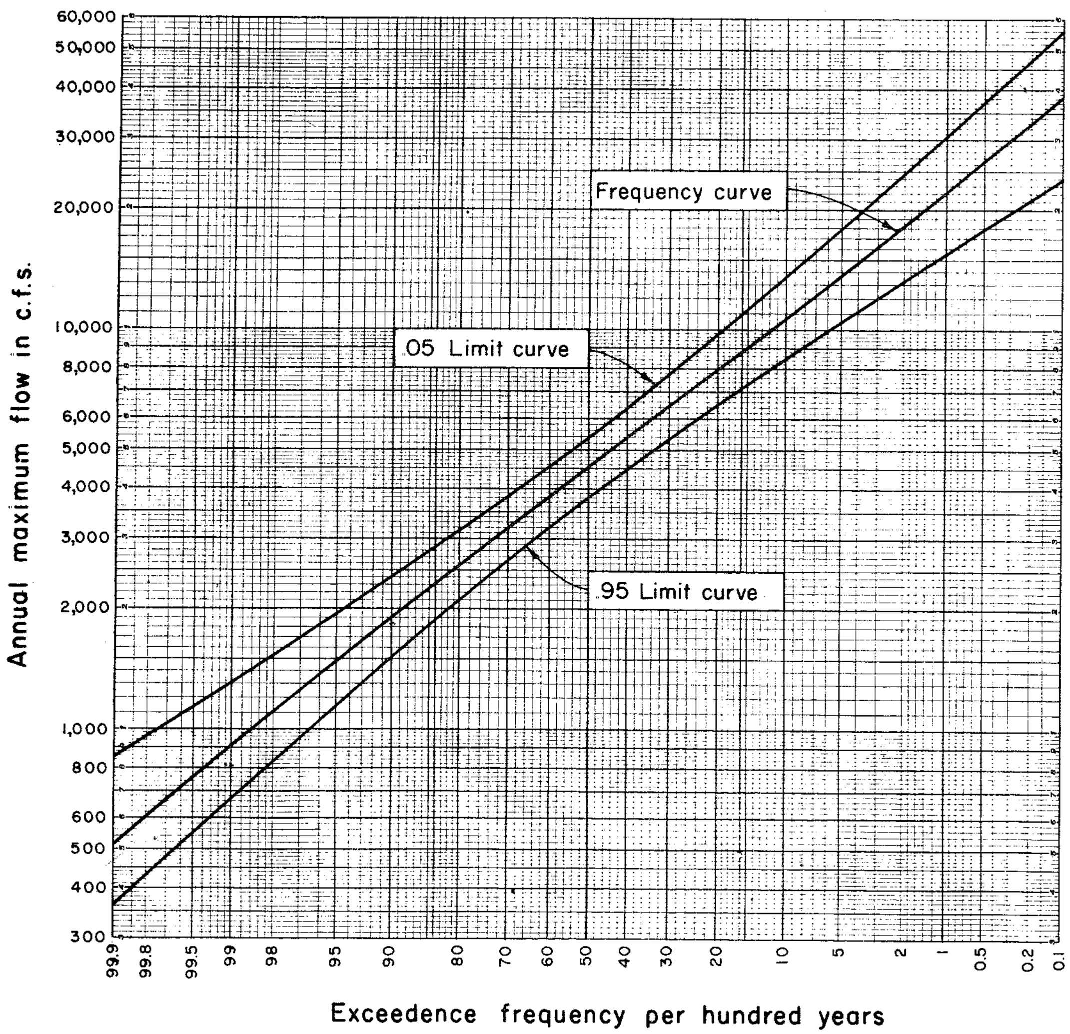
The reliability of analytical frequency determinations can best be illustrated by establishing error-limit curves. The error of the estimated flow for any given frequency is a function of the errors in estimating the mean and standard deviation, assuming that the skew coefficient is known. Criteria for construction of error-limit curves are based on the distribution of the "non-central t." Selected values transformed for convenient use are given on Exhibit 6. By use of this exhibit and equation 4-5, error-limit curves shown on fig. 9.01 were calculated as illustrated on that figure. While the expected frequency is that shown by the middle curve, there is one chance in 20 that the true value for any given frequency is greater than that indicated by the .05 curve and one chance in 20 that it is smaller than the value indicated by the .95 curve. There are, therefore, nine chances in 10 that the true value lies between the .05 and .95 curves.

LIMIT-CURVE COMPUTATION

(Based on equivalent of 41-year record)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(9) P_∞	0.1	1.0	10	50	90	99	99.9
(10) k (Ex 39)	3.09	2.33	1.28	0	-1.28	-2.33	-3.09
(11) Log Q (Eq 5)	4.529	4.314	4.017	3.655	3.293	2.996	2.781
(12) Q, cfs	33,800	20,600	10,400	4,520	1,960	991	604
(13) P_N, (Ex 40) (Plot Q vs. P_N)	0.20	1.33	10.6	50	89.4	98.67	99.80
(14) .05 error in S units (Ex 6)	.76	.60	.41	.27	.31	.43	.53
(15) .05 error, log	.215	.170	.116	.076	.088	.122	.150
(16) .05 limit-curve value, log	4.744	4.484	4.133	3.731	3.381	3.118	2.931
(17) .05 limit-curve value, cfs (Plot vs. P_∞)	55,500	30,500	13,600	5,380	2,400	1,310	853
(18) .95 error in S units (Ex 6)	-.53	-.43	-.31	-.27	-.41	-.60	-.76
(19) .95 error, log	-.150	-.122	-.088	-.076	-.116	-.170	-.215
(20) .95 limit-curve value, log	4.379	4.192	3.929	3.579	3.177	2.826	2.566
(21) .95 limit-curve value, cfs (Plot vs. P_∞)	23,900	15,600	8,490	3,790	1,500	670	368

NOTE: See fig. 4-10 for supporting computations for the frequency curve.



ILLUSTRATIVE EXAMPLE

ERROR-LIMIT CURVES

Selected References

SELECTED REFERENCES

Papers

(1) Beard, L. R., "Estimation of Flood Probabilities," ASCE Proceedings V80, Separate No. 438, May 1954

(2) Beard, L. R., "Probability Estimates Based on Small Normal-Distribution Samples," AGU Journal of Geophysical Research, V65, No. 7, July 1960

(3) Chow, V. T., "Frequency Analysis of Hydrologic Data with Special Application to Rainfall Intensities," University of Illinois Engineering Experiment Station Bulletin Series No. 414, July 1953

(4) Chow, V. T., "The Log Probability Law and Its Engineering Applications," ASCE Proceedings, V80, Separate No. 536, November 1954

(5) Cochran, A. L., "Flood Hydrograph-Damage Integration Method of Estimating Flood Damages in Agricultural Areas," OCE Memo, 30 May 1961

(6) Lane and Kai Lei, "Stream Flow Variability," ASCE Transactions, V115, 1950, p. 1084

(7) Langbein, W. B., "Annual Floods and the Partial Duration Series," Transactions, American Geophysical Union, Vol. 30, p. 879, 1949

(8) Schwarz, "Determination of Flood Frequencies in a Major Drainage Basin," IUGG, International Association of Scientific Hydrology, General Assembly of Toronto, VIII, p. 174, September 1975

(9) Snyder, Franklin F., "Synthetic Flood Frequency," ASCE Proceedings, Journal of Hydraulics Division, V84, Paper 1808, October 1958

(10) Stall, J. B., "Low Flows of Illinois Streams for Impounding Reservoir Design," 1964

(11) Todd, "Frequency Analysis of Stream Flow Data," ASCE Proceedings, V84, Paper No. 1166, February 1957

Civil Works Investigations Reports, Washington District, Corps of Engineers

(12) Technical Report No. 1, "Stream Flow Volume-Duration Frequency Studies," June 1955

(13) Technical Bulletin No. 1, "Stream-Flow Volume-Duration Frequency Studies," June 1960

Civil Works Investigations Reports, Sacramento District, Corps of Engineers

(14) Technical Bulletin No. 1, "Analysis of Flood Frequencies by Season," October 1954

(15) Technical Bulletin No. 2, "Coordination of Flood Volume Frequency Statistics," July 1955

(16) Technical Bulletin No. 3, "Flood Volume Frequency Statistics for Pacific Coast Streams," March 1956

(17) Technical Bulletin No. 4, "Ten-Year Storm Precipitation in California and Oregon Coastal Basins," May 1957

(18) Technical Bulletin No. 5, "Long-Duration Runoff Volumes," July 1958

(19) Technical Bulletin No. 6, "Runoff Frequency Curves for Great Basin Streams," December 1959

(20) Technical Bulletin No. 7, Generalized Rainflood Frequencies for California, Oregon and Washington," September 1960

(21) Research Note No. 1, "Frequency of New England Floods," July 1958

(22) Research No. 2, "Probability Estimates Based on Small Normal-Distribution Samples," March 1959

(23) Research No. 3, "Frequency Curves from Incomplete Samples," September 1959

(24) Research No. 4, "A Reliability Test of Hydrologic Frequency Methods," July 1960

Texts

(25) Benjamin, J. R., & Cornell, C. A., "Probability, Statistics, and Decision for Civil Engineers," McGraw-Hill, 1970

(26) Croxton & Cowden, "Applied General Statistics," Prentice-Hall, 1955

(27) Ezekiel, M., "Methods of Correlation Analysis," Wiley & Sons, 1941

(28) Fryer, H. C., "Concepts and Methods of Experimental Statistics," Allyn & Bacon, 1966

(29) Guenther, W. C., "Concepts of Statistical Inference," McGraw-Hill, 1965

(30) Gumbel, E. J., "Statistics of Extremes," Columbia University Press, 1958

(31) Nielson, "Methods in Numerical Analysis," Macmillan, 1956

(32) Sterling, T. D., & Pollack, S. V., "Introduction to Statistical Data Processing," Prentice-Hall, 1956

