

# Evolution of Methods for Evaluating the Occurrence of Floods

---

GEOLOGICAL SURVEY WATER-SUPPLY PAPER 1580-A



# Evolution of Methods for Evaluating the Occurrence of Floods

By MANUEL A. BENSON

FLOOD HYDROLOGY

---

GEOLOGICAL SURVEY WATER-SUPPLY PAPER 1580-A



**UNITED STATES DEPARTMENT OF THE INTERIOR**

**STEWART L. UDALL, *Secretary***

**GEOLOGICAL SURVEY**

**Thomas B. Nolan, *Director***

## CONTENTS

---

	Page
Abstract.....	A-1
Introduction.....	1
Methods of analyzing flood events.....	2
Empirical formulas.....	3
Empirical formulas involving frequency.....	5
Formulas involving rainfall.....	5
Statistical methods.....	6
Techniques of statistical flood-frequency analysis.....	6
Flood-frequency curves.....	6
Mathematically fitted curves.....	6
Graphical curves.....	10
Types of flood data analyzed.....	11
Plotting position.....	12
Plotting paper.....	14
Regional flood-frequency analysis.....	15
Significance of a single station record.....	15
Synthetic 1,000-year flood study.....	15
Index-flood method.....	16
Base-frequency curves.....	16
Mean annual flood.....	17
Factors related to mean annual flood.....	18
Special treatment for large streams.....	19
Multiple-correlation methods.....	20
Evaluation of the index-flood method.....	20
Design frequency relations at gaged sites.....	23
Predictive value of the flood-frequency relation.....	25
References.....	29

## **FLOOD HYDROLOGY**

---

### **EVOLUTION OF METHODS FOR EVALUATING THE OCCURRENCE OF FLOODS**

---

By **MANUEL A. BENSON**

---

#### **ABSTRACT**

A brief summary is given of the history of methods of expressing flood potentialities, proceeding from simple flood formulas to statistical methods of flood-frequency analysis on a regional basis. Current techniques are described and evaluated. Long-term flood records in the United States show no justification for the adoption of a single type of theoretical distribution of floods. The significance and predictive values of flood-frequency relations are considered. Because of the length of flood records available and the interdependence of flood events within a region, the probable long-term average magnitudes of floods of a given recurrence interval are uncertain. However, if the magnitudes defined by the records available are accepted, the relative effects of drainage-basin characteristics and climatic variables can be determined with a reasonable degree of assurance.

#### **INTRODUCTION**

One of the tasks of the hydrologist is to synthesize a flood hydrograph from the factors combining to produce a flood. This synthesis will enable him to predict or reconstruct a flood occurrence, given the preceding and concurrent physical and meteorological conditions. The task is accomplished by the use of several well-known hydrologic principles, such as hydrograph separation, unit-hydrograph theory, infiltration theory, and flood-routing.

Another related but different task of the hydrologist is the definition of a peak discharge or flood volume of a given frequency or probability in terms of the drainage-basin characteristics and the prevailing meteorologic factors. This problem involves a statistical concept, the  $T$ -year flood, and its solution requires statistical as well as hydrologic techniques.

The two problems are alike in some respects. For example, some factors, such as basin slope or channel slope, may be expected to in-

fluence both the individual flood peak and the peak of a given recurrence interval,  $T$ , called the " $T$ -year" flood peak. However, whereas the individual peak discharge is influenced by the rainfall directly associated with the event, the  $T$ -year peak is related to the prevailing precipitation characteristics. Such characteristics are best expressed by a statistical factor, such as average annual precipitation or precipitation of a short duration and a given frequency. For the individual flood event, antecedent ground conditions as affected by previous precipitation must be taken into account. This and other variable conditions are not relevant in defining the  $T$ -year peak, because the  $T$ -year peak is related to the average of such conditions over a long period of time.

The techniques for construction of an individual flood hydrograph have been studied for a long time, and they are therefore fairly well standardized. The definition of flood-frequency relations is a newer problem, about which there is much yet to be learned. The U.S. Geological Survey plans to study the relation of the  $T$ -year flood to drainage-basin characteristics and to climatic factors. Both terrain and climate vary over the United States and a study of the relations of flood frequency and magnitude over a wide variety of conditions is to be made.

Techniques of flood-frequency analysis are not fixed but are still in a state of change for some of the practices are somewhat controversial. As a preliminary to the study of the relation of flood frequencies to physical and climatic factors, a brief historical review has been made of methods of flood-frequency analysis. The general principles of current methods of statistical analysis and their deficiencies are described in this report. The significance and the predictive value of flood-frequency relations is also discussed.

The material in part A of the present report serves as a general introduction to specific studies that are to follow: a second part B will describe a study of factors that influence the occurrence of floods in a humid region of diverse terrain, and a third part C will describe a study of factors that influence the occurrence of floods in semiarid and arid regions.

This study has been a part of a research project on areal flood frequency. The project leader was M. A. Benson, assisted by D. R. Dawdy, J. Davidian, and M. W. Busby, all engineers.

## METHODS OF ANALYZING FLOOD EVENTS

Information on flood discharges was very meager when the earliest attempts were made to analyze flood events. The flood formulas first devised were therefore simple and generalized. They provided an

estimate of the maximum flood to be expected. Rainfall data was used in many earlier formulas to estimate discharge, as more information was available on rainfall than on discharge. Statistical methods were introduced about 1914 to improve the analysis of flood events. These methods represent a major advancement and recognize that, to determine the most efficient design, it is proper to use something other than the maximum flood to be expected. More recently, it has been recognized that when statistical analysis is applied to the records at a single point, a large sampling error is involved. Knowledge of this has led to the present analytical methods, in which data from a wide region are combined to establish generalized relationships that may be applied anywhere within the region, to gaged or ungaged sites. In flood-frequency analysis, however, knowledge of hydrology has not advanced as fast as statistical applications, and at present the most fruitful field of investigation lies in relating the hydrology of floods to their magnitudes and frequencies.

The desirability of analyzing floods on a frequency basis is now generally recognized, because economic considerations influence decisions made in relation to planning and to design of structures (Langbein and Hoyt, 1959, p. 127-142). The magnitude-frequency relation of floods must be known for the hydroeconomic decisions involved in the construction of highway and railway bridges and culverts; the location of highways, railways, industries, farms, and residences on flood plains; the design of water-supply systems; the design of storm drains; and the design of flood-control structures, such as reservoirs and levees. Economic considerations of flood insurance, which is demanded increasingly by the public, can be based only on a knowledge of the magnitude-frequency relation of floods.

Other factors that control the design of structures on or near rivers are the useful life of the structure, the necessity of maintaining it in continuous operation, and the danger to life and property in case of its failure. However, even to consider these factors, frequency relations must be known.

A comprehensive history of flood-frequency methods was written by Jarvis and others (1936). A brief summary of the basic approaches is given in this report to show the evolution of methods. The principal deficiencies of each are discussed.

#### EMPIRICAL FORMULAS

The earliest approach, and still a prevalent one, to solving the problem of determining floodflows has been by means of floodflow formulas. The earlier formulas, which are entirely empirical, are usually based on few data for a particular region and contain one or

more constants whose values are based on judgment. Such a formula, in generalized form, is:

$$Q = CA^n,$$

where  $Q$  is the floodflow,  $C$  is a coefficient related to the region,  $A$  is the size of drainage area, and  $n$  is a constant with a value between 0.5 and 1.0.

The development of a formula like this is very simple. The peak discharges of some particular region are plotted against drainage area on logarithmic paper. A line is drawn that either averages or envelops the points. The equation of the line will then be of the general form shown above, where  $n$  is the slope of the line and  $C$  is the intercept when  $A$  equals 1. Although the formula is easy to derive, its significance is not clear. It represents either a high average or the maximum of experience in the region, depending on how the line is drawn. There is no frequency significance attached, and the effects of variation in precipitation or terrain have no bearing. Search of European and American engineering literature reveals many formulas of this type.

