Journal of the
HYDRAULICS DIVISION
Proceedings of the American Society of Civil Engineers

FREQUENCY OF NATURAL EVENTS

By H. C. Riggs,[1] M. ASCE

SYNOPSIS

Magnitude-frequency relationships are used frequently in design problems, yet the precise meaning of the relation is not widely understood. This paper begins with the development of a cumulative frequency curve and its statistical interpretation. From the frequency curve, a relation between magnitude, design period in years, and probability of not exceeding that magnitude in the design period is derived. The relation is presented graphically for easy use, and the applicability of the general procedure is shown by sampling from a 1,023-yr period of tree-ring indexes.

INTRODUCTION

An estimate of the frequency of an event, such as a flood or a drought, usually is obtained from a cumulative frequency curve. That curve, based on observed events, and constructed by one of several standard methods, relates the magnitudes of events to mean recurrence intervals or to probabilities. Both the magnitude and the recurrence interval in such a plot are subject to sampling errors. The sampling error of the magnitude can be reduced only by increasing the sample size. A sampling error of the recurrence interval is present because the recurrence interval is not a fixed value but is the mean length of the intervals between events that exceed a given magnitude. The two sampling errors are dependent and a procedure for combining them has not been

Note.—Discussion open until June 1, 1961. To extend the closing date one month, a written request must be filed with the Executive Secretary, ASCE. This paper is part of the copyrighted Journal of the Hydraulics Division, Proceedings of the American Society of Civil Engineers, Vol. 87, No. HY 1, January, 1961.

[1] Hydraulic Engr., U. S. Geological Survey, Washington 25, D. C.

developed except by non-parametric methods. However, a relation between magnitude, probability of exceedance (the occurrence of an event greater than that magnitude), and design period (defined later) will allow the variability of the recurrence interval to be assessed and will therefore provide more complete information than is given by the conventional frequency curve. Such a relation is developed in this paper.

The theory used and the results obtained are not original. Beard,[2] Davenport,[3] Thomas,[4] Court,[5] Gumbel,[6] Kendall,[7] and others have made similar analyses. However, the method of presenting the results is new, and the importance of the theory would seem to justify its wider notice.

The construction of a cumulative frequency curve from a probability density function and from a small sample of observations is described first. Next is given the statistical interpretation of the cumulative frequency curve and the derivation of the relation between magnitude, probability, and design period. Finally, the theoretical results are substantiated by sampling from a list of 1,023 tree-ring indexes.

CONSTRUCTION OF CUMULATIVE FREQUENCY CURVES

Consider the histogram of Fig. 1 that shows the frequency of events for several ranges of magnitude. If the number of observations is allowed to approach infinity at the same time that the class interval (width of the rectangles) approaches zero, the enveloping line of the histogram will approach a smooth curve. Then if the ordinate values are divided by a number, such that the area under the curve becomes one, the resulting curve is a probability density curve, also shown in Fig. 1.

For the theoretical development of the cumulative frequency curve, assume that the probability density curve is known to be that of Fig. 1. By definition, the probability of a random event falling in any particular interval is the ratio of the area under the curve within that interval to the total area under the curve. The hatched area under the curve of Fig. 1 is one-tenth of the total, and by the preceding definition the probability is 0.1 that a random event will be greater than E. There is no probability associated with the exact event E. Probabilities in continuous distributions refer only to an event being within a certain range or of being larger or smaller than some magnitude E.

In hydrology it is conventional to interpret the cumulative frequency curve as giving the probability of occurrence of an event "equal to or greater (less) than." The "equal to" portion of the statement is not supported by theory, has no practical meaning, and, therefore, is not used in this paper.

If the area under the curve of Fig. 1 is divided into many vertical strips, the relative area of each determined and these relative areas plotted cumulatively

---

2 "Statistical Analysis in Hydrology," by L. R. Beard, Transactions, ASCE, No. 108, 1943, pp. 1110-1121.

3 "Discussion of Characteristics of Heavy Rainfall in New Mexico and Arizona by Luna B. Leopold," by R. W. Davenport, Transactions, ASCE, No. 109, 1944, pp. 877-878.

4 "Frequency of Minor Floods," by Harold A. Thomas, Jr., Journal, Boston Soc. of Civ. Engrs., Vol. 35, No. 4, 1948, pp. 425-442.