Computer Program Documents, Hydrologic Engineering Center

(33) Regional Frequency Computation, Computer Program 723-X6-L2350, July 1972

Exhibits

PLOTTING POSITIONS IN PERCENT
(EXCEEDENCE FREQUENCY IN EVENTS PER HUNDRED YEARS)

m \ N	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	N \ m
	50.0	29.3	20.6	15.9	13.0	10.9	9.43	8.30	7.43	6.70	6.11	5.62	5.20	4.83	4.52	4.25	4.00	3.78	3.58	3.41	3.25	3.10	1
		70.7	50.0	38.6	31.5	26.5	22.9	20.2	18.1	16.3	14.9	13.7	12.7	11.8	11.0	10.3	9.8	9.2	8.8	8.3	7.9	7.6	2
			79.4	61.4	50.0	42.2	36.5	32.1	28.7	25.9	23.7	21.8	20.1	18.7	17.5	16.4	15.5	14.7	13.9	13.2	12.6	12.0	3
44	98.4			84.1	68.5	57.8	50.0	44.0	39.4	35.6	32.4	29.8	27.6	25.7	24.0	22.5	21.2	20.1	19.1	18.1	17.3	16.5	4
43	96.2	98.4			87.0	73.5	63.5	56.0	50.0	45.2	41.2	37.9	35.1	32.6	30.5	28.6	27.0	25.5	24.2	23.0	22.0	21.0	5
42	93.9	96.1	98.4			89.1	77.1	67.9	60.6	54.8	50.0	46.0	42.5	39.6	37.0	34.7	32.8	31.0	29.4	27.9	26.6	25.4	6
41	91.7	93.8	96.0	98.3			90.6	79.8	71.3	64.4	58.8	54.0	50.0	46.5	43.5	40.8	38.5	36.4	34.5	32.8	31.3	29.9	7
40	89.4	91.5	93.6	95.9	98.3			91.7	81.9	74.1	67.6	62.1	57.5	53.5	50.0	46.9	44.2	41.8	39.7	37.7	36.0	34.4	8
39	87.2	89.2	91.3	93.5	95.8	98.2			92.6	83.7	76.3	70.2	64.9	60.4	56.5	53.1	50.0	47.3	44.8	42.6	40.6	38.8	9
38	84.9	86.9	88.9	91.1	93.3	95.7	98.2			93.3	85.1	78.2	72.4	67.4	63.0	59.2	55.8	52.7	50.0	47.5	45.3	43.3	10
37	82.7	84.6	86.6	88.7	90.8	93.2	95.6	98.1			93.9	86.3	79.9	74.3	69.5	65.3	61.5	58.2	55.2	52.5	50.0	47.8	11
36	80.4	82.3	84.2	86.2	88.4	90.6	93.0	95.5	98.1			94.4	87.3	81.3	76.0	71.4	67.2	63.6	60.3	57.4	54.7	52.2	12
35	78.2	80.0	81.9	83.8	85.9	88.1	90.4	92.8	95.3	98.0			94.8	88.2	82.5	77.5	73.0	69.0	65.5	62.3	59.4	56.7	13
34	75.9	77.4	79.5	81.4	83.4	85.5	87.8	90.1	92.6	95.2	98.0			95.2	89.0	83.6	78.8	74.5	70.6	67.2	64.0	61.2	14
33	73.7	75.4	77.1	79.0	81.0	83.0	85.2	87.4	89.8	92.4	95.1	97.9			95.5	89.7	84.5	79.9	75.8	72.1	68.7	65.6	15
32	71.4	73.0	74.8	76.6	78.5	80.5	82.6	84.8	87.1	89.6	92.2	94.9	97.9			95.8	90.2	85.3	80.9	77.0	73.4	70.1	16
31	69.2	70.7	72.4	74.2	76.0	77.9	80.0	82.1	84.3	86.7	89.3	91.9	94.8	97.8			96.0	90.8	86.1	81.9	78.0	74.6	17
30	66.9	68.4	70.1	71.7	73.5	75.4	77.4	79.4	81.6	83.9	86.3	88.9	91.7	94.6	97.7			96.2	91.2	86.8	82.7	79.0	18
29	64.6	66.1	67.7	69.3	71.0	72.9	74.7	76.7	78.9	81.1	83.4	85.9	88.6	91.4	94.4	97.6			96.4	91.7	87.4	83.5	19
28	62.4	63.8	65.3	66.9	68.6	70.3	72.1	74.1	76.1	78.3	80.5	82.9	85.5	88.2	91.1	94.2	97.6			96.6	92.1	88.0	20
27	60.1	61.5	63.0	64.5	66.1	67.8	69.5	71.4	73.4	75.4	77.6	79.9	82.4	85.0	87.8	90.8	94.1	97.5			96.8	92.4	21
26	57.9	59.2	60.6	62.1	63.6	65.2	66.9	68.7	70.6	72.6	74.7	77.0	79.3	81.9	84.6	87.4	90.5	93.8	97.4			96.9	22
25	55.6	56.9	58.3	59.7	61.1	62.7	64.3	66.0	67.9	69.8	71.8	74.0	76.2	78.7	81.3	84.0	87.0	90.2	93.6	97.3			
24	53.4	54.6	55.9	57.2	58.7	60.2	61.7	63.4	65.1	67.0	68.9	71.0	73.2	75.5	78.0	80.6	83.5	86.5	89.8	93.3	97.2		
23	51.1	52.3	53.5	54.8	56.2	57.6	59.1	60.7	62.4	64.1	66.0	68.0	70.1	72.3	74.7	77.2	80.0	82.9	86.0	89.4	93.0	97.0	23
22	48.9	50.0	51.2	52.4	53.7	55.1	56.5	58.0	59.6	61.3	63.1	65.0	67.0	69.1	71.4	73.8	76.4	79.2	82.2	85.5	89.0	92.8	22
21	46.6	47.7	48.8	50.0	51.2	52.5	53.9	55.3	56.9	58.5	60.2	62.0	63.9	65.9	68.1	70.4	72.9	75.6	78.4	81.5	84.8	88.5	21
20	44.4	45.4	46.5	47.6	48.8	50.0	51.3	52.7	54.1	55.7	57.3	59.0	60.8	62.7	64.8	67.0	69.4	71.9	74.6	77.6	80.8	84.2	20
19	42.1	43.1	44.1	45.2	46.3	47.5	48.7	50.0	51.4	52.8	54.4	56.0	57.9	59.6	61.5	63.6	65.9	68.3	70.8	73.6	76.6	79.9	19
18	39.9	40.8	41.7	42.8	43.8	44.9	46.1	47.3	48.6	50.0	51.5	53.0	54.6	56.4	58.2	60.2	62.3	64.6	67.1	69.7	72.6	75.7	18
17	37.6	38.5	39.4	40.3	41.3	42.4	43.5	44.7	45.9	47.2	48.5	50.0	51.5	53.2	54.9	56.8	58.8	61.0	63.3	65.8	68.4	71.4	17
16	35.4	36.2	37.0	37.9	38.9	39.8	40.9	42.0	43.1	44.3	45.6	47.0	48.5	50.0	51.6	53.4	55.3	57.3	59.5	61.8	64.4	67.1	16
15	33.1	33.9	34.7	35.5	36.4	37.3	38.3	39.3	40.4	41.5	42.7	44.0	45.4	46.8	48.4	50.0	51.8	53.7	55.7	57.9	60.2	62.8	15
14	30.8	31.6	32.3	33.1	33.9	34.8	35.7	36.6	37.6	38.7	39.8	41.0	42.1	43.6	45.1	46.6	48.2	50.0	51.9	53.9	56.2	58.6	14
13	28.6	29.3	29.9	30.7	31.4	32.2	33.1	34.0	34.9	35.9	36.9	38.0	39.2	40.4	41.8	43.2	44.7	46.3	48.1	50.0	52.0	54.3	13
12	26.3	27.0	27.6	28.3	29.0	29.7	30.5	31.3	32.1	33.0	34.0	35.0	36.1	37.3	38.5	39.8	41.2	42.7	44.3	46.1	48.0	50.0	12
11	24.1	24.6	25.2	25.8	26.5	27.1	27.9	28.6	29.4	30.2	31.1	32.0	33.0	34.1	35.2	36.4	37.7	39.0	40.5	42.1	43.8	45.7	11
10	21.8	22.3	22.9	23.4	24.0	24.6	25.3	25.9	26.6	27.4	28.2	29.0	29.9	30.9	31.9	33.0	34.1	35.4	36.7	38.2	39.8	41.4	10
9	19.6	20.0	20.5	21.0	21.5	22.1	22.6	23.3	23.9	24.6	25.3	26.0	26.8	27.7	28.6	29.6	30.6	31.7	32.9	34.2	35.6	37.2	9
8	17.3	17.7	18.1	18.6	19.0	19.5	20.0	20.6	21.1	21.7	22.4	23.0	23.8	24.5	25.3	26.2	27.1	28.1	29.2	30.3	31.6	32.9	8
7	15.1	15.4	15.8	16.2	16.6	17.0	17.4	17.9	18.4	18.9	19.5	20.1	20.7	21.3	22.0	22.8	23.6	24.4	25.4	26.4	27.4	28.6	7
6	12.8	13.1	13.4	13.8	14.1	14.5	14.8	15.2	15.7	16.1	16.6	17.1	17.6	18.1	18.7	19.4	20.0	20.8	21.6	22.4	23.4	24.3	6
5	10.6	10.8	11.1	11.3	11.6	11.9	12.2	12.6	12.9	13.3	13.7	14.1	14.5	15.0	15.4	16.0	16.5	17.1	17.8	18.5	19.2	20.1	5
4	8.3	8.5	8.7	8.9	9.2	9.4	9.6	9.9	10.2	10.4	10.7	11.1	11.4	11.8	12.2	12.6	13.0	13.5	14.0	14.5	15.2	15.8	4
3	6.1	6.2	6.4	6.5	6.7	6.8	7.0	7.2	7.4	7.6	7.8	8.1	8.3	8.6	8.9	9.2	9.5	9.8	10.2	10.6	11.0	11.5	3
2	3.8	3.9	4.0	4.1	4.2	4.3	4.4	4.5	4.7	4.8	4.9	5.1	5.2	5.4	5.6	5.8	6.0	6.2	6.4	6.7	7.0	7.2	2
1	1.56	1.60	1.64	1.68	1.72	1.76	1.81	1.86	1.91	1.96	2.02	2.08	2.15	2.22	2.29	2.36	2.45	2.54	2.64	2.74	2.85	2.97	1
m \ N	44	43	42	41	40	39	38	37	36	35	34	33	32	31	30	29	28	27	26	25	24	23	N \ m