The Jarvis or modified Myer formula is one such equation:

$$Q = 10,000 \ p \sqrt{A},$$

in which  $C$  is equal to 10,000  $p$ , and  $n$  equals  $\frac{1}{2}$ . In this formula,  $p$  is the ratio of the maximum flood on the given stream to a supposed extreme maximum there. The term "Myers rating," which is commonly used, is a percentage equal to 100  $p$ . For a drainage area of 1 square mile, the 100-percent Myers-rating flood would be 10,000 cfs (cubic feet per second), and the 50-percent flood would be 5,000 cfs.

The various "culvert formulas" used by railroad and highway engineers, such as those by Talbot, Fanning, and McMath, are of this general type. Of these, the formula by Talbot is widely used and is in the form

$$a = CA^{0.75},$$

where  $a$  gives directly the required cross-sectional area of the culvert, in square feet;  $C$  is a coefficient; and  $A$  is the drainage area in acres. This formula gives the required culvert area rather than the peak flow by the simple expedient of dividing the typical flood-flow formula by 10, an assumed allowable velocity in feet per second through the culvert. This formula is a variation of the most primitive type, and the results are subject to much uncertainty because personal judgment is a major factor in the determination of the value of  $C$ .

The simple floodflow formula can be further modified by introducing the effect of other basin characteristics, such as shape, length, or width. In spite of these modifications, the formulas are deficient because they provide no information on flood frequencies.



## EMPIRICAL FORMULAS INVOLVING FREQUENCY

Various empirical formulas have been devised that attempt to introduce the factor of frequency into the computed flood peak. One of these is Fuller's formula:

$$q = \bar{q}(1 + c \log_{10} T_p),$$

where  $q$ , in cubic feet per second, is the flood with a return period of  $T_p$  years,  $\bar{q}$  is the average annual flood, and  $c$  is a coefficient that has been found to vary between 0.7 and 4.5 in different localities.

Another such formula is that of Horton:

$$q = \frac{kT_r^n}{A},$$

in which  $q$  is the flood, in cubic feet per second per square mile, with a recurrence interval of  $T_r$ , in years;  $A$  is the drainage area, in square miles;  $k$ , a constant, has a value of about 4,000; and  $n$  varies with locality, as for example, 0.25 in eastern Pennsylvania.

Both frequency formulas are too general to be of much use. In Fuller's formula, it is necessary to know the mean annual flood from actual records and to determine the  $c$  coefficient empirically. Because no hydrologic factors are used,  $c$  will vary considerably, even within short distances. In Horton's formula, there are two coefficients to be determined empirically. The only hydrologic factor used is drainage-area size, so that the coefficients will remain constant only within regions in which other hydrologic factors vary little or not at all—such regions must of necessity be fairly small.

## FORMULAS INVOLVING RAINFALL

Many floodflow formulas include terms that involve rainfall factors and some measure of the physical characteristics of the basin. Among the better known of these is the Pettis formula. Another, and perhaps the most widely used of such formulas is the rational formula:

$$Q = ci A_d,$$

where  $c$  is a runoff coefficient dependent on various basin characteristics;  $i$  is the rainfall intensity, in inches per hour, during a period equal to the concentration time of the basin and for the same recurrence interval as the discharge; and  $A_d$  is the drainage area, in acres. Linsley, Kohler, and Paulhus (1958, p. 213) comment about this formula as follows:

Of all the flood formulas, the rational formula has the advantage that its physical meaning is reasonably clear. However, it should be used with extreme caution since it does not adequately recognize all the complications of the runoff process.

**STATISTICAL METHODS**

The methods most recently developed and considered to be most promising apply statistical analysis to data of peak discharge. Any list or array of values for a variable, such as the annual flood peak, may be analyzed in several different ways. Statistics provides a means of reducing a mass of data to a few useful and meaningful figures, such as the average, the middle value, or the most common value, which may be used to evaluate the central tendency; or the standard deviation, which may be used to evaluate the spread or dispersion of the data. The distribution of the data may be described by a density function or curve that defines the frequency or relative frequency of occurrence of values of the variable throughout the range. A cumulative distribution function or curve describes the frequency of probability of all values of the variable equal to or greater than some particular value. The term "frequency curve" as used by hydrologists usually refers to what statisticians call a cumulative distribution curve.

Statistics, in addition to providing meaningful figures for describing the magnitude distribution of a variable, provides methods of relating a dependent variable to one or more independent variables. This procedure results in an explanation of the variations in the dependent variable in terms of the other factors and permits predictions of the dependent variable when the other factors are known. This statistical approach is known as correlation or multiple correlation. Johnstone and Cross (1949, p. 236-264) present an excellent summary of statistical methods applicable to hydrologic problems.

**TECHNIQUES OF STATISTICAL FLOOD-FREQUENCY ANALYSIS****FLOOD-FREQUENCY CURVES**

The frequency curve, or cumulative distribution curve, of flood data from a given site is used to determine the floods of specific recurrence intervals or probabilities, such as the 25-year flood or its equivalent, the 0.04-probability flood. A frequency curve may be constructed to represent some type of theoretical curve based on statistical treatment of the data, or it may be a graphical curve based on plotted points.

**MATHEMATICALLY FITTED CURVES**

Various methods of fitting theoretical curves to flood data have been developed by Foster, Hazen, Goodrich, Slade, Gumbel, Chow, and others. The earlier applications have been explained in detail by Jarvis and others (1936). Most of these methods involve the computation of the mean, the standard deviation, and the coefficient of skew of the

flood magnitudes. All require an initial assumption that flood data should fall into a pattern represented by some particular form of theoretical distribution. Proponents of each method have been able to show some degree of confirmation for their particular distributions by comparing theoretical results with known data.

In an attempt to evaluate some of these methods, about 100 records representing the longest flood records obtained in the United States were plotted on various types of graph paper. Each graph paper represented a theoretical distribution. Log-probability paper and extreme-values graph paper were among those used. Plotted data should fit a straight line, if the theoretical and the actual distributions coincide. No one type of plotting gave consistent results over the country. For any type of plotting, the variations from straight lines appeared, in part, to be random and, in part, to show a geographical pattern. There were regional tendencies for curves to be either concave upward or downward. It is concluded that no distribution achieved general confirmation by comparison with actual data, because of the geographical pattern in the departures from straight lines.

One of the reasons advanced for use of a mathematically fitted line is that identical results can be obtained by anyone. This, of course, is true; but experience has shown that the need for use of personal judgment is paramount at times, as for example, in appraising a short record containing data on one or two extraordinarily high floods with true return periods much longer than the period of record. Application here of a least-squares or similar method that abrogates judgment may lead to absurd results. A single high flood may easily outweigh all others and lead to a fitted line that departs from all plotted points. If, on the other hand, judgment is used and the high flood is ignored, curves fitted graphically by several individuals might vary, yet each would be better than a mathematically fitted line. A method is not better simply because it leads to uniform answers, if those answers are uniformly unsound.

Another reason for using a mathematically-fitted function is that the resulting line or formula has no apparent bounds. The assumption, whether or not stated, is that the same statistical distribution holds far beyond the range of the data. Floods of 100, 1,000, or 10,000 years have been evaluated by this method in published reports that are based on less than 50 years of record; and 50-year floods have been computed from 10-year records. Ezekiel and Fox (1959, p. 116), state that:

When there is a good logical basis for the selection of a particular equation, the equation and the corresponding curve can provide a definite logical measurement of the nature of the relationship. When no such logical basis can be developed, a curve fitted by a definite equation yields only an empirical statement of the

relationship and may fail to show the true relation. In such cases a curve fitted freehand by graphic methods, and conforming to logical limitations on its shape, may be even more valuable as a description of the facts of the relationship than a definite equation and corresponding curve selected empirically, but fitting less well.

In any event, estimates of the probable value of the dependent variable cannot be made with any degree of accuracy for values of the independent variable beyond the limits of the cases observed; and can be made most accurately only within the range where a considerable number of observations is available. It may be possible to extrapolate the curve if its equation is based on a logical analysis of the relation as well as on the cases observed; but in that case the logical analysis, and not the statistical examination, must bear the responsibility for the validity of the procedure.

Similarly, Thomas (1948) states:

Indeed, in so far as extrapolation made beyond the range of the recorded floods is concerned, the most elaborate analytical procedure of curve-fitting gives results that are no more reliable than those obtained by a simple extension by eye of the flood-frequency curve on any kind of probability paper.