5 "Some New Statistical Techniques in Geophysics," by Arnold Court, Advances in Geophysics, Academic Press, Inc., New York, Vol. 1, 1952, pp. 45-85.

6 "The Calculated Risk in Flood Control," by E. J. Gumbel, Applied Science Research, The Hague, Holland, Sect. A, Vol. 5, 1955, pp. 273-280.

7 "Statistical Analysis of Extreme Values," by G. R. Kendall, First Canadian Hydrology Symposium, Natl. Research Council of Canada, November 4 and 5, 1959.

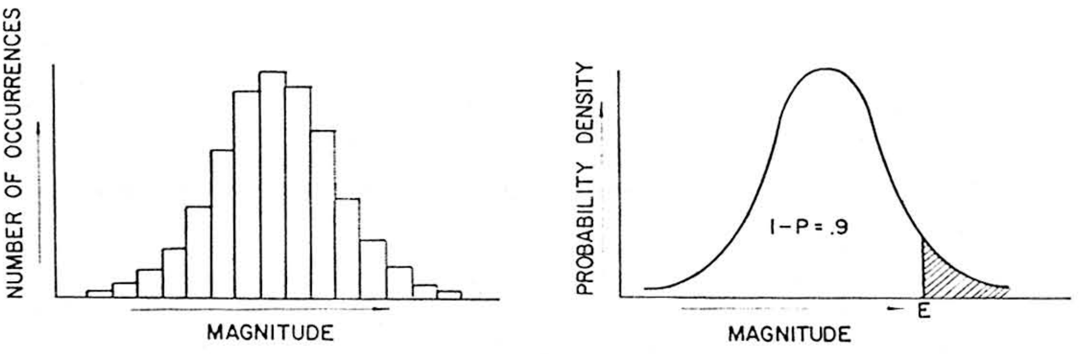

FIG. 1.—HISTOGRAM AND PROBABILITY DENSITY CURVE.

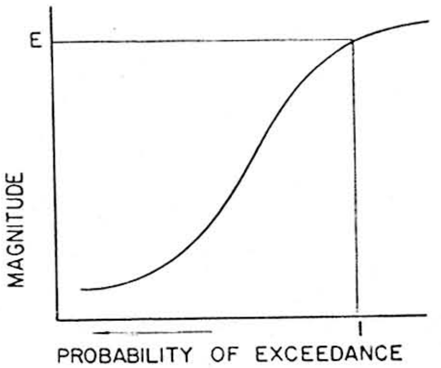

FIG. 2.—CUMULATIVE FREQUENCY CURVE FROM THE PROBABILITY DENSITY CURVE OF FIG. 1

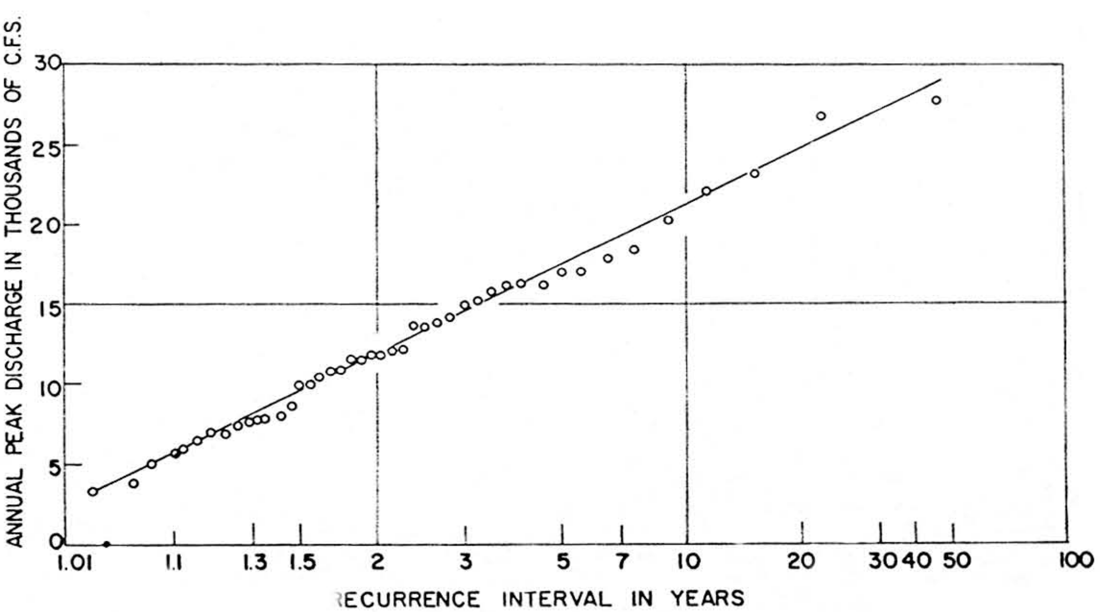

FIG. 3.—CUMULATIVE FREQUENCY CURVE

against magnitude, the result is a cumulative frequency curve such as is shown in Fig. 2. The cumulative curve is the integral of the density curve.