EXHIBIT 1 SHEET 1

PLOTTING POSITIONS IN PERCENT
(EXCEEDENCE FREQUENCY IN EVENTS PER HUNDRED YEARS)

m \ N	45	46	47	48	49	50	51	52	53	54	55
1	1.53	1.50	1.47	1.43	1.41	1.38	1.35	1.32	1.30	1.28	1.25
2	3.7	3.7	3.6	3.5	3.4	3.4	3.3	3.2	3.2	3.1	3.1
3	5.9	5.8	5.7	5.6	5.4	5.3	5.2	5.1	5.0	5.0	4.9
4	8.1	8.0	7.8	7.6	7.5	7.3	7.2	7.0	6.9	6.8	6.7
5	10.3	10.1	9.9	9.7	9.5	9.3	9.1	9.0	8.8	8.6	8.5
6	12.5	12.3	12.0	11.8	11.5	11.3	11.1	10.9	10.7	10.5	10.3
7	14.7	14.4	14.1	13.8	13.6	13.3	13.1	12.8	12.5	12.3	12.1
8	17.0	16.7	16.2	15.9	15.6	15.3	15.0	14.7	14.4	14.1	13.9
9	19.2	18.7	18.3	18.0	17.6	17.2	16.9	16.6	16.3	16.0	15.7
10	21.4	20.9	20.5	20.0	19.6	19.2	18.9	18.5	18.2	17.8	17.5
11	23.6	23.1	22.6	22.1	21.6	21.2	20.8	20.4	20.0	19.7	19.3
12	25.8	25.2	24.7	24.2	23.7	23.2	22.8	22.3	21.9	21.5	21.1
13	28.0	27.4	26.8	26.2	25.7	25.2	24.7	24.2	23.8	23.3	22.9
14	30.2	29.5	28.9	28.3	27.7	27.2	26.6	26.1	25.7	25.2	24.7
15	32.4	31.7	31.0	30.4	29.8	29.2	28.6	28.0	27.5	27.0	26.5
16	34.6	33.8	33.1	32.4	31.8	31.1	30.5	30.0	29.4	28.9	28.3
17	36.8	36.0	35.2	34.5	33.8	33.1	32.5	31.9	31.3	30.7	30.1
18	39.0	38.1	37.3	36.6	35.8	35.1	34.4	33.8	33.1	32.5	31.9
19	41.2	40.3	39.4	38.6	37.8	37.1	36.4	35.7	35.0	34.4	33.8
20	43.4	42.5	41.6	40.7	39.9	39.1	38.3	37.6	36.9	36.2	35.6
21	45.6	44.6	43.7	42.8	41.9	41.1	40.3	39.5	38.8	38.0	37.4
22	47.8	46.8	45.8	44.8	43.9	43.1	42.2	41.4	40.6	39.9	39.2
23	50.0	48.9	47.9	46.9	46.0	45.0	44.2	43.3	42.5	41.7	41.0
			50.0	49.0	48.0	47.0	46.1	45.2	44.4	43.6	42.8
					50.0	49.0	48.1	47.1	46.3	45.4	44.6
							50.0	49.0	48.1	47.2	46.4
									50.0	49.1	48.2
											50.0

56	57	58	59	60	61	62	63	64	65	66	N \ m
1.23	1.21	1.19	1.17	1.15	1.13	1.11	1.09	1.08	1.06	1.05	1
3.0	3.0	2.9	2.8	2.8	2.7	2.7	2.7	2.6	2.6	2.5	2
4.8	4.7	4.6	4.5	4.5	4.4	4.3	4.2	4.2	4.1	4.1	3
6.6	6.4	6.3	6.2	6.1	6.0	5.9	5.8	5.7	5.6	5.6	4
8.3	8.2	8.0	7.9	7.8	7.6	7.5	7.4	7.3	7.2	7.1	5
10.1	9.9	9.8	9.6	9.4	9.3	9.1	9.0	8.8	8.7	8.6	6
11.9	11.7	11.5	11.3	11.1	10.9	10.7	10.6	10.4	10.2	10.1	7
13.6	13.4	13.2	13.0	12.7	12.5	12.3	12.1	12.0	11.8	11.6	8
15.4	15.2	14.9	14.6	14.4	14.2	13.9	13.7	13.5	13.3	13.1	9
17.2	16.9	16.6	16.3	16.1	15.8	15.5	15.3	15.1	14.8	14.6	10
19.0	18.6	18.3	18.0	17.7	17.4	17.1	16.9	16.6	16.4	16.1	11
20.7	20.4	20.0	19.7	19.4	19.0	18.7	18.4	18.2	17.9	17.6	12
22.5	22.1	21.7	21.4	21.0	20.7	20.3	20.0	19.7	19.4	19.1	13
24.3	23.9	23.5	23.1	22.7	22.3	21.9	21.6	21.3	20.9	20.6	14
26.1	25.6	25.2	24.7	24.3	23.9	23.6	23.2	22.8	22.5	22.1	15
27.8	27.3	26.9	26.4	26.0	25.6	25.2	24.8	24.4	24.0	23.6	16
29.6	29.1	28.6	28.1	27.6	27.2	26.8	26.3	25.9	25.5	25.1	17
31.4	30.8	30.3	29.8	29.3	28.8	28.4	27.9	27.5	27.1	26.7	18
33.2	32.6	32.0	31.5	31.0	30.5	30.0	29.5	29.0	28.6	28.2	19
34.9	34.3	33.7	33.2	32.6	32.1	31.6	31.1	30.6	30.1	29.7	20
36.7	36.1	35.4	34.8	34.3	33.7	33.2	32.6	32.1	31.6	31.2	21
38.5	37.8	37.2	36.5	35.9	35.3	34.8	34.2	33.7	33.2	32.7	22
40.2	39.5	38.9	38.2	37.6	37.0	36.4	35.8	35.2	34.7	34.2	23
42.0	41.3	40.6	39.9	39.2	38.6	38.0	37.4	36.8	36.2	35.7	24
43.8	43.0	42.3	41.6	40.9	40.2	39.6	39.0	38.4	37.8	37.2	25
45.6	44.8	44.0	43.3	42.5	41.9	41.2	40.5	39.9	39.3	38.7	26
47.3	46.5	45.7	44.9	44.2	43.5	42.8	42.1	41.5	40.8	40.2	27
49.1	48.3	47.4	46.6	45.9	45.1	44.4	43.7	43.0	42.4	41.7	28
	50.0	49.1	48.3	47.5	46.7	46.0	45.3	44.6	43.9	43.2	29
			50.0	49.2	48.4	47.6	46.9	46.1	45.4	44.7	30
					50.0	49.2	48.4	47.7	46.9	46.2	31
							50.0	49.2	48.5	47.7	32
									50.0	49.2	33

NOTES:

1. Plotting positions are symmetrical about 50 percent.

2. For arrays exceeding 66 values, plotting positions can readily be obtained by use of desk calculator and constants given below.

N	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	N
P_1	1.03	1.01	1.00	.99	.97	.96	.94	.93	.92	.91	.90	.88	.87	.86	.85	.84	.83	P_1
ΔP	1.4839	1.4624	1.4412	1.4206	1.4008	1.3814	1.3628	1.3444	1.3265	1.3091	1.2921	1.2758	1.2597	1.2440	1.2288	1.2138	1.1993	ΔP

N	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	N
P_1	.82	.81	.80	.79	.78	.78	.77	.76	.75	.74	.73	.73	.72	.71	.71	.70	.69	P_1
ΔP	1.1851	1.1712	1.1576	1.1444	1.1315	1.1186	1.1063	1.0942	1.0824	1.0709	1.0596	1.0483	1.0375	1.0269	1.0163	1.0061	.9962	ΔP