All methods applied so far have been in a sense empirical, whether or not a theoretical distribution is involved, because it has not been possible to start with the basic hydrologic factors responsible for flood peaks and to demonstrate the nature of the statistical distribution of the peaks. A particular statistical distribution is considered correct only because it shows some agreement with flood data. This is an empirical process of obtaining an expression for flood distribution.

A recently developed and now widely used application of statistical analysis to floods is the Gumbel method. Gumbel (1954, 1958) considered that the annual floods at a station represent a set of extreme values to which the theory of extreme values may be applied. The theory, when applied to floods, starts first with the daily discharges during a given year, which have a skewed distribution with a lower limit of zero and no practical limit at the upper end. The peak daily discharges from each year are then considered; when arrayed, they also have a skewed distribution. They are considered as extreme values of large sets of independent events, to which the extreme-value analysis can be applied. This theory demonstrates a constant value of the coefficient of skew, so that the floods for all stations would follow the same pattern. An extreme-values plotting paper has been designed to straighten out the plotting of this particular skewed distribution, just as arithmetic-probability paper straightens out a normal distribution, and as log-probability paper straightens out a log-normal distribution.

The Gumbel theory cannot be expected to provide an exact distribution for flood peaks for the following reasons:

1. It is assumed that the treatment derived for daily discharges can also be applied to peaks.
2. The daily discharges are not independent events.
3. The 365 daily discharges in a year do not constitute a large number as predicated by the theory.
4. An assumption underlying the extreme-value theory is that all the events are part of the same statistical population. The annual peaks under consideration may not be of the same population. Some peaks are caused by ordinary seasonal rains, some by snow-melt, others by hurricanes. There are different physical factors influencing each type. Therefore, even though all other considerations are in accord with the assumptions, not all of the peaks at one station would necessarily fit the extreme-value theory.

The most recent work by Gumbel (1958, p. 236, 272) describes three basic "asymptotic distributions of extremes," of which the first two, and possibly the third, may be used for floods, depending on the distribution of the data in the sample. The free choice between the three, on the basis of best fit, emphasizes that this statistical method is merely an empirical process of curve fitting. In spite of this deficiency, the Gumbel method of fitting flood data is useful because in some ranges, particularly the low range, the frequency curve tends toward a straight line. An index flood chosen in that range is well defined.

Another theoretical distribution in fairly wide use at present is the log-normal distribution. Chow (1954) has attempted to show justification for this distribution as applied to hydrologic data by use of the central-limit theorem. According to this theorem, the distribution of the sum of a number of independent variables tends toward the normal distribution as the number of the variables becomes large. In flood-frequency analysis, these variables correspond to the logarithms of the factors influencing flood peaks. Part of the requirements for an approximation to the normal distribution is that a large number of independent variables should be involved, each having only a small effect on the result. This condition is not met by the variables related to flood peaks. One of the findings of a recent investigation by Benson (1962) relating flood peaks to hydrologic factors is that only a small number of independent variables can be related to flood peaks. Of the variables found significant, only two or three which are not entirely independent, have by far the most effect. According to Hald (1952), "the central-limit theorem is not of much value for predicting which processes can be expected to generate a normal distribution." In the absence of a theoretical justification, use of the

log-normal distribution remains an empirical method of curve fitting.

It is equally difficult to justify the use of the log-normal distribution for flood peaks empirically. Dixon and Massey (1957, p. 321) say:

In fact, no method which uses the normal distribution can be relied on to estimate extreme percentage points because such estimates depend critically on the assumption of normality. In most experimental research, it is possible to find simple transformations which make the variate essentially normal in the region of the mean, but to make it normal in the tails is quite another matter. Nothing short of an extensive exploration of the distribution, involving perhaps thousands of observations, will suffice in that case.

Obviously the definition of the extreme percentage points is the principal objective of flood-frequency analysis.

In the literature of flood-frequency analysis, where mathematical methods of curve fitting have been used, it has been rediscovered many times that the inclusion or exclusion of a single large flood will lead to considerably different frequency relations. For example, the large flood may occur within a 40-year period but may actually represent a 1,000-year flood. In statistical terms, the large flood is not homogeneous with the rest of the sample. The homogeneity of the sample may be determined by statistical tests, but such tests depend on an assumed normal distribution, and are therefore of limited use. However, both hydrologic considerations and engineering judgment can be used to draw conclusions regarding homogeneity faster and with equal validity. Engineering judgment may make use of historical data at or near the site in question to estimate the probability of the large flood. The use of sound judgment and hydrologic knowledge is demonstrably preferable to a system that relies solely on mechanical methods.

#### GRAPHICAL CURVES

When a set of data is fitted by means of a theoretical curve, more attention is given to the form of the fitted curve than to the actual data. Such a procedure may obscure the relationships that are sought. The graphical procedure, which is much simpler than the analytical, may provide a better indication of the underlying relations.

It was decided that for the studies to follow frequency curves at individual stations would be drawn graphically. The recommendation of the Committee on Floods of the Boston Society of Civil Engineers (1942) on this matter were similar, although perhaps for different reasons. They stated that "Simple plotting with a curve fitted to the points as near as possible without any attempt to extend it much beyond the period of record would seem to be sufficiently accurate for almost any flood frequency computation."

In the graphical method, frequency curves are drawn by eye to average the plotted points. Straight lines are not drawn, regardless of the type of plotting paper used, unless so indicated by the data. There is always the possibility that the highest flood or floods within a short period of record may have a larger or smaller recurrence interval than actually computed. For this reason the curve is not drawn through or near the highest flood unless it follows the trend of the lower points.

If many points in the upper range define a break in the frequency relation, the curve should pass through them. A few points, say 3 or 4, at the upper end that are not in line with the lower points pose a problem. Historical information may help solve the problem by providing more realistic recurrence intervals for these upper points. Examination of other frequency curves for the same locality may also provide a clue as to whether the upper points represent a regional tendency. Comparison with other frequency curves is simplified if a logarithmic discharge scale is used.

#### TYPES OF FLOOD DATA ANALYZED

In the past, various types of data representing flood peaks have been analyzed. Maximum daily discharges have frequently been used, probably because such data are readily available in the U.S. Geological Survey Water-Supply Papers. Daily, weekly, or monthly peak flow rates are actually flood volumes, and a study of their frequency is useful when storage is under consideration. Maximum instantaneous rates of flow are the direct cause of most flood damage, and most flood-frequency studies are made by use of maximum-discharge data.

It is possible either to use the maximum instantaneous flood peak in each year (the annual-flood series) or to consider all flood peaks above some selected base (the partial-duration series). The present tendency is to use the annual-flood series, although some engineers still prefer to use the partial-duration series. At first glance, it would seem that the series that includes the larger number of events would be preferable. Further consideration shows that the additional data increase the definition in only the lower part of the frequency range, a range wherein there are the most data and the least interest. Use of the partial-duration series always involves the necessity for setting arbitrary rules to determine whether adjacent rises should be considered as one or more rises. It has been shown (Langbein, 1949) that the two methods give virtually identical results above a frequency of about 10 years. For lower values, it is possible to convert from one series to the other. For these reasons, and because the annual-flood series is

much simpler to compile and to analyze, it alone is used in flood-frequency studies currently made by the Survey.

### PLOTTING POSITION

The analysis of a series of flood data is begun by arranging the peaks under consideration, such as all the annual peaks at a gaging station, in order of magnitude, with the highest peak as 1. Some measure of frequency must be computed so that a plotting position is obtained for each flood. In traditional statistics, the plotting position would be the probability of a flood as high as, or higher than, each. In flood-frequency analysis, it is customary to use the reciprocal of the probability in terms of years, which is called a return period or recurrence interval. The definition of a continuous distribution by means of a small set of values, as in an array of flood events, presents some difficulties. There are differences in the way that the probabilities (recurrence intervals) may be computed, particularly for the extreme events, in which there is most interest. Some of the methods of computing plotting position are given below.