In practical work, the probability density curve is never known. The cumulative frequency curve must be developed directly from the data by one of two methods. The first requires the mathematical fitting of data to an arbitrarily-selected theoretical distribution. Procedures are described in the literature. The second method is semi-graphical and requires no assumptions as to the type of distribution. It is described in the following paragraphs as an aid in understanding subsequent developments.

The semi-graphical method of obtaining the cumulative frequency curve requires (1) arranging the data in order of magnitude, (2) computing the plotting position of each item, (3) plotting each item against its corresponding plotting position on probability paper, and (4) fitting a line to the plotted points.

Plotting positions may be computed by one of several formulas. The more common ones are

$$p = \frac{M}{N+1} \qquad (1)$$

$$p = \frac{M - \frac{1}{2}}{N} \qquad (2)$$

and

$$p = \frac{M}{N} \qquad (3)$$

in which p is the probability of exceedance, N is the number of events (or years, for annual events) used in preparing the frequency curve, and M is the order number of an event when the events are arranged in order of magnitude from the largest to the smallest, with M = 1 for the largest. More commonly the reciprocal of the probability, called the return period or recurrence interval, is plotted, and that approach is used in this paper. The formula used here is

$$R.I. = \frac{N+1}{M} \qquad (4)$$

in which R. I. is recurrence interval. The following developments are not affected by the equation chosen to compute the recurrence intervals.

Only three types of probability plotting paper are commonly used; normal, log-normal, and extreme value. The frequency curves given herein are plotted to a recurrence interval scale based on an extreme-value distribution proposed by Gumbel.[8] When graphic interpretation of the plotted points is used the resulting line is not always straight; if not, the line does not represent an extreme-value distribution of the Gumbel[8] type. An example of a cumulative frequency curve is given in Fig. 3 for annual floods on John Day River at McDonald Ferry, Oregon, from 1905 to 1948.

INTERPRETATION OF THE CUMULATIVE FREQUENCY CURVE

As previously stated, the density curve of Fig. 1 defines the probability of a random event being smaller than E as 0.9 and larger than E as 0.1. Similarly, on the cumulative frequency curve of Figure 2, the event E corresponds to a probability of exceedance of 0.1. Thus, the cumulative frequency curve

8 "On the Plotting of Flood Discharges," by E. J. Gumbel, Transactions, American Geophysical Union, Part II, 1943, pp. 699-719.

shows the probability that a single random event will exceed a given magnitude. If the cumulative frequency curve is based on annual events (that is, only the largest event per year is used), then the probability given is that of an annual event exceeding a certain magnitude. Now suppose the abscissa scale of Fig. 2 is changed to a recurrence-interval scale by taking reciprocals of the probabilities. The 0.1 probability becomes a 10-yr recurrence interval, and E is called the 10-yr event. This means that the average time between annual events that exceeds E is 10 yr. This is illustrated in Fig. 4 by considering the