EXHIBIT 1 SHEET 2

PEARSON TYPE III COORDINATES

g (Skew coefficient)	k = Magnitude in standard deviations from mean for exceedence percentages of:												
	0.01	0.1	1.0	5	10	30	50	70	90	95	99	99.9	99.99
1.0	5.92	4.54	3.03	1.87	1.34	0.38	-0.16	-0.61	-1.12	-1.31	-1.59	-1.80	-1.88
0.8	5.48	4.25	2.90	1.83	1.34	0.42	-0.13	-0.60	-1.16	-1.38	-1.74	-2.03	-2.18
0.6	5.04	3.96	2.77	1.79	1.33	0.45	-0.09	-0.58	-1.19	-1.45	-1.88	-2.28	-2.53
0.4	4.60	3.67	2.62	1.74	1.32	0.48	-0.06	-0.57	-1.22	-1.51	-2.03	-2.54	-2.92
0.2	4.16	3.38	2.48	1.69	1.30	0.51	-0.03	-0.55	-1.25	-1.58	-2.18	-2.81	-3.32
0.0	3.73	3.09	2.33	1.64	1.28	0.52	0.00	-0.52	-1.28	-1.64	-2.33	-3.09	-3.73
-0.2	3.32	2.81	2.18	1.58	1.25	0.55	0.03	-0.51	-1.30	-1.69	-2.48	-3.38	-4.16
-0.4	2.92	2.54	2.03	1.51	1.22	0.57	0.06	-0.48	-1.32	-1.74	-2.62	-3.67	-4.60
-0.6	2.53	2.28	1.88	1.45	1.19	0.58	0.09	-0.45	-1.33	-1.79	-2.77	-3.96	-5.04
-0.8	2.18	2.03	1.74	1.38	1.16	0.60	0.13	-0.42	-1.34	-1.83	-2.90	-4.25	-5.48
-1.0	1.88	1.80	1.59	1.31	1.12	0.61	0.16	-0.38	-1.34	-1.87	-3.03	-4.54	-5.92
Skew Coefficients Commonly Used													
.00	3.73	3.09	2.33	1.64	1.28	0.52	0.00	-0.52	-1.28	-1.64	-2.33	-3.09	-3.73
-.04	3.65	3.03	2.30	1.63	1.27	0.53	0.01	-0.52	-1.28	-1.65	-2.36	-3.15	-3.82
-.12	3.48	2.92	2.24	1.60	1.26	0.54	0.02	-0.51	-1.29	-1.67	-2.42	-3.26	-3.99
-.23	3.26	2.77	2.16	1.57	1.25	0.55	0.03	-0.50	-1.30	-1.70	-2.50	-3.42	-4.23
-.32	3.08	2.68	2.09	1.54	1.23	0.56	0.05	-0.49	-1.31	-1.72	-2.56	-3.55	-4.42
-.37	2.98	2.58	2.05	1.52	1.22	0.57	0.06	-0.48	-1.32	-1.73	-2.60	-3.63	-4.53
-.40	2.92	2.54	2.03	1.51	1.22	0.57	0.06	-0.48	-1.32	-1.74	-2.62	-3.67	-4.60

NOTE: Approximate transformations between normal deviate (X) and Pearson Type III deviate k can be accomplished with the following equation:

$$k = \frac{2}{g}\left\{\left[\frac{g}{6}\left(X - \frac{g}{6}\right) + 1\right]^{3} - 1\right\}$$

EXHIBIT 2

TABLE OF P_N VERSUS P_∞ IN PERCENT

For use with samples drawn from a normal population

N-1 \ P_∞	50.0	30.0	10.0	5.0	1.0	0.1	0.01
1	50.0	37.2	24.3	20.4	15.4	12.1	10.2
2	50.0	34.7	19.3	14.6	9.0	5.7	4.3
3	50.0	33.6	16.9	11.9	6.4	3.5	2.3
4	50.0	33.0	15.4	10.4	5.0	2.4	1.37
5	50.0	32.5	14.6	9.4	4.2	1.79	.92
6	50.0	32.2	13.8	8.8	3.6	1.38	.66
7	50.0	31.9	13.5	8.3	3.2	1.13	.50
8	50.0	31.7	13.1	7.9	2.9	.94	.39
9	50.0	31.6	12.7	7.6	2.7	.82	.31
10	50.0	31.5	12.5	7.3	2.5	.72	.25
11	50.0	31.4	12.3	7.1	2.3	.64	.21
12	50.0	31.3	12.1	6.9	2.2	.58	.18
13	50.0	31.2	11.9	6.8	2.1	.52	.16
14	50.0	31.1	11.8	6.7	2.0	.48	.14
15	50.0	31.1	11.7	6.6	1.96	.45	.13
16	50.0	31.0	11.6	6.5	1.90	.42	.12
17	50.0	31.0	11.5	6.4	1.84	.40	.11
18	50.0	30.9	11.4	6.3	1.79	.38	.10
19	50.0	30.9	11.3	6.2	1.74	.36	.091
20	50.0	30.8	11.3	6.2	1.70	.34	.084
21	50.0	30.8	11.2	6.1	1.67	.33	.078
22	50.0	30.8	11.1	6.1	1.63	.31	.073
23	50.0	30.7	11.1	6.0	1.61	.30	.068
24	50.0	30.7	11.0	6.0	1.58	.29	.064
25	50.0	30.7	11.0	5.9	1.55	.28	.060
26	50.0	30.6	10.9	5.9	1.53	.27	.057
27	50.0	30.6	10.9	5.9	1.51	.26	.054
28	50.0	30.6	10.9	5.8	1.49	.26	.051
29	50.0	30.6	10.8	5.8	1.47	.25	.049
30	50.0	30.6	10.8	5.8	1.45	.24	.046
40	50.0	30.4	10.6	5.6	1.33	.20	.034
60	50.0	30.3	10.4	5.4	1.22	.16	.025
120	50.0	30.2	10.2	5.2	1.11	.13	.017
∞	50.0	30.0	10.0	5.0	1.00	.10	.010

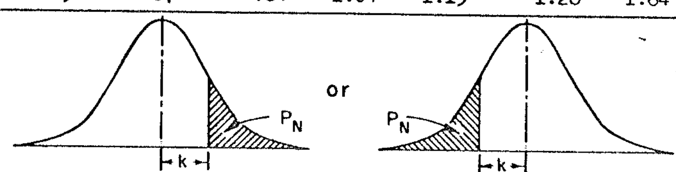
NOTE: P_N values above are usable approximately with Pearson Type III distributions having small skew coefficients.

TABLE OF k VERSUS P_N

For use with samples drawn from a normal population.

N-1 \ P_N (%)	50.0	45.0	40.0	35.0	30.0	25.0	20.0	15.0	12.5	10.0	5.0	2.5	1.25	1.0	0.5	0.25	0.05
1	.00	.19	.40	.62	.89	1.22	1.69	2.40	2.96	3.77	7.73	15.56	31.17	38.97	77.96	155.93	779.70
2	.00	.16	.33	.51	.71	.94	1.23	1.60	1.85	2.18	3.37	4.97	7.16	8.04	11.46	16.27	36.49
3	.00	.15	.31	.47	.65	.86	1.09	1.40	1.59	1.83	2.63	3.56	4.67	5.08	6.53	8.33	14.47
4	.00	.15	.30	.45	.62	.81	1.03	1.30	1.47	1.68	2.34	3.04	3.83	4.10	5.04	6.13	9.43
5	.00	.14	.29	.44	.60	.79	.99	1.25	1.41	1.59	2.18	2.78	3.42	3.63	4.36	5.16	7.41
6	.00	.14	.28	.43	.59	.77	.97	1.21	1.36	1.54	2.08	2.62	3.17	3.36	3.96	4.62	6.37
7	.00	.14	.28	.43	.58	.75	.95	1.19	1.33	1.50	2.01	2.51	3.01	3.18	3.71	4.27	5.73
8	.00	.14	.28	.42	.58	.74	.94	1.17	1.31	1.47	1.96	2.43	2.90	3.05	3.54	4.04	5.31
9	.00	.14	.27	.42	.57	.74	.93	1.15	1.29	1.45	1.92	2.37	2.82	2.96	3.41	3.87	5.01
10	.00	.13	.27	.41	.57	.73	.92	1.14	1.28	1.43	1.89	2.33	2.75	2.89	3.31	3.74	4.79
11	.00	.13	.27	.41	.56	.73	.91	1.13	1.26	1.42	1.87	2.29	2.70	2.83	3.23	3.64	4.62
12	.00	.13	.27	.41	.56	.72	.91	1.12	1.25	1.41	1.85	2.26	2.66	2.78	3.17	3.56	4.48
13	.00	.13	.27	.41	.56	.72	.90	1.12	1.25	1.40	1.83	2.24	2.62	2.74	3.12	3.49	4.37
14	.00	.13	.27	.41	.55	.71	.90	1.11	1.24	1.39	1.82	2.22	2.59	2.71	3.07	3.44	4.28
15	.00	.13	.27	.41	.55	.71	.89	1.11	1.23	1.38	1.81	2.20	2.57	2.68	3.04	3.39	4.20
16	.00	.13	.27	.40	.55	.71	.89	1.10	1.23	1.38	1.80	2.18	2.54	2.66	3.01	3.35	4.13
17	.00	.13	.26	.40	.55	.71	.89	1.10	1.22	1.37	1.79	2.17	2.53	2.64	2.98	3.31	4.07
18	.00	.13	.26	.40	.55	.71	.88	1.09	1.22	1.36	1.78	2.16	2.51	2.62	2.95	3.28	4.02
19	.00	.13	.26	.40	.55	.70	.88	1.09	1.22	1.36	1.77	2.14	2.49	2.60	2.93	3.25	3.98
20	.00	.13	.26	.40	.55	.70	.88	1.09	1.21	1.36	1.77	2.14	2.48	2.59	2.91	3.23	3.94
21	.00	.13	.26	.40	.54	.70	.88	1.09	1.21	1.35	1.76	2.13	2.47	2.57	2.89	3.21	3.90
22	.00	.13	.26	.40	.54	.70	.88	1.08	1.21	1.35	1.75	2.12	2.46	2.56	2.88	3.19	3.87
23	.00	.13	.26	.40	.54	.70	.88	1.08	1.20	1.35	1.75	2.11	2.45	2.55	2.86	3.17	3.84
24	.00	.13	.26	.40	.54	.70	.87	1.08	1.20	1.34	1.74	2.10	2.44	2.54	2.85	3.15	3.82
25	.00	.13	.26	.40	.54	.70	.87	1.08	1.20	1.34	1.74	2.10	2.43	2.53	2.84	3.14	3.80
26	.00	.13	.26	.40	.54	.70	.87	1.08	1.20	1.34	1.74	2.09	2.42	2.52	2.83	3.12	3.78
27	.00	.13	.26	.40	.54	.70	.87	1.08	1.20	1.34	1.73	2.09	2.42	2.52	2.82	3.11	3.76
28	.00	.13	.26	.40	.54	.69	.87	1.07	1.20	1.34	1.73	2.08	2.41	2.51	2.81	3.10	3.74
29	.00	.13	.26	.40	.54	.69	.87	1.07	1.19	1.33	1.73	2.08	2.40	2.50	2.80	3.09	3.72
30	.00	.13	.26	.40	.54	.69	.87	1.07	1.19	1.33	1.72	2.07	2.40	2.50	2.79	3.08	3.70
40	.00	.13	.26	.39	.54	.69	.86	1.06	1.18	1.32	1.70	2.05	2.36	2.45	2.74	3.01	3.59
60	.00	.13	.26	.39	.53	.68	.85	1.05	1.17	1.31	1.68	2.02	2.32	2.41	2.68	2.94	3.49
120	.00	.13	.26	.39	.53	.68	.85	1.05	1.16	1.29	1.66	1.99	2.28	2.37	2.63	2.87	3.39
∞	.00	.13	.25	.38	.52	.67	.84	1.04	1.15	1.28	1.64	1.96	2.24	2.33	2.58	2.81	3.29