1. The California method is the simplest form of computation, in which the recurrence interval,  $T_r$ , equals  $n/m$ , where  $n$  is the number of years of record and  $m$  is the rank with the highest peak as 1. There is a minor objection to this method when considering the probability of the event, which is the reciprocal of the recurrence interval. The probability of the lowest flood occurring is computed as 1, which means that the occurrence of any flood lower than this is impossible. The lowest flood could not be plotted on probability paper.
2. The Hazen method computes the return period,  $T_p$ , as  $2n/(2m-1)$ . This formula results in a recurrence interval of  $2n$  for the highest flood of record. This constitutes an artificial lengthening of the period of record. In terms of probability, it signifies that a flood as large as the highest within a 10-year period of record is believed to have only one chance in 20 of occurring within any year, although it is known to have occurred in 1 of 10 years.
3. The theoretical Gumbel method of computing the plotting position is based on the assumption that the observed  $m$ th value is the most probable, or modal, value of this rank of flood. The return period is therefore skewed toward the mode of the theoretical distribution, and it is difficult to compute. Other methods that are equally valid but simpler are recommended.
4. The method of Beard (1943) assumes that the  $m$ th value is the median value for all floods of rank  $m$  within the same period of time. This method results in a recurrence interval of approxi-



mately  $1.44n$  for the highest flood in  $n$  years. This means that we are willing to call the top flood in 100 years a 144-year flood, because a 144-year flood has an equal chance of being exceeded or not exceeded in any 100-year period. Results from this method are intermediate between those from the California method and those from the Hazen method. In terms of probability, the top flood in a 10-year record would be assigned one chance in 15 of occurring within any one year, although experience has shown that it occurred in one out of 10 years. The justification for this reasoning is that, given a large number of 10-year periods, the top flood would be a 15-year flood 50 percent of the time.

5. The formula for recurrence interval currently used by the Survey is

$$T = \frac{n+1}{m}.$$

This formula was first proposed by Kimball (1946). It is almost identical with the California method, but lacks the theoretical deficiencies of the latter method. It gives virtually the same results as Gumbel's recurrence intervals, which are computed by use of the theory of extreme values, and is much simpler to use. It has been adopted by Gumbel and by many others. See Thomas (1948), Gumbel (1954), Chow (1953), Dalrymple (1960, p. 49), and Velz (1950), for discussions and various derivations of the same formula.

This formula has a theoretical basis in that the reciprocal,  $m/(n+1)$ , is the average of the probabilities of all floods with rank  $m$  in a series of  $n$ -year periods. The recurrence interval  $(n+1)/m$  is conservative in that it is equal to  $n+1$  for the highest flood in  $n$  years, thus conforming closely to the period of record. When used in reciprocal form in the annual flood series, it represents the mean probability of occurrence for an annual flood of that size or larger in any one year. Thus the 50-year flood so computed has one chance in 50 or a 0.02-average chance of occurring as an annual flood in any year.

6. When historical flood data are incorporated with a period of recent record, the plotting positions for the historical floods are computed from the longer period of years. For example, consider a 40-year record containing an extraordinary flood that is known to be the highest in 300 years. Its plotting position, by use of the formula  $(n+1)/m$ , is computed as 301 years. The second highest flood within the 40-year period would then be computed as  $20\frac{1}{2}$  years.

The concept of the probability of occurrence within a single year is probably the simplest to understand and the least subject to misinterpretation. The term "50-year flood" may be misunderstood by those unfamiliar with probability concepts, and hence they may be surprised at the not-uncommon occurrence of two "50-year floods" within successive years. It would be more apparent that this flood has a small though equal chance of occurring within any year, if the terms "0.02-probability flood" or "2-percent chance flood" were used instead.

The probability of occurrence for a given flood within any year is not the only possible basis for design. For example, the designer may wish to consider the flood magnitude that has a given probability of being exceeded one or more times within some definite period of time, such as the first 5 years after construction or the estimated life of the structure. This is a compound rather than a simple probability, involving consideration of the distribution of the return period. The subject has been treated by Thomas (1948), Alexander (1957), and others. Linsley, Kohler, and Paulhus (1958, p. 248, 258) present a clear exposition of "the significance of the  $N$ -yr event" and the "selection of the design frequency," as related to the distribution of the return period. Linsley, Kohler, and Paulhus (1958) also present tables (11-2 and 11-6) that may be used in the selection of a design flood.

In flood-frequency studies currently made by the Survey, the recurrence interval used for a given flood magnitude is the reciprocal of the simple probability that the flood magnitude may be exceeded within any year. Translation from the simple to the compound probability, if desired, may be made by the appropriate applications of probability theory.

#### PLOTTING PAPER

After computing the plotting positions in terms of recurrence interval or probability, the flood magnitudes are plotted against these positions on some type of plotting graph. The plotting paper serves as a smoothing graph; thus any paper on which the frequency relation tends toward a straight line or a smooth curve is satisfactory. Log-log paper may be used; if so, the resulting frequency curve will bend sharply (concave downward) in the lower range but may be satisfactory elsewhere. Semilog paper may also be used, and is most desirable when the partial-duration series is used. See Langbein (1949) and Chow (1950). Graph paper specifically designed to fit the theory of extreme values is satisfactory. Recurrence interval, in years, is used as the horizontal scale with plotting papers already mentioned. If probabilities are used (they are the reciprocal of the

recurrence interval), the data may be plotted on arithmetic- or logarithmic-probability paper, but the latter is preferable.

The vertical or discharge scale may be either linear or logarithmic. As previously mentioned, comparison of frequency curves is facilitated if logarithmic discharge scales are used, because equal vertical distances on each graph represent equal percentage changes in discharge. For this reason, either logarithmic-probability or logarithmic Gumbel graph paper is desirable.

## REGIONAL FLOOD-FREQUENCY ANALYSIS

### SIGNIFICANCE OF A SINGLE STATION RECORD

The foregoing discussion on methods of frequency analysis applies to the analysis of the records at an individual site. The results of such a study represent an exact description of what has happened at the site in the past, for some definite period of time.

However, frequency analysis of peak discharges is most commonly used to estimate the basic or long-time distribution. This distribution can then be considered as the best representation of future expectations. In the use of past records to predict the future, it must be assumed that there has been no change in the nature of the factors influencing the size of flood peaks. If conditions are known to have changed, the changes must be considered if possible, otherwise the changing conditions will be a source of error in the final results. The past record is considered as a sample of the total population of peak discharge, consisting of both past and future floods. When used for predicting the future, the sample must be considered as only an approximation, because it may vary from the group as a whole.

In general, flood records represent relatively short samples. For this reason, the record at any individual station may depart considerably from a true representation of the overall long-time flood-frequency relation.

### SYNTHETIC 1,000-YEAR FLOOD STUDY

A study was made by Benson (1952) to demonstrate the variability inherent in short records. This study was begun by use of a theoretical 1,000 years of record with annual peaks so distributed as to define exactly a straight-line graph on extreme-values graph paper. Individual peaks were then drawn at random to form groups of samples equivalent to 10-year, 25-year, 50-year, and 100-year periods of record. Benson's study was incorporated into a paper by Dalrymple (1960, p. 51-74). Figures 20 to 23 of Dalrymple's paper show graphically the variation in frequency curves based on the short-term samples. (The curves were extended beyond the period of record, based on their

lower parts, in order to test the reliability of such extensions.) Even the 100-year records show more variation than expected. The significance is that an actual gaging-station record of 25 years, for example, could by chance alone take the position of any of the curves shown in figure 21 of Dalrymple's paper. The range in variation away from the known base value at any recurrence interval can be estimated from these curves. Even the mean annual flood (recurrence interval, 2.33 years, see p. A-17) can vary significantly from the overall value. The study shows that 12 years of record are required to define the mean annual flood within 25 percent (with expectation of correct results 95 percent of the time). The quantitative results are, of course, dependent on the initial assumptions, but the study indicates the large variation possible from chance alone in single-station frequency curves.

#### INDEX-FLOOD METHOD

Because of the unreliability of the single-station curve, other methods have been developed (Dalrymple, 1960). These involve a study of frequency relations on a regional basis. Such studies are used in an attempt to reduce the large sample errors inherent in individual records and to arrive at the basic long-term relations. The resulting general relations are then applicable to all points within the region studied, at sites previously gaged or ungaged. Many regional frequency reports have been prepared for individual States. (See p. A-18, 19.)