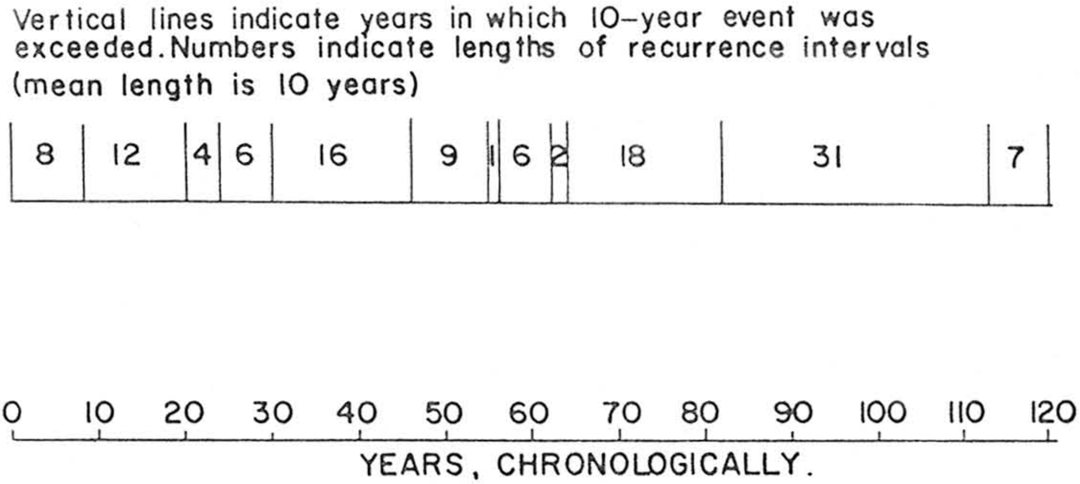

FIG. 4.—HYPOTHETICAL SEQUENCE OF RECURRENCE INTERVALS DURING A 120-YEAR PERIOD

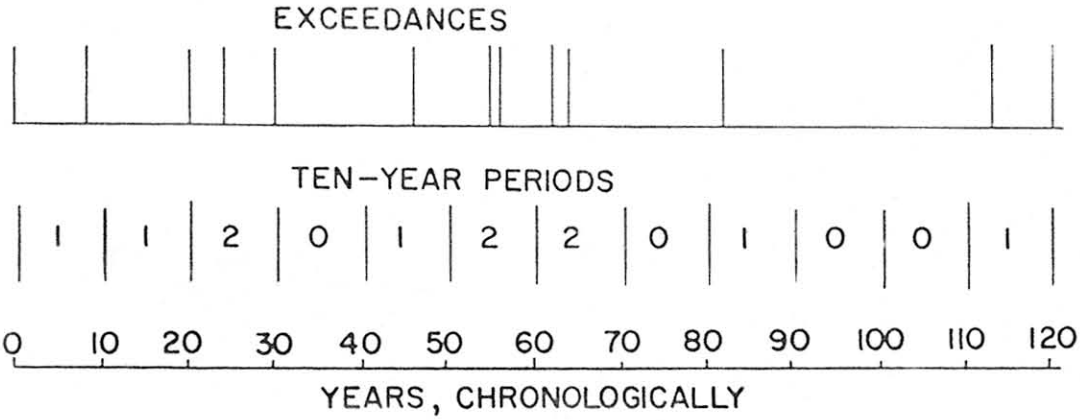

FIG. 5.—RANDOM SEQUENCE OF RECURRENCE INTERVALS (FROM FIG. 4), SHOWING THE NUMBER OF EXCEEDANCES IN EACH 10-YEAR PERIOD

random sequence of recurrence intervals in 120 yr. Although these intervals between exceedances range from one to 31 yr the average is 10 yr. Therefore, the event exceeded could be called a 10-yr event. Neglecting one of the two end exceedances, there are 12 exceedances in 120 yr for a probability of exceedance in any one year of 12/120 or 0.1, as previously defined. If both or neither of the end exceedances were neglected, the computed probabilities would approximate the 0.1 value.

Recurrence interval is an average value, thus, the n-year event will be exceeded at intervals averaging n years in length, but will also be exceeded in more than half of a series of n-year periods. If clarification is needed, refer again to Figs. 1 and 2. The probability of not exceeding E is 0.9, Fig. 1, and of exceeding E, 0.1, Fig. 2. If E is assigned a value equal to the 10-yr event, the probabilities remain the same. The probability of 0.9 applies to one annual event not exceeding E. The probability of not exceeding E in 10 yr is, by the multiplicative law

$$(0.9)^{10} = 0.35$$

Similarly, the probabilities of exactly one, exactly two, etc., exceedances in the 10-yr period can be obtained by solving the binomial equation

$$f(x) = {}_nC_x \, p^x \, (1-p)^{n-x} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (5)$$

in which $f(x)$ is the probability of x exceedances in n trials, ${}_nC_x$ is the number of combinations of n things taken x at a time, and p is the probability of an exceedance in one trial (see Mood[9]). For this problem, the probabilities are:

$$P\ (1 \text{ exceedance in } 10 \text{ yr}) = {}_{10}C_1(.1)\ (.9)^9 = .3874$$

$$P\ (2 \text{ exceedances in } 10 \text{ yr}) = {}_{10}C_2(.1)^2\ (.9)^8 = .1935$$

$$P\ (3 \text{ exceedances in } 10 \text{ yr}) = {}_{10}C_3(.1)^3\ (.9)^7 = .0576$$

$$P\ (4 \text{ exceedances in } 10 \text{ yr}) = {}_{10}C_4(.1)^4\ (.9)^6 = .0105$$

$$P\ (5 \text{ exceedances in } 10 \text{ yr}) = {}_{10}C_5(.1)^5\ (.9)^5 = .0015$$

The sum of the probabilities of 0, 1, 2, ... , 10 exceedances (in a 10-yr period) equals one.