This table shows k for values of P_N as illustrated:



Note: Because of the symmetry illustrated at the left, k values for frequencies greater than 50% are obtained by subtracting the desired frequency from 100, entering the table with the difference and the proper value of N-1, reading the k value, and multiplying the value by -1.

EXHIBIT 4

TRANSFORMED PLOTTING POSITIONS (1)
k VALUES HAVING ZERO MEAN AND UNIT STANDARD DEVIATION

N.	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
Event No. 1	1.71	1.75	1.79	1.82	1.85	1.88	1.91	1.93	1.95	1.97	1.99	2.01	2.03	2.05	2.07	2.09	2.10	2.11	2.12	2.13	2.15	2.16	2.17	2.18	2.19
2	1.11	1.16	1.21	1.26	1.30	1.33	1.36	1.39	1.42	1.45	1.48	1.50	1.52	1.54	1.56	1.58	1.60	1.62	1.64	1.66	1.67	1.68	1.70	1.71	1.72
3	0.76	0.82	0.88	0.93	0.98	1.02	1.06	1.10	1.13	1.16	1.19	1.22	1.24	1.26	1.28	1.30	1.32	1.34	1.36	1.38	1.40	1.42	1.44	1.46	1.47
4	0.48	0.56	0.62	0.68	0.73	0.78	0.83	0.87	0.91	0.94	0.97	1.00	1.03	1.06	1.09	1.11	1.13	1.15	1.17	1.19	1.21	1.23	1.25	1.27	1.29
5	0.25	0.32	0.40	0.47	0.53	0.59	0.64	0.68	0.72	0.76	0.80	0.83	0.86	0.89	0.92	0.95	0.97	0.99	1.01	1.04	1.06	1.08	1.10	1.12	1.14
6	0	0.11	0.20	0.27	0.34	0.41	0.47	0.52	0.56	0.60	0.64	0.68	0.71	0.74	0.77	0.80	0.83	0.86	0.88	0.90	0.93	0.95	0.97	0.99	1.01
7			0	0.09	0.17	0.24	0.30	0.36	0.41	0.46	0.50	0.54	0.58	0.61	0.64	0.67	0.70	0.73	0.76	0.79	0.81	0.83	0.85	0.87	0.89
8					0	0.08	0.15	0.21	0.27	0.32	0.37	0.41	0.45	0.49	0.52	0.55	0.58	0.61	0.64	0.67	0.70	0.73	0.75	0.77	0.79
9							0	0.07	0.13	0.19	0.24	0.29	0.34	0.38	0.42	0.45	0.48	0.51	0.54	0.57	0.60	0.63	0.65	0.67	0.70
10									0	0.07	0.12	0.17	0.22	0.27	0.31	0.35	0.38	0.41	0.44	0.47	0.50	0.53	0.56	0.59	0.61
11											0	0.06	0.11	0.16	0.20	0.24	0.28	0.32	0.35	0.38	0.41	0.44	0.47	0.50	0.52
12													0	0.05	0.10	0.15	0.19	0.23	0.26	0.30	0.33	0.36	0.39	0.42	0.45
13															0	0.05	0.10	0.14	0.17	0.21	0.24	0.28	0.31	0.34	0.37
14																	0	0.06	0.10	0.13	0.16	0.20	0.23	0.26	0.29
15																			0	0.04	0.08	0.12	0.15	0.18	0.22
16																					0	0.04	0.08	0.11	0.14
17																							0	0.04	0.07
18																									0

N	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
Event No. 1	2.20	2.22	2.23	2.24	2.25	2.26	2.27	2.28	2.28	2.29	2.30	2.31	2.31	2.32	2.33	2.33	2.34	2.35	2.36	2.37	2.37	2.38	2.39	2.39	2.40
2	1.74	1.75	1.76	1.77	1.79	1.80	1.81	1.82	1.83	1.84	1.85	1.86	1.87	1.88	1.88	1.89	1.90	1.91	1.92	1.93	1.93	1.94	1.95	1.96	1.97
3	1.48	1.50	1.51	1.53	1.54	1.55	1.56	1.57	1.59	1.60	1.61	1.62	1.63	1.64	1.65	1.66	1.67	1.68	1.69	1.69	1.70	1.71	1.72	1.73	1.74
4	1.30	1.31	1.33	1.35	1.36	1.37	1.39	1.40	1.41	1.42	1.43	1.44	1.45	1.47	1.48	1.49	1.50	1.51	1.52	1.53	1.54	1.55	1.56	1.56	1.57
5	1.15	1.17	1.19	1.20	1.22	1.23	1.25	1.26	1.27	1.28	1.30	1.31	1.32	1.33	1.34	1.35	1.36	1.37	1.38	1.39	1.40	1.41	1.42	1.43	1.44
6	1.03	1.05	1.06	1.08	1.10	1.11	1.12	1.14	1.15	1.17	1.18	1.19	1.20	1.21	1.22	1.24	1.25	1.26	1.27	1.28	1.29	1.30	1.31	1.32	1.33
7	0.91	0.93	0.95	0.97	0.99	1.01	1.02	1.04	1.05	1.06	1.08	1.09	1.10	1.12	1.13	1.14	1.15	1.16	1.18	1.19	1.20	1.21	1.22	1.23	1.24
8	0.81	0.83	0.85	0.87	0.89	0.91	0.92	0.94	0.95	0.97	0.98	1.00	1.01	1.03	1.04	1.05	1.06	1.07	1.09	1.10	1.11	1.12	1.13	1.14	1.15
9	0.72	0.74	0.76	0.78	0.80	0.82	0.84	0.85	0.87	0.88	0.90	0.91	0.93	0.94	0.96	0.97	0.98	0.99	1.00	1.02	1.03	1.04	1.05	1.06	1.07
10	0.63	0.65	0.67	0.70	0.72	0.74	0.75	0.77	0.79	0.80	0.82	0.84	0.85	0.86	0.88	0.89	0.91	0.92	0.93	0.94	0.96	0.97	0.98	0.99	1.00
11	0.55	0.57	0.59	0.62	0.64	0.66	0.68	0.70	0.71	0.73	0.75	0.76	0.78	0.79	0.80	0.82	0.84	0.85	0.86	0.88	0.89	0.90	0.91	0.92	0.94
12	0.47	0.49	0.52	0.54	0.56	0.58	0.60	0.62	0.64	0.66	0.68	0.69	0.71	0.72	0.74	0.75	0.77	0.78	0.80	0.81	0.82	0.84	0.85	0.86	0.87
13	0.39	0.42	0.45	0.47	0.49	0.51	0.53	0.55	0.57	0.59	0.61	0.63	0.64	0.66	0.68	0.69	0.71	0.72	0.74	0.75	0.76	0.78	0.79	0.80	0.81
14	0.32	0.35	0.37	0.40	0.42	0.44	0.46	0.49	0.51	0.53	0.54	0.56	0.58	0.60	0.61	0.63	0.64	0.66	0.68	0.69	0.70	0.72	0.73	0.74	0.76
15	0.25	0.27	0.30	0.33	0.36	0.38	0.40	0.42	0.44	0.46	0.48	0.50	0.52	0.54	0.56	0.57	0.59	0.60	0.62	0.63	0.65	0.66	0.68	0.69	0.70
16	0.17	0.21	0.24	0.26	0.29	0.31	0.34	0.36	0.38	0.40	0.42	0.44	0.46	0.48	0.50	0.51	0.53	0.55	0.56	0.58	0.59	0.61	0.62	0.64	0.65
17	0.11	0.14	0.17	0.20	0.22	0.25	0.27	0.30	0.32	0.34	0.36	0.38	0.40	0.42	0.44	0.46	0.48	0.49	0.51	0.52	0.54	0.56	0.57	0.58	0.60
18	0.04	0.07	0.10	0.13	0.16	0.19	0.21	0.24	0.26	0.28	0.31	0.33	0.35	0.37	0.39	0.40	0.42	0.44	0.46	0.47	0.49	0.51	0.52	0.53	0.55
19		0	0.03	0.06	0.09	0.12	0.15	0.18	0.20	0.22	0.25	0.27	0.29	0.31	0.33	0.35	0.37	0.39	0.41	0.42	0.44	0.46	0.47	0.49	0.50
20				0	0.03	0.06	0.09	0.12	0.14	0.17	0.19	0.22	0.24	0.26	0.28	0.30	0.32	0.34	0.36	0.37	0.39	0.41	0.42	0.44	0.45
21						0	0.03	0.06	0.09	0.11	0.14	0.16	0.18	0.21	0.23	0.25	0.27	0.29	0.31	0.32	0.34	0.36	0.38	0.39	0.41
22								0	0.03	0.06	0.08	0.11	0.13	0.15	0.18	0.20	0.22	0.24	0.26	0.28	0.30	0.31	0.33	0.35	0.36
23										0	0.03	0.05	0.08	0.10	0.13	0.15	0.17	0.19	0.21	0.23	0.25	0.27	0.29	0.30	0.32
24												0	0.03	0.05	0.08	0.10	0.12	0.14	0.16	0.18	0.20	0.22	0.24	0.26	0.28
25														0	0.03	0.05	0.07	0.10	0.12	0.14	0.16	0.18	0.20	0.21	0.23
26																0	0.03	0.05	0.07	0.09	0.11	0.13	0.15	0.17	0.19
27																		0	0.03	0.05	0.07	0.09	0.11	0.13	0.15
28																				0	0.02	0.04	0.07	0.09	0.11
29																						0	0.02	0.04	0.06
30																								0	0.02