The index-flood method is currently used by the U.S. Geological Survey for regional frequency analysis. There are two major parts to such an analysis. The first is the development of basic dimensionless frequency curves representing the ratio of the flood of any frequency to an index flood (the mean annual flood). The second is the development of relations between hydrologic characteristics of drainage areas and the mean annual flood, for predicting the mean annual flood at any point within the region. Combining the mean annual flood with the regional frequency curve, which is in terms of the mean annual flood, provides a frequency curve for any location.

#### BASE-FREQUENCY CURVES

In fairly large regions that are homogeneous with respect to flood-producing characteristics, individual streams whose drainage areas vary greatly in size have frequency curves of approximately equal slope or steepness, if discharge is expressed as a ratio of the mean. The peak floods at each gaging station are divided by an index flood (the mean annual flood at the station is commonly used), thus reducing them to dimensionless ratios. The individual curves that are

plotted by use of the dimensionless flood ratios may be superimposed and will nearly coincide with each other.

These curves will all pass through the frequency of 2.33 years at the ratio of 1.0, but will have somewhat different slopes. The variation of these slopes may be tested to see whether the spread of the slopes could occur by chance alone among samples from the same population (this is the homogeneity test, which makes use of the ratio of the 10-year flood to the 2.33-year flood as the slope). If so, the assumed homogeneity of the region where all the stations are located is considered to be satisfactory. Within such a homogeneous region, a general representation of the flood-frequency relation may be obtained by combining all dimensionless curves. The resulting average frequency curve is then applicable throughout the region, and is called the base or regional frequency curve.

If the spread in the slopes of the dimensionless curves is greater than may be attributed to chance alone, the region is subdivided so as to produce two or more flood-frequency regions. Statewide studies usually indicate two to three such regions within a State; however, these will tie in with like regions in adjoining States, so that the area of a particular flood-frequency region may be large.

The combining of individual curves to obtain the base frequency curve is done by first determining the median of flood ratios for each station for the same recurrence intervals. These median values are then plotted to define the base frequency curve.

#### MEAN ANNUAL FLOOD

The mean annual flood, as used in the index-flood method, is the value of the graphical frequency curve at a recurrence interval of 2.33 years. Generally, an adjustment is made so that the mean annual floods are related to the same base period of time. The 2.33-year flood is used because it may be demonstrated by use of the theory of extreme values that the mean of many annual flood peaks has a magnitude equivalent to the flood of a 2.33-year recurrence interval. In other theoretical approaches, such as the Foster types I and III and the log-normal distribution, the recurrence interval of the mean is nearly 2.33 years, provided the coefficient of skew is not much different from 1.0. The actual arithmetic mean is not used, because of the possible bias due to the occurrence of rare floods within the period of record. The median rather than the mean annual flood may be used with equal justification; however, engineers have traditionally used the mean annual flood.

The magnitude of the mean annual flood may be affected by many factors, which may be classed as either topographic or meteorologic.

The problem is to determine what mean annual flood may be expected to recur within a drainage basin of given physical characteristics, located in a region subject to given meteorologic conditions. The answer is obtained by correlating the known mean annual floods for gaged drainage areas within a region with the known basin and climatic characteristics. For simplicity, the combined topographic and meteorologic characteristics are called hydrologic characteristics in this report.

Where the study region is inadequately mapped or where the terrain is uniform over large areas, a method used successfully in frequency studies has been to divide the region into various parts called hydrologic areas. Within each area, a separate curve of mean annual flood is correlated with the drainage area and perhaps some other significant factor. Within each area, general factors, such as rainfall and geology, probably have the same overall effect. This method has been found satisfactory in many of the statewide frequency studies. However, the location of boundaries between the areas is at times vague. Drawing boundaries produces an unknown number of lost degrees of freedom, so that computation of confidence limits becomes uncertain.

Correlation of the mean annual flood with other basin or meteorologic factors may be studied by use of either mathematical or graphical correlation techniques. Techniques of multiple correlation are given in most textbooks on statistics. Both mathematical and graphical procedures are described by Ezekiel and Fox (1959).

Results of correlation studies can be shown entirely by graphs, in which a set of curves shows all the necessary relationships. This method has been used in most of the regional studies made so far. It furnishes the simplest means of applying the results to ungaged areas. Results may be shown by nomographs, as was done by Bodhaine and Thomas (1960). Results may also be shown by equations. Flood formulas have been developed for Massachusetts by Kinnison and Colby (1945), for Connecticut by Bigwood and Thomas (1955), and for eastern Montana by Berwick (1958).

#### **FACTORS RELATED TO MEAN ANNUAL FLOOD**

States in which regional frequency studies have been made and hydrologic factors that have significant correlation with the mean annual flood are given below. The factors are listed approximately in order of importance.

Alabama (Peirce, 1954) : Drainage area, hydrologic areas (some evident relation of peak discharges to shape of basin and types of soil).

Connecticut (Bigwood and Thomas, 1955) : Drainage area, slope of principal streams, degree of urbanization, channel storage.

- Delaware River Basin (Tice, 1958) : Drainage area, hydrologic areas, storage area.
- Florida (Pride, 1958) : Drainage area, hydrologic areas, storage area.
- Georgia (Carter, 1951) : Drainage area, hydrologic areas.
- Illinois (Mitchell, 1954) : Drainage area, hydrologic areas, basin lag.
- Iowa (Schwob, 1953) : Drainage area, hydrologic areas (some evident relation of areas to types of soil).
- Kentucky (McCabe, 1958) : Drainage area, hydrologic areas.
- Louisiana (Cragwall, 1952) : Drainage area, hydrologic areas.
- Massachusetts (Kinnison and Colby, 1945) : Drainage area, mean distance to outlet, mean altitude above gage, storage area.
- Minnesota (Prior, 1949) : Drainage area, hydrologic areas, annual runoff, storage area.
- Missouri (Searcy, 1955) : Drainage area, hydrologic areas.
- Eastern Montana (Berwick, 1958) : Drainage area, mean altitude.
- Nebraska (Furness, 1955) : Drainage area, geographical factor.
- Western New Mexico and eastern Arizona (Kennon, 1954) : Drainage area, hydrologic areas.
- North Carolina (Riggs, 1955) : Drainage area, hydrologic areas.
- North and South Dakota (McCabe and Crosby, 1959) : Drainage areas.
- Ohio (Cross and Webber, 1959) : Drainage area, main-channel slope, types of soil.
- Pennsylvania (Youghiogheny and Kiskiminetas basins, U.S. Geol. Survey, 1952) : Drainage area.
- Washington (Bodhaine and Thomas, 1960) : Drainage area, hydrologic areas (geographical factors), mean annual runoff, area of lakes.
- Wisconsin (Ericson, 1961) : Drainage area, main-channel slope, geographic factor, storage area.

#### SPECIAL TREATMENT FOR LARGE STREAMS

Many large streams traverse more than one flood-frequency region or hydrologic area. The occurrence of floods for these streams is commonly not the same as that of the small tributary streams. For this reason, the large streams cannot be combined with the tributary streams to define general relationships and may, therefore, require separate treatment to define both the base-frequency curve and the mean-annual-flood relation.

For large streams, there are generally a few gaging-station records that can be used to define the flood-frequency relation. A single composite base-frequency curve for the stream, such as for the Mississippi River (Searcy, 1955, fig. 12) may be sufficient. For other streams the frequency relation may be changing progressively downstream, and may require treatment as given for the Red River in Louisiana (Cragwall, 1952, fig. 22).

The mean annual floods for large streams may require definition by curves based on records for stations only along these streams. The curves serve to smooth the individual station records and are a means

of interpolation between the stations on the basis of either drainage area or river mileage.

### **MULTIPLE-CORRELATION METHODS**

An alternative to the index-flood method of regional frequency analysis may be used, if there are sufficient data covering both the flood peaks of high recurrence intervals and the hydrologic characteristics of the drainage basins. Studies by Benson (1962) have shown that it is possible to use statistical multiple-correlation techniques that will relate the floods at any defined level of recurrence interval to the hydrologic characteristics. This procedure eliminates some uncertainties of the index-flood method and makes it unnecessary to assume any fixed distributions for peak discharges.

It is probable that different factors may be acting at the separate flood levels or that the same factors may have varying effects at different levels. The multiple-correlation method, when used independently at specific flood levels, allows complete flexibility and does not require making assumptions about the relation between floods of different recurrence intervals.

Multiple correlation means simply the relation between more than two variables. A relation between flood peaks and drainage area would be a simple correlation, whereas a relation between the flood peak, drainage area, and precipitation would be a multiple correlation. Standard procedures for developing such relations are described in most textbooks on statistics. The procedures may be either graphical or analytical. Both procedures are explained by Ezekiel and Fox (1959).

Multiple-correlation methods have been used by some to relate hydrologic characteristics to parameters of an assumed distribution—mean, standard deviation, and skew. Another procedure is to use magnitudes of floods for selected percentiles, rather than to use statistical parameters as the dependent variables. This constitutes a non-parametric rather than a parametric statistical approach. These two approaches have been discussed previously in this report.

### **EVALUATION OF THE INDEX-FLOOD METHOD**

Some deficiencies have recently been recognized in the use of the index-flood method for developing regional flood-frequency relations.

1. The flood ratios for comparable streams may differ, because of large differences in the index flood. For example, the 50-year flood ratio for two areas of the same size may vary two- or three-fold, because, although the 50-year floods are of the same size, the mean annual floods are not. This points up a deficiency in the index-



flood method. If the index flood is not typical and is used as obtained from a short period of record, the remainder of the frequency curve, if based on the index, may be faulty. This deficiency has been recognized and described in previous reports, although no solution could be offered.

2. The homogeneity test is used to determine whether the differences in slopes of frequency curves are greater than may be attributed to chance alone. This test uses the ratio of the 10-year flood to the mean annual flood as the slope. The test cannot practicably be applied at a level much higher than that of the 10-year flood because many individual records are too short to adequately define the frequency curve at higher levels. It has been found in some studies that although homogeneity is apparently established at the 10-year level, the individual curves show wide and sometimes systematic differences at higher levels.
3. In the use of the index-flood method, it has been accepted that within a flood-frequency region, frequency curves may be combined for all sizes of drainage areas, excluding only the largest. Although the variation in the slope of the frequency curve with drainage area had been investigated at the time of each study, it was studied at the 10-year point, where the effect is small. The error of neglecting this drainage-area effect has been reduced by giving separate and special treatment to large streams, as described previously. Recent studies for which ratios of less frequent floods were used have shown in all regions where such data are available that the ratios of any specified flood to the mean annual flood will vary inversely with the drainage area. In general, the larger the drainage area, the flatter the frequency curve. The effect of drainage area is relatively greater for floods of higher recurrence intervals.

Investigations by Benson (1962) have shown that the flood ratios vary not only with drainage area but with main-channel slope and climatic factors as well. In his studies, using multiple-correlation methods, the resulting flood formulas were in the form :

$$Q_T = aA^b S^c St^d I^e t^f O^g,$$

in which the  $T$ -year flood varies with the drainage-area,  $A$ ; the main-channel slope,  $S$ ; the surface storage,  $St$ ; the rainfall intensity,  $I$ ; the mean number of degrees below freezing,  $t$ ; and an orographic factor,  $O$ .

If the formula for any peak is divided by the formula for the 2.33-year peak, the resulting flood ratio is in terms of the drainage area, main-channel slope, storage, rainfall intensity, degrees below

freezing, and orographic factor. For example by use of the results of Benson's study (1962).

$$\frac{Q_{25}}{Q_{2.33}} = \frac{2.410A^{0.972}S^{0.468}St^{-0.227}I^{0.511}t^{0.415}O^{1.142}}{3.316A^{0.998}S^{0.357}St^{-0.279}I^0t^{0.347}O^{0.845}} \\ = 0.727A^{-0.026}S^{0.111}St^{0.052}I^{0.511}t^{0.068}O^{0.297}.$$

Statistical tests show that the exponents in the last equation and in similar equations for other flood ratios are significantly different (at the 5-percent level) from zero through part or all the range of recurrence interval, for all variables except storage.

The results of Benson's (1962) study also show the variation in the flood ratio if flood peaks are related only to drainage area. For example,

$$\frac{Q_{25}}{Q_{2.33}} = \frac{196.9A^{0.740}}{44.4A^{0.954}} = 4.43A^{-0.114}.$$

The exponent of  $A$ ,  $-0.114$ , can be shown to be highly significant statistically (at the 1-percent level), as are all similar exponents for ratios for higher recurrence intervals. Exponents for ratios for recurrence intervals of 10 years or less cannot be shown to be statistically significant; however, they follow the trend shown by the higher floods and therefore undoubtedly show that the flood ratio varies with drainage area throughout the range of recurrence interval.

It may be concluded that the flood ratio is not a constant, and that it varies with five of the independent variables. For any one station, the ratio and, hence, the slope of the frequency curve, depends upon the specific values of the five variables at that station.

Three of the variables—area, slope, and rainfall—are interrelated. In general, as drainage area increases, slope and rainfall intensity decrease. Examination of the second formula on p. A-22 shows that increase in area, decrease in slope, and decrease in intensity all tend to decrease the flood ratio. This explains why the frequency curves of large streams are generally flatter than those of small streams, if the peak discharge is shown as the ratio to mean discharge.

In the index-flood method, the base-frequency curve is defined by the medians of the flood ratios at individual stations. An equivalent for the 10-year flood ratio for New England may be obtained by determining the median values of all six variables and by substituting them in an equation for the 10-year flood ratio. It then becomes apparent that the median flood ratio at any level depends upon the range (or median values) of each of the variables for the stations that are combined to define the base-frequency curve.

For example, consider an area, such as New England, throughout which the same basic flood-frequency relations apply as expressed in the general formulas already developed. Consider what might happen if the index-flood method is used. If within a large part of the area, the gaging stations were on predominantly large drainage areas, the 10-year (and other) flood ratios would be predominantly small. This part of the area would then probably be considered as a separate frequency region, for which a flatter base-frequency curve would be determined than for the remainder of the area. Although this curve would represent the median position of the station curves from which it had been derived, it would not represent the true median position of all the ungaged sites to which it might be applied. The shape of the base-frequency curve based on 1000-square-mile areas would be different from one based on 500-square-mile areas. Yet each curve would be taken to represent the basic flood-frequency relation that would be applicable to any drainage basin.

In spite of the deficiencies previously mentioned, the regional frequency relations developed by using the index-flood method have led to the most useful results so far available. The method will remain useful where inadequate mapping or lack of meteorologic data make it unfeasible to compute the significant hydrologic factors. It will also remain the most practicable solution where there are too few long-term records of flood peaks to allow adequate multiple-correlation studies of the peaks of higher recurrence intervals. The underlying relations found in studies by Benson (1962) may be useful in refining the index-flood method. For example, adjustments may be made in the flood ratios for the size of drainage area, and adjustments may be made for other significant variables, if they can be evaluated.

### DESIGN FREQUENCY RELATIONS AT GAGED SITES

There is no question but that generalized flood-frequency relations are the best means for determining the flood-frequency curve at an ungaged site. But what of a gaged site? Are the general relationships developed more or less reliable for design purposes than is the actual record collected at the site? There are two possible reasons for variation from the normal condition. One is simply chance variation. Even a fairly long record may by chance vary considerably from the normal or median condition (see p. A-16). The other reason for variation may be some factor or factors that are not accounted for in the flood-frequency relations, but that may have a large effect on the magnitude of flood peaks at the design site. Benson's study (1962) has an easily recognized example of this. Such a condition may exist at the gaging station on Otter Creek at Middlebury, Vt.

The actual values for flood peaks there are only about one-third of those computed from the generalized relations. The report of the Committee on Floods of the Boston Society of Civil Engineers (1930, p. 375) says:

An unusual case of channel pondage also occurs upon Otter Creek above Middlebury in Vermont. The dam at Middlebury is at about El. 336, while for many miles upstream the valley is wide and flat so that it overflows in high water every year. The water level reached in the 1927 flood above Middlebury was about El. 352, corresponding to a pondage of about 4.5 billion cubic feet or 7.0 million cubic feet per square mile of drainage area. As a result in this flood the maximum flow at and below Middlebury was only about 23 second feet per square mile compared with 100 or more on many other rivers of this size at that time.