N	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85
Event No. 1	2.41	2.41	2.42	2.42	2.43	2.43	2.44	2.44	2.45	2.45	2.46	2.46	2.47	2.47	2.48	2.48	2.49	2.49	2.50	2.50	2.51	2.51	2.51	2.52	2.52
2	1.97	1.98	1.99	1.99	2.00	2.01	2.01	2.02	2.02	2.03	2.04	2.04	2.05	2.05	2.06	2.06	2.07	2.07	2.08	2.08	2.09	2.09	2.10	2.10	2.11
3	1.74	1.75	1.76	1.76	1.77	1.78	1.79	1.79	1.80	1.81	1.81	1.82	1.83	1.83	1.84	1.84	1.85	1.85	1.86	1.86	1.87	1.87	1.88	1.88	1.89
4	1.58	1.59	1.60	1.60	1.61	1.62	1.63	1.63	1.64	1.65	1.66	1.66	1.67	1.68	1.68	1.69	1.69	1.70	1.71	1.71	1.72	1.72	1.73	1.73	1.74
5	1.45	1.46	1.47	1.48	1.48	1.49	1.50	1.51	1.52	1.52	1.53	1.54	1.55	1.55	1.56	1.57	1.57	1.58	1.59	1.59	1.60	1.60	1.61	1.61	1.62
6	1.34	1.35	1.36	1.37	1.38	1.39	1.39	1.40	1.41	1.42	1.43	1.43	1.44	1.45	1.45	1.46	1.47	1.48	1.48	1.49	1.49	1.50	1.50	1.51	1.52
7	1.25	1.26	1.27	1.28	1.29	1.29	1.30	1.31	1.32	1.33	1.33	1.34	1.35	1.36	1.36	1.37	1.38	1.39	1.39	1.40	1.40	1.41	1.42	1.42	1.43
8	1.16	1.17	1.18	1.19	1.20	1.21	1.22	1.23	1.24	1.24	1.25	1.26	1.27	1.28	1.28	1.29	1.30	1.31	1.31	1.32	1.33	1.33	1.34	1.35	1.35
9	1.09	1.10	1.11	1.12	1.13	1.14	1.14	1.15	1.16	1.17	1.18	1.19	1.20	1.20	1.21	1.22	1.23	1.23	1.24	1.25	1.26	1.26	1.27	1.27	1.28
10	1.01	1.02	1.03	1.04	1.05	1.06	1.07	1.08	1.09	1.10	1.11	1.12	1.13	1.14	1.15	1.15	1.16	1.17	1.18	1.18	1.19	1.20	1.20	1.21	1.22
11	0.95	0.96	0.97	0.98	0.99	1.00	1.01	1.02	1.03	1.04	1.05	1.06	1.07	1.07	1.08	1.09	1.10	1.11	1.12	1.12	1.13	1.14	1.14	1.15	1.16
12	0.88	0.90	0.91	0.92	0.93	0.94	0.95	0.96	0.97	0.98	0.99	1.00	1.01	1.02	1.02	1.03	1.04	1.05	1.06	1.06	1.07	1.08	1.09	1.10	1.10
13	0.83	0.84	0.85	0.86	0.87	0.88	0.89	0.90	0.91	0.92	0.93	0.94	0.95	0.96	0.97	0.98	0.99	1.00	1.01	1.01	1.02	1.03	1.04	1.04	1.05
14	0.77	0.78	0.79	0.80	0.82	0.83	0.84	0.85	0.86	0.87	0.88	0.89	0.90	0.91	0.92	0.93	0.94	0.94	0.95	0.96	0.97	0.98	0.99	0.99	1.00
15	0.71	0.73	0.74	0.75	0.76	0.77	0.79	0.80	0.81	0.82	0.83	0.84	0.85	0.86	0.87	0.88	0.89	0.90	0.91	0.91	0.92	0.93	0.94	0.94	0.95
16	0.66	0.68	0.69	0.70	0.71	0.72	0.73	0.75	0.76	0.77	0.78	0.79	0.80	0.81	0.82	0.83	0.84	0.85	0.86	0.86	0.87	0.88	0.89	0.90	0.91
17	0.61	0.63	0.64	0.65	0.66	0.68	0.69	0.70	0.71	0.72	0.73	0.74	0.75	0.77	0.78	0.78	0.79	0.80	0.81	0.82	0.83	0.84	0.85	0.86	0.86
18	0.56	0.58	0.59	0.60	0.62	0.63	0.64	0.65	0.66	0.68	0.69	0.70	0.71	0.72	0.73	0.74	0.75	0.76	0.77	0.78	0.79	0.80	0.81	0.81	0.82
19	0.52	0.53	0.54	0.56	0.57	0.58	0.59	0.61	0.62	0.63	0.64	0.65	0.67	0.68	0.69	0.70	0.71	0.72	0.73	0.74	0.75	0.76	0.77	0.77	0.78
20	0.47	0.48	0.50	0.51	0.53	0.54	0.55	0.56	0.58	0.59	0.60	0.61	0.62	0.63	0.64	0.66	0.67	0.68	0.69	0.70	0.71	0.72	0.73	0.73	0.74
21	0.42	0.44	0.45	0.47	0.48	0.50	0.51	0.52	0.53	0.55	0.56	0.57	0.58	0.59	0.60	0.62	0.63	0.64	0.65	0.66	0.67	0.68	0.69	0.69	0.70
22	0.38	0.39	0.41	0.42	0.44	0.45	0.47	0.48	0.49	0.50	0.52	0.53	0.54	0.55	0.56	0.58	0.59	0.60	0.61	0.62	0.63	0.64	0.65	0.66	0.67
23	0.34	0.35	0.37	0.38	0.40	0.41	0.42	0.44	0.45	0.46	0.48	0.49	0.50	0.51	0.52	0.54	0.55	0.56	0.57	0.58	0.59	0.60	0.61	0.62	0.63
24	0.29	0.31	0.33	0.34	0.35	0.37	0.38	0.40	0.41	0.42	0.44	0.45	0.46	0.48	0.49	0.50	0.51	0.52	0.53	0.54	0.55	0.56	0.57	0.58	0.59
25	0.25	0.27	0.28	0.30	0.31	0.33	0.34	0.36	0.37	0.39	0.40	0.41	0.43	0.44	0.45	0.46	0.47	0.48	0.49	0.51	0.52	0.53	0.54	0.55	0.56
26	0.21	0.23	0.24	0.26	0.27	0.29	0.30	0.32	0.33	0.35	0.36	0.38	0.39	0.40	0.41	0.43	0.44	0.45	0.46	0.47	0.48	0.49	0.50	0.51	0.52
27	0.17	0.18	0.20	0.22	0.23	0.25	0.27	0.28	0.30	0.31	0.32	0.34	0.35	0.36	0.38	0.39	0.40	0.41	0.42	0.44	0.45	0.46	0.47	0.48	0.49
28	0.13	0.14	0.16	0.18	0.19	0.21	0.23	0.24	0.26	0.27	0.29	0.30	0.32	0.33	0.34	0.36	0.37	0.38	0.39	0.40	0.41	0.43	0.44	0.45	0.46
29	0.08	0.10	0.12	0.14	0.16	0.17	0.19	0.20	0.22	0.24	0.25	0.27	0.28	0.29	0.31	0.32	0.33	0.34	0.36	0.37	0.38	0.39	0.40	0.41	0.43
30	0.04	0.06	0.08	0.10	0.12	0.13	0.15	0.17	0.18	0.20	0.21	0.23	0.24	0.26	0.27	0.29	0.30	0.31	0.32	0.33	0.35	0.36	0.37	0.38	0.39
31	0	0.02	0.04	0.06	0.08	0.10	0.11	0.13	0.15	0.16	0.18	0.19	0.21	0.22	0.24	0.25	0.26	0.28	0.29	0.30	0.32	0.33	0.34	0.35	0.36
32			0	0.02	0.04	0.06	0.08	0.09	0.11	0.13	0.14	0.16	0.17	0.19	0.20	0.22	0.23	0.24	0.26	0.27	0.28	0.30	0.31	0.32	0.33
33					0	0.02	0.04	0.06	0.07	0.09	0.11	0.12	0.14	0.15	0.17	0.18	0.20	0.21	0.22	0.24	0.25	0.26	0.28	0.29	0.30
34							0	0.02	0.04	0.06	0.07	0.09	0.11	0.12	0.14	0.15	0.16	0.18	0.19	0.21	0.22	0.23	0.24	0.26	0.27
35									0	0.02	0.04	0.05	0.07	0.09	0.10	0.12	0.13	0.15	0.16	0.17	0.19	0.20	0.21	0.23	0.24
36											0	0.02	0.04	0.05	0.07	0.08	0.10	0.11	0.13	0.14	0.16	0.17	0.18	0.20	0.21
37													0	0.02	0.03	0.05	0.07	0.08	0.10	0.11	0.12	0.14	0.15	0.17	0.18
38															0	0.02	0.03	0.05	0.06	0.08	0.09	0.11	0.12	0.14	0.15
39																	0	0.02	0.03	0.05	0.06	0.08	0.09	0.11	0.12
40																			0	0.02	0.03	0.05	0.06	0.08	0.09
41																					0	0.02	0.03	0.05	0.06
42																							0	0.02	0.03
43																									0