Channel storage was recognized as a factor that might significantly affect flood peaks, but it was not feasible to make a comprehensive study of it in Benson's (1962) investigation. Where such a factor has extreme variation from the normal condition, it may have a considerable effect on the magnitude of peaks, as apparently was the case here.

If the variation from the general relation at an individual station were known to be due to chance alone, it would be better to use the general relation. If the variation were large enough and if some hydrologic factor consistent with the direction of variation were known or suspected to be operative, the individual station frequency curve might be preferable for design purposes. Even though no reasons are apparent, a large variation from the general relation might be considered as sufficient reason for giving consideration to the actual record, despite the very large variations that may occur through chance alone.

There is some point in every engineering investigation where engineering judgment must take over. It appears that this is the point here. For design at or near a gaging-station site, both the actual and the theoretical curves should be compared. The decision of which to use should be made by consideration of the length of record, the amount and direction of the variation between the actual and computed curves, the existence of otherwise unaccounted-for variables consistent with the variation, the risk involved, and the size of the design flood. If the design is for a high flood, the probability of chance variation at a single station is large, and the general relation is likely to be better. If the two curves are consistently far apart throughout the range, the curve based on the station record may have more validity. If the record is short, variation due to chance is likely to be large in the curve for the station and the general relation likely to be better.

## PREDICTIVE VALUE OF THE FLOOD-FREQUENCY RELATION

In the study of flood-frequency relations, there are many unresolved problems that require consideration. The variation of flood occurrences with time may be either random or systematic. The variation is systematic if long-term climatic changes are effective within the period of record or if natural or man-made changes occur that affect the characteristics of the drainage basin. The period of record is generally too short to establish with certainty whether the variations are truly random and our knowledge of the effect of man-made changes or the trend of climate is insufficient to establish that the variations are systematic.

In generalizing flood-frequency relations over a region, uncertainties exist because of the lack of independence in the flood data. Most floods occur under conditions that affect wide areas. The same storm may produce peak discharges at many gaging stations within a region. Though peak discharges may not occur on the same day, the period and magnitude of thundershowers of small areal extent are related for basins within wide regions. Correlation may be rather high between annual peak discharges of stations affected by thunderstorms and having but few annual peaks of the same date.

In a research study of regional flood-frequency relations in New England (Benson, 1962), 164 basins were used in the analysis. Five or six major storms were responsible for the 3 or 4 highest flood peaks at most of the 164 stations. Flood peaks at each station correlated to some degree with peaks at most other stations within New England. There were 13,366 possible combinations of the 164 stations. By random selection, a sample of 200 pairs was used to obtain a distribution of distances between stations and the median distance—94 miles. A further sampling, stratified for distance, provided 54 pairs of stations. The correlation coefficients between the stations in each of the 54 pairs were computed and plotted against the distances between the stations. A correlation coefficient of 0.26 corresponded to the median distance of 94 miles, and was taken to be the general overall correlation coefficient between the 164 stations.

Insofar as the mean of peak discharges is concerned, such a correlation coefficient has the effect that the 164 stations are the equivalent of only about 3 independent stations. What is the effect of this dependence upon the results obtained in the regional flood-frequency analysis?

For the purpose of analyzing this effect, a model is set up with even less independence than is usually found.

1. Assume 100 stations within a region of widely varying physical characteristics.
2. Suppose that in each year there is only one rainstorm of any consequence in this area, having a uniform intensity over the entire region and lasting exactly 24 hours.
3. Suppose that this storm produces a peak discharge of the same frequency, for example, a 10-year peak at each station. This is not possible, but the assumption is made at this time as a starting point for speculation.
4. Suppose that there is a long-term or basic distribution of floods within the region.
5. Assume a 50-year record over this area with a representative distribution of floods; that is, both the rainfalls and the resulting peak discharges would define frequency curves with little or no scatter of individual points.

There is only one experience of the so-called 50-year flood level, in general, over the area. For this reason the 50-year flood, in general, or the 50-year peak discharge at any station is not known with any confidence. The so-called 50-year flood level may be actually (if indeed there is a long-time "actual" flood level) the 80-year level or the 25-year level.

Suppose that the 50-year peak discharges (dependent variables) at the 100 stations are correlated with the physical characteristics of the basins (independent variables), and a multiple-correlation equation is developed. What then is the significance of such an equation and what confidence may be placed in it? The equation is based on 100 sets of data, one set at each station. The dependent variables are all known to be at the same flood level, but what that level may be is not known with any confidence. Yet the relation between peak discharges and physical characteristics for that flood level has been determined with some exactness, based on 100 sets of data and nearly 100 degrees of freedom. A fairly dependable relation has been derived between known quantities, the physical characteristics, and peak discharges whose magnitudes are known but whose true recurrence interval is not known. The uncertainty of the recurrence interval does not diminish the confidence limits of the relation with the physical characteristics; on the other hand, the reliability of the regression equation does not add to the knowledge of the true recurrence interval of the discharges.

Of what use then is such an equation? Because of the shortness of the records, the 50-year flood is only nominal. Yet this figure must serve as the basis for the problems of engineering design. It is al-

most necessary to assume that the available period of record has defined the true distribution of floods, though this is not true. For the design flood, whether 50-year flood or not, the regression equation indicates how the so-called 50-year peak discharges vary for each drainage basin. For design purposes, a 50-year record with a given distribution of floods has been postulated to serve as a basis for assigning values to the 50-year, 25-year, 10-year, and other peaks. Regression equations relating the  $T$ -year peak discharge,  $Q_T$ , to the hydrologic characteristics may be determined for each of these flood levels. These equations have the form

$$Q_T = aB^bC^cD^d \dots,$$

where  $B, C, D, \dots$  are the independent variables characterizing the hydrologic conditions and  $a, b, c, d, \dots$  are constants of the regression equation.

Suppose now that there is a second 50-year period during which the annual peaks average twice the magnitude of the peaks in the first period. From data for this period, a second set of values could be assigned to the 50-, 25-, and 10-year peaks and regression equations relating them to the hydrologic characteristics could be developed. The resulting equations could then be expected to have  $a$  values approximately twice those of the first period; the remaining regression constants  $b, c, d, \dots$  would be the same as those defined for the first period.

What is the relation of the actual condition to the model? In regions where snowmelt peaks are common, there may be only one peak each year; but in other regions there are generally more than one. Annual peaks at all stations will not be caused by the same storm. The annual peak discharges of each year do not have the same recurrence intervals at all stations; instead, the 25-year peaks, for example, are determined from the frequency curves at each station. Yet the list of 25-year peaks that are used to set up a regression equation defining the 25-year peak in terms of basin and climatic characteristics is based on the experience during a given period with a given average level of flood occurrences. It seems clear that the  $a$  constant in the regression equation is dependent on the general level of the flood activity and that the other regression constants are based on the relations between that general level of flood activity and the hydrologic characteristics. The equivalent of three independent stations have provided three independent measures of the 50-year flood level rather than one, as in the model. The 50-year flood is less uncertain but still an uncertain quantity.

The vague concept of the flood level that has been used may be made more concrete by starting from the regression equation for the  $T$ -year peak discharge.

$$Q_T = aB^b C^c D^d \dots$$

Consider a set of 25-year peaks. If the actual peaks can be adjusted to unit or standard conditions, a set of comparable unit discharges for all basins will be obtained, whose average can be called the 25-year flood level. This can be done by assigning a value of 1 to all the physical basin characteristics or climatic variables and then by adjusting the known  $T$ -year discharge for all the variables thus:

$$Q_T \times \frac{1^b}{B^b} \times \frac{1^c}{C^c} \times \frac{1^d}{D^d} \dots,$$

which is equal to

$$Q_T / B^b C^c D^d$$

This is theoretically equal to  $a$  at each station, and differs only by a random error—the average for all stations should in fact equal  $a$ . The value  $a$  can thus be treated as an actual measure of the general or unit flood level for any recurrence interval.