NOTE:

Same footnotes as sheet 2

TRANSFORMED PLOTTING POSITIONS (1)
k VALUES HAVING ZERO MEAN AND UNIT STANDARD DEVIATION

Event No. \ N.	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	110	120	130	140	150	160	170	180	190	200
1	2.53	2.53	2.53	2.54	2.54	2.55	2.55	2.55	2.56	2.56	2.56	2.57	2.57	2.57	2.58	2.61	2.64	2.67	2.69	2.71	2.73	2.75	2.77	2.79	2.81
2	2.11	2.12	2.12	2.13	2.13	2.14	2.14	2.14	2.15	2.15	2.16	2.16	2.16	2.17	2.17	2.21	2.24	2.27	2.30	2.32	2.35	2.37	2.39	2.41	2.43
3	1.89	1.90	1.90	1.91	1.91	1.92	1.92	1.93	1.93	1.94	1.94	1.95	1.95	1.96	1.96	2.00	2.04	2.07	2.10	2.13	2.15	2.18	2.20	2.22	2.24
4	1.74	1.75	1.75	1.76	1.77	1.77	1.78	1.78	1.79	1.79	1.80	1.80	1.80	1.81	1.81	1.85	1.89	1.93	1.96	1.99	2.02	2.04	2.07	2.09	2.11
5	1.62	1.63	1.63	1.64	1.65	1.65	1.66	1.66	1.67	1.67	1.68	1.68	1.69	1.69	1.70	1.74	1.78	1.82	1.85	1.88	1.91	1.94	1.96	1.98	2.00
6	1.52	1.53	1.53	1.54	1.55	1.55	1.56	1.56	1.57	1.58	1.58	1.59	1.59	1.60	1.60	1.64	1.69	1.72	1.76	1.79	1.82	1.85	1.87	1.90	1.92
7	1.44	1.44	1.45	1.45	1.46	1.47	1.47	1.48	1.48	1.49	1.49	1.50	1.50	1.51	1.51	1.56	1.61	1.64	1.68	1.71	1.74	1.77	1.80	1.82	1.85
8	1.36	1.37	1.37	1.38	1.38	1.39	1.40	1.40	1.41	1.41	1.42	1.42	1.43	1.43	1.44	1.49	1.53	1.57	1.61	1.65	1.68	1.70	1.73	1.76	1.78
9	1.29	1.30	1.30	1.31	1.32	1.32	1.33	1.33	1.34	1.34	1.35	1.36	1.36	1.37	1.37	1.42	1.47	1.51	1.55	1.58	1.62	1.65	1.67	1.70	1.72
10	1.22	1.23	1.24	1.25	1.25	1.26	1.26	1.27	1.28	1.28	1.29	1.29	1.30	1.30	1.31	1.36	1.41	1.45	1.49	1.53	1.56	1.59	1.62	1.65	1.67
11	1.17	1.17	1.18	1.19	1.19	1.20	1.21	1.21	1.22	1.22	1.23	1.24	1.24	1.25	1.25	1.31	1.36	1.40	1.44	1.48	1.51	1.54	1.57	1.60	1.62
12	1.11	1.12	1.12	1.13	1.14	1.14	1.15	1.16	1.17	1.17	1.18	1.18	1.19	1.19	1.20	1.26	1.31	1.35	1.39	1.43	1.46	1.49	1.52	1.55	1.58
13	1.06	1.07	1.07	1.08	1.09	1.09	1.10	1.11	1.11	1.12	1.12	1.13	1.14	1.14	1.15	1.21	1.26	1.30	1.34	1.38	1.42	1.45	1.48	1.51	1.53
14	1.01	1.02	1.02	1.03	1.04	1.04	1.05	1.06	1.07	1.07	1.08	1.09	1.09	1.10	1.10	1.16	1.21	1.26	1.30	1.34	1.38	1.41	1.44	1.47	1.49
15	0.96	0.97	0.98	0.98	0.99	1.00	1.01	1.01	1.02	1.03	1.03	1.04	1.05	1.05	1.06	1.12	1.17	1.22	1.26	1.30	1.34	1.37	1.40	1.43	1.46
16	0.92	0.92	0.93	0.94	0.95	0.95	0.96	0.97	0.98	0.98	0.99	1.00	1.00	1.01	1.01	1.08	1.13	1.18	1.22	1.26	1.30	1.33	1.36	1.39	1.42
17	0.87	0.88	0.89	0.90	0.91	0.91	0.92	0.93	0.93	0.94	0.95	0.95	0.96	0.97	0.97	1.04	1.09	1.14	1.19	1.23	1.26	1.30	1.33	1.36	1.39
18	0.83	0.84	0.85	0.86	0.86	0.87	0.88	0.89	0.89	0.90	0.91	0.92	0.92	0.93	0.93	1.00	1.05	1.11	1.15	1.19	1.23	1.27	1.30	1.33	1.36
19	0.79	0.80	0.81	0.82	0.82	0.83	0.84	0.85	0.85	0.86	0.87	0.88	0.88	0.89	0.90	0.96	1.02	1.07	1.12	1.16	1.20	1.23	1.27	1.30	1.33
20	0.75	0.76	0.77	0.78	0.78	0.79	0.80	0.81	0.82	0.82	0.83	0.84	0.85	0.85	0.86	0.93	0.98	1.04	1.08	1.13	1.17	1.20	1.24	1.27	1.30
21	0.71	0.72	0.73	0.74	0.75	0.76	0.77	0.77	0.78	0.79	0.80	0.81	0.81	0.82	0.82	0.89	0.95	1.00	1.05	1.10	1.14	1.17	1.21	1.24	1.27
22	0.68	0.69	0.69	0.70	0.71	0.72	0.73	0.74	0.75	0.75	0.76	0.77	0.78	0.78	0.79	0.86	0.92	0.97	1.02	1.07	1.11	1.14	1.18	1.21	1.24
23	0.64	0.65	0.66	0.67	0.68	0.69	0.70	0.70	0.71	0.72	0.73	0.73	0.74	0.75	0.76	0.83	0.89	0.94	0.99	1.04	1.08	1.12	1.15	1.18	1.21
24	0.60	0.61	0.62	0.63	0.64	0.65	0.66	0.67	0.68	0.68	0.69	0.70	0.71	0.71	0.72	0.79	0.86	0.91	0.96	1.01	1.05	1.09	1.12	1.16	1.19
25	0.57	0.58	0.59	0.60	0.61	0.62	0.63	0.63	0.64	0.65	0.66	0.67	0.68	0.68	0.69	0.76	0.83	0.88	0.93	0.98	1.02	1.06	1.10	1.13	1.16
26	0.53	0.55	0.56	0.57	0.57	0.58	0.59	0.60	0.61	0.62	0.63	0.64	0.64	0.65	0.66	0.73	0.80	0.86	0.91	0.95	1.00	1.04	1.07	1.11	1.14
27	0.50	0.51	0.52	0.53	0.54	0.55	0.56	0.57	0.58	0.59	0.60	0.60	0.61	0.62	0.63	0.70	0.77	0.83	0.88	0.93	0.97	1.01	1.05	1.08	1.11
28	0.47	0.48	0.49	0.50	0.51	0.52	0.53	0.54	0.55	0.56	0.57	0.57	0.58	0.59	0.60	0.67	0.74	0.80	0.85	0.90	0.95	0.99	1.02	1.06	1.09
29	0.44	0.45	0.46	0.47	0.48	0.49	0.50	0.51	0.52	0.53	0.54	0.54	0.55	0.56	0.57	0.65	0.71	0.77	0.83	0.88	0.92	0.96	1.00	1.04	1.07
30	0.41	0.42	0.43	0.44	0.45	0.46	0.47	0.48	0.49	0.50	0.51	0.51	0.52	0.53	0.54	0.62	0.69	0.75	0.80	0.85	0.90	0.94	0.98	1.01	1.05
31	0.37	0.39	0.40	0.41	0.42	0.43	0.44	0.45	0.46	0.47	0.48	0.49	0.49	0.50	0.51	0.59	0.66	0.72	0.78	0.83	0.88	0.92	0.96	0.99	1.03
32	0.34	0.35	0.36	0.38	0.39	0.40	0.41	0.42	0.43	0.44	0.45	0.46	0.46	0.47	0.48	0.56	0.64	0.70	0.76	0.80	0.85	0.90	0.93	0.97	1.00
33	0.31	0.32	0.33	0.35	0.36	0.37	0.38	0.39	0.40	0.41	0.42	0.43	0.44	0.45	0.46	0.54	0.61	0.67	0.73	0.78	0.83	0.87	0.91	0.95	0.98
34	0.28	0.29	0.30	0.32	0.33	0.34	0.35	0.36	0.37	0.38	0.39	0.40	0.41	0.42	0.43	0.51	0.59	0.65	0.71	0.76	0.81	0.85	0.89	0.93	0.96
35	0.25	0.26	0.27	0.29	0.30	0.31	0.32	0.33	0.34	0.35	0.36	0.37	0.38	0.39	0.40	0.49	0.56	0.63	0.69	0.74	0.79	0.83	0.87	0.91	0.94
36	0.22	0.23	0.24	0.26	0.27	0.28	0.29	0.30	0.31	0.32	0.33	0.34	0.35	0.36	0.37	0.46	0.54	0.60	0.66	0.72	0.77	0.81	0.85	0.89	0.92
37	0.19	0.20	0.22	0.23	0.24	0.25	0.26	0.27	0.29	0.30	0.31	0.32	0.33	0.34	0.35	0.43	0.51	0.58	0.64	0.70	0.74	0.79	0.83	0.87	0.91
38	0.16	0.18	0.19	0.20	0.21	0.22	0.23	0.25	0.26	0.27	0.28	0.29	0.30	0.31	0.32	0.41	0.49	0.56	0.62	0.67	0.72	0.77	0.81	0.85	0.89
39	0.13	0.15	0.16	0.17	0.18	0.19	0.21	0.22	0.23	0.24	0.25	0.26	0.27	0.28	0.29	0.39	0.47	0.54	0.60	0.65	0.70	0.75	0.79	0.83	0.87
40	0.10	0.12	0.13	0.14	0.16	0.17	0.18	0.19	0.20	0.21	0.23	0.24	0.25	0.26	0.27	0.36	0.44	0.51	0.58	0.63	0.68	0.73	0.77	0.81	0.85
41	0.07	0.09	0.10	0.11	0.13	0.14	0.15	0.16	0.18	0.19	0.20	0.21	0.22	0.23	0.24	0.34	0.42	0.49	0.55	0.61	0.66	0.71	0.75	0.79	0.83
42	0.04	0.06	0.07	0.09	0.10	0.11	0.12	0.14	0.15	0.16	0.17	0.18	0.19	0.20	0.21	0.31	0.40	0.47	0.53	0.59	0.65	0.69	0.74	0.78	0.82
43	0.02	0.03	0.04	0.06	0.07	0.08	0.10	0.11	0.12	0.13	0.15	0.16	0.17	0.18	0.19	0.29	0.37	0.45	0.51	0.57	0.63	0.68	0.72	0.76	0.80
44		0	0.02	0.03	0.04	0.06	0.07	0.08	0.10	0.11	0.12	0.13	0.14	0.15	0.16	0.27	0.35	0.43	0.49	0.55	0.61	0.66	0.70	0.74	0.78
45				0	0.02	0.03	0.04	0.06	0.07	0.08	0.09	0.10	0.12	0.13	0.14	0.24	0.33	0.41	0.47	0.53	0.59	0.64	0.68	0.73	0.76
46						0	0.02	0.03	0.04	0.05	0.07	0.08	0.09	0.10	0.10	0.22	0.31	0.39	0.45	0.51	0.57	0.62	0.67	0.71	0.75
47								0	0.02	0.03	0.04	0.05	0.06	0.08	0.09	0.19	0.29	0.36	0.43	0.50	0.55	0.60	0.65	0.69	0.73
48										0	0.02	0.03	0.04	0.05	0.06	0.17	0.26	0.34	0.41	0.48	0.53	0.58	0.63	0.67	0.71
49												0	0.01	0.03	0.04	0.15	0.24	0.32	0.40	0.46	0.52	0.57	0.61	0.66	0.70
50														0	0.01	0.13	0.22	0.30	0.38	0.44	0.50	0.55	0.60	0.64	0.68
51																0.10	0.20	0.28	0.36	0.42	0.48	0.53	0.58	0.63	0.67
52																0.08	0.18	0.26	0.34	0.40	0.46	0.52	0.56	0.61	0.65
53																0.06	0.16	0.24	0.32	0.39	0.45	0.50	0.55	0.59	0.64
54																0.03	0.14	0.22	0.30	0.37	0.43	0.48	0.53	0.58	0.62
55																0.01	0.12	0.20	0.28	0.35	0.41	0.47	0.52	0.56	0.60
56																	0.09	0.18	0.26	0.33	0.39	0.45	0.50	0.55	0.59
57																	0.07	0.16	0.24	0.31	0.38	0.43	0.48	0.53	0.58
58																	0.05	0.14	0.23	0.30	0.36	0.42	0.47	0.52	0.56
59																	0.03	0.13	0.21	0.28	0.34	0.40	0.45	0.50	0.55
60																	0.01	0.11	0.19	0.26	0.33	0.39	0.44	0.49	0.53
61																		0.09	0.17	0.24	0.31	0.37	0.42	0.47	0.52
62																		0.07	0.15	0.23	0.29	0.35	0.41	0.46	0.50
63																		0.05	0.13	0.21	0.28	0.34	0.39	0.44	0.49
64																		0.03	0.12	0.19	0.26	0.32	0.38	0.43	0.47
65																		0.01	0.10	0.18	0.24	0.31	0.36	0.41	0.46
66																			0.08	0.16	0.23	0.29	0.35	0.40	0.45
67																			0.06	0.14	0.21	0.28	0.33	0.39	0.43
68																			0.05	0.13	0.20	0.26	0.32	0.37	0.42
69																			0.03	0.11	0.18	0.25	0.30	0.36	0.41
70																			0.01	0.09	0.17	0.23	0.29	0.34	0.39
71																				0.07	0.15	0.22	0.27	0.33	0.38
72																				0.06	0.13	0.20	0.26	0.32	0.37
73																				0.04	0.12	0.19	0.25	0.30	0.35
74																				0.02	0.10	0.17	0.23	0.29	0.34
75																				0.01	0.09	0.16	0.22	0.27	0.33
76																					0.07	0.14	0.20	0.26	0.31
77																					0.05	0.13	0.19	0.25	0.30
78																					0.04	0.11	0.17	0.23	0.29
79																					0.02	0.10	0.16	0.22	0.27
80																					0.01	0.08	0.15	0.21	0.26
81																						0.07	0.13	0.19	0.25
82																						0.05	0.12	0.18	0.23
83																						0.04	0.10	0.17	0.22
84																						0.02	0.09	0.15	0.21
85																						0.01	0.08	0.14	0.20
86																							0.06	0.13	0.18
87																							0.05	0.11	0.17
88																							0.03	0.10	0.16
89																							0.02	0.09	0.15
90																							0.01	0.07	0.13
91																								0.06	0.12
92																								0.05	0.11
93																								0.03	0.09
94																								0.02	0.08
95																								0.01	0.07
96																									0.06
97																									0.04
98																									0.03
99																									0.02
100																									0.01