The derivation of the term " $a$ " also makes it evident that  $a$  is directly related to the mean of the  $T$ -year discharges:

$$\log a = \bar{Q}' - b\bar{B}' - c\bar{C}' \dots,$$

where  $Q', B', C', \dots$  are the logarithms of the original variables.

The confidence limits that may be computed for  $a$ , by use of a number of degrees of freedom nearly equal to the total number of stations, is a measure of the variation in  $a$  based on the mean of the  $Q_T$  values selected as the  $T$ -year peaks. Those confidence limits are not, however, the limits for the true  $T$ -year flood level. Such limits must be based on the number of independent determinations of the  $T$ -year flood level.

The equivalent number of independent stations,  $N'$ , according to Alexander (1954), equals  $K/[1 + \bar{r}(K-1)]$ , where  $r$  is the overall or mean correlation coefficient and  $K$  is the number of stations. The number of independent determinations is apt to be quite small, regardless of the total number of station records available. No appreciable improvement in defining the population value is obtained by using 300 instead of 50 stations, if they are all within the same general area.

The conclusion is that floods cannot be predicted within narrow confidence bands if, by predicting floods, is meant predicting the true  $T$ -year peak discharge either at a gaged or ungaged site, from the usual length of records available. For example, for 50 stations—and a mean correlation coefficient of 0.3, the mean flood would be



defined by 3.2 independent events. The confidence limits for a mean flood determined by 3.2 independent events would necessarily be rather wide.

However, all data within the period of record can be used as though they represent the true distribution, and consistent if not absolute standards can be attained for design by differentiating the peak discharges on various basins by means of the physical and climatic characteristics.

### REFERENCES

- Alexander, G. N., 1954, Some aspects of time series in hydrology: Jour. Inst. Engineers [Australia], Sept., p. 196.
- 1957, Flood flow estimation, probability, and the return period: Jour. Inst. Engineers [Australia], Oct.-Nov., p. 263-280.
- Beard, L., 1943, Statistical analysis in hydrology: Am. Soc. Civil Engineers Trans., v. 108, p. 1110-1160.
- Benson, M. A., 1952, Characteristics of frequency curves based on a theoretical 1,000-year record: U.S. Geol. Survey open-file report, 21 p.
- 1962, Factors influencing the occurrence of floods in a humid region of diverse terrain: U.S. Geol. Survey Water-Supply Paper 1580-B (in press).
- Berwick, V. K., 1958, Floods in eastern Montana, magnitude and frequency: U.S. Geol. Survey open-file report, 23 p.
- Bigwood, B. L., and Thomas, M. P., 1955, A flood-flow formula for Connecticut: U.S. Geol. Survey Circ. 365, 16 p.
- Bodhaine, G. L., and Thomas, D. M., 1960, Floods in Washington, magnitude and frequency: U.S. Geol. Survey open-file report, 25 p., 3 pls.
- Boston Society of Civil Engineers, 1930, Report of the Committee on Floods: Boston Soc. Civil Engineers Jour., v. 17, p. 293-466.
- 1942, Report of the Committee on Floods: Boston Soc. Civil Engineers Jour., v. 25, p. 1-160.
- Carter, R. W., 1951, Floods in Georgia, frequency and magnitude: U.S. Geol. Survey Circ. 100, 127 p.
- Chow, V. T., 1950, Discussion of "Annual floods and the partial-duration series", by W. B. Langbein: Am. Geophys. Union Trans., v. 31, p. 939-941.
- 1953, Frequency analysis of hydrologic data with special application to rainfall intensities: Illinois Univ., Bull. Ser. 414, 80 p.
- 1954, The log-probability law and its engineering applications: Am. Soc. Civil Engineers, separate no. 536, 25 p.
- Cragwall, J. S., Jr., 1952, Floods in Louisiana, magnitude and frequency: Louisiana Dept. Highways, 281 p.
- Cross, W. P., and Webber, E. E., 1959, Floods in Ohio, magnitude and frequency: Ohio Dept. Nat. Resources, Div. of Water Bull. 32, 325 p.
- Dalrymple, Tate, 1960, Flood-frequency analyses: U.S. Geol. Survey Water-Supply Paper 1543-A, 80 p.
- Dixon, W. J., and Massey, F. J., Jr., 1957, Introduction to statistical analysis: 2nd ed., New York, McGraw-Hill, 488 p.
- Ericson, D. W., 1961, Floods in Wisconsin, magnitude and frequency: U.S. Geol. Survey open-file report.
- Ezekiel, M., and Fox, K. A., 1959, Methods of correlation and regression analysis: 3rd ed., New York, John Wiley and Sons, 548 p.

- Furness, L. W., 1955, Floods in Nebraska, magnitude and frequency: Nebraska Dept. Roads and Irrig., 103 p.
- Gumbel, E. J., 1954, Statistical theory of extreme values and some practical applications: Natl. Bur. of Standards, Appl. Math. Ser. 33, p. 13-15.
- 1958, Statistics of extremes: New York, Columbia Univ. Press, 375 p.
- Hald, A., 1952, Statistical theory with engineering applications: New York, John Wiley and Sons, p. 190.
- Jarvis, C. S., and others, 1936, Floods in the United States, magnitude and frequency: U.S. Geol. Survey Water-Supply Paper 771, 497 p.
- Johnstone, D., and Cross, W. P., 1949, Elements of applied hydrology: New York, Ronald Press Co., 276 p.
- Kennon, F. W., 1954, Magnitude and frequency of summer floods in western New Mexico and eastern Arizona: U.S. Geol. Survey open-file report, 14 p., 10 figs.
- Kimball, B. F., 1946, Assignment of frequencies to a completely ordered set of sample data: Am. Geophys. Union Trans., v. 27, p. 843-846.
- Kinnison, H. B., and Colby, B. R., 1945, Flood formulas based on drainage-basin characteristics: Am. Soc. Civil Engineers Trans., v. 110, p. 849-904.
- Langbein, W. B., 1949, Annual floods and the partial-duration flood series: Am. Geophys. Union Trans., v. 30, p. 879-881.
- Langbein, W. B., and Hoyt, W. G., 1959, Water facts for the nation's future: New York, Ronald Press Co., 288 p.
- Linsley, R. K., Kohler, M. A., and Paulhus, J. L., 1958, Hydrology for engineers: New York, McGraw-Hill, 340 p.
- McCabe, J. A., 1958, Floods in Kentucky, magnitude and frequency: U.S. Geol. Survey open-file report, 13 p.
- McCabe, J. A., and Crosby, O. A., 1959, Floods in North and South Dakota, frequency and magnitude: U.S. Geol. Survey open-file report, 132 p.
- Mitchell, W. D., 1954, Floods in Illinois, magnitude and frequency: Illinois Dept. Public Works and Buildings, Div. of Waterways, 386 p.
- Peirce, L. B., 1954, Floods in Alabama, magnitude and frequency: U.S. Geol. Survey Circ. 342, 105 p.
- Pride, R. W., 1958, Floods in Florida, magnitude and frequency: U.S. Geol. Survey open-file report, 136 p.
- Prior, C. H., 1949, Magnitude and frequency of floods in Minnesota: Minnesota Dept. Conserv., Div. of Waters Bull. 1, 128 p.
- Riggs, H. C., 1955, Floods in North Carolina, magnitude and frequency: U.S. Geol. Survey open-file report, 59 p.
- Schwob, H. H., 1953, Iowa floods, magnitude and frequency: Iowa Highway Research Board Bull. 1, 171 p.
- Searcy, J. K., 1955, Floods in Missouri, magnitude and frequency: U.S. Geol. Survey Circ. 370, 126 p.
- Thomas, H. A., 1948, Frequency of minor floods: Boston Soc. Civil Engineers Jour., v. 35, p. 428.
- Tice, R. H., 1958, Delaware River Basin flood frequency: U.S. Geol. Survey open-file report, 10 p.
- U.S. Geological Survey, 1952, Floods in Youghiogheny and Kiskiminetas river basins, Pennsylvania and Maryland, frequency and magnitude: U.S. Geol. Survey Circ. 204, 22 p.
- Velz, C. J., 1950, Normal probability paper, pt. 2 of Graphical approach to statistics: Chicago, Ill., Water and Sewage Works, p. 309-314.