(1) Linear distance on normal probability grid.

NOTES: k values for lower half of series are negative and equal in magnitude to above values if these are numbered in reverse order starting with N.
N is number of years of record.

ERRORS OF ESTIMATED VALUES

As Coefficients of Standard Deviation

Level of Significance*	Years of Record (N)	Exceedence Frequency in Percent						
		0.1	1	10	50	90	99	99.9
.05	5	4.41	3.41	2.12	.95	.76	1.00	1.22
	10	2.11	1.65	1.07	.58	.57	.76	.94
	15	1.52	1.19	.79	.46	.48	.65	.80
	20	1.23	.97	.64	.39	.42	.58	.71
	30	.93	.74	.50	.31	.35	.49	.60
	40	.77	.61	.42	.27	.31	.43	.53
	50	.67	.54	.36	.24	.28	.39	.49
	70	.55	.44	.30	.20	.24	.34	.42
	100	.45	.36	.25	.17	.21	.29	.37
.25	5	1.41	1.09	.68	.33	.31	.41	.49
	10	.77	.60	.39	.22	.24	.32	.39
	15	.57	.45	.29	.18	.20	.27	.34
	20	.47	.37	.25	.15	.18	.24	.30
	30	.36	.29	.19	.12	.15	.20	.25
	40	.30	.24	.16	.11	.13	.18	.22
	50	.27	.21	.14	.10	.12	.16	.20
	70	.22	.17	.12	.08	.10	.14	.18
	100	.18	.14	.10	.07	.09	.12	.15
.75	5	- .49	- .41	- .31	-.33	- .68	-1.09	-1.41
	10	- .39	- .32	- .24	-.22	- .39	- .60	- .77
	15	- .34	- .27	- .20	-.18	- .29	- .45	- .57
	20	- .30	- .24	- .18	-.15	- .25	- .37	- .47
	30	- .25	- .20	- .15	-.12	- .19	- .29	- .36
	40	- .22	- .18	- .13	-.11	- .16	- .24	- .30
	50	- .20	- .16	- .12	-.10	- .14	- .21	- .27
	70	- .18	- .14	- .10	-.08	- .12	- .17	- .22
	100	- .15	- .12	- .09	-.07	- .10	- .14	- .18
.95	5	-1.22	-1.00	- .76	-.95	-2.12	-3.41	-4.41
	10	- .94	- .76	- .57	-.58	-1.07	-1.65	-2.11
	15	- .80	- .65	- .48	-.46	- .79	-1.19	-1.52
	20	- .71	- .58	- .42	-.39	- .64	- .97	-1.23
	30	- .60	- .49	- .35	-.31	- .50	- .74	- .93
	40	- .53	- .43	- .31	-.27	- .42	- .61	- .77
	50	- .49	- .39	- .28	-.24	- .36	- .54	- .67
	70	- .42	- .34	- .24	-.20	- .30	- .44	- .55
	100	- .37	- .29	- .21	-.17	- .25	- .36	- .45

* Chance of true value being greater than sum of normal-curve value and given error.

Hydrologic Engineering Methods for Water Resources Development

Volume 1	Requirements and General Procedures, 1971
Volume 2	Hydrologic Data Management, 1972
Volume 3	Hydrologic Frequency Analysis, 1975
Volume 4	Hydrograph Analysis, 1973
Volume 5	Hypothetical Floods, 1975
Volume 6	Water Surface Profiles, 1975
Volume 7	Flood Control by Reservoir, 1976
Volume 8	Reservoir Yield, 1975
Volume 9	Reservoir System Analysis for Conservation, 1977
Volume 10	Principles of Groundwater Hydrology, 1972
Volume 11	Water Quality Determinations, 1972
Volume 12	Sediment Transport, 1977