The probability of not exceeding the n-year event in n years is, for other values of n and for known values of the probability of exceedance in one year:

$$(0.80)^5 = 0.33$$

$$(0.95)^{20} = .36$$

$$(0.98)^{50} = .364$$

$$(0.99)^{100} = .366$$

$$(0.999)^{1,000} = .368$$

Therefore, the n-year event has a probability of approximately 1 - 0.36 = 0.64 of being exceeded one or more times in an n-year period; or 64 of 100 n-year periods would include at least one exceedance of the n-year event; or the n-year event will be exceeded at least once in about 64% of a series of n-year periods. The last is shown empirically in Fig. 5 in which the sequence of exceedances of Figure 4 is repeated. The 120-yr period is divided into 12

9 "Introduction to the Theory of Statistics," by A. M. Mood, McGraw-Hill Book Co., Inc., New York, 1950, pp. 54-58.

ten-yr periods and 8 of the 12 periods (67%) include exceedances of the ten-yr event.

DESIGN-PROBABILITY CURVES

It has been shown that the probability of the n-year event (from a cumulative frequency curve of annual values) being exceeded in an n-year period is about 0.64. Additional information regarding the probabilities of exceedances in a definite period of years would be useful and can be obtained. Consider the use to which a frequency relationship is put. Design of a project logically might begin with selection of a design period, the number of years for which the project is expected to operate. Having fixed that period, the designer would inquire as to the probability of occurrence of damaging floods or of inadequate supply during the period. The conventional frequency curve cannot answer this inquiry adequately. A relation between (1) magnitude, (2) probability of not exceeding that magnitude, and (3) design period would supply the answer needed.

TABLE 1.—VALUES OF p AND R. I.

n (Design period) (1)	p (2)	R.I. = 1/p (3)
2	.293	3.4
3	.206	4.8
5	.129	7.8
10	.067	14.9
20	.034	29.4
30	.023	43.0
40	.017	59.0
50	.0138	72.0
100	.0069	145.0
1,000	.00069	1,450.0

Such a relationship can be obtained by modifying the conventional frequency curve. The modifications to be described are not confidence limits on the position of the frequency curve. They provide a more complete interpretation of the frequency curve as defined by the data.

Assume it is desired to define a relationship between magnitude of an event E and a design period such that there is a 0.5 probability of not exceeding the event in the design period. Further assume that the magnitude-frequency relationship is exactly defined to large recurrence intervals and that the events are randomly distributed in time. Let 1-p equal the probability of not exceeding E in one year and let n equal the number of years. Then

$$(1 - p)^n = 0.5$$

from which the values of p and R. I. given in Table 1 are computed.

These results indicate that the 29.4-yr event from the cumulative frequency curve is the one that has a 0.5 probability of not being exceeded in a single 20-yr period. Likewise it is the 72-yr event that has an even chance of not being exceeded in a 50-yr period.

These results may be used to modify a frequency curve by plotting the magnitude of the 3.4-yr recurrence interval event at the 2-yr design period, the

7.8 at 5, the 14.9 at 10, and so forth. Similar adjustments for different probabilities of not being exceeded can be computed by substituting the desired probability instead of 0.5 in the above formula and recomputing p and R.I. Results for the probabilities of 0.25 and 0.75 are given in Table 2.

The conventional frequency curve of Fig. 3 and the data in Tables 1 and 2 are used as the basis for constructing the curves of Fig. 6. Fig. 6 shows the design-probability curves for annual floods on John Day River at McDonald Ferry, Oregon (based on the curve of Fig. 3). For example, the magnitude of the event having 0.5 probability of not being exceeded in a 10-yr design period is that corresponding to the 14.9-yr recurrence interval (see Table 1) on Fig. 3. Other points are obtained similarly. Notice that the abscissa on Fig. 6 is the length of the design period and not a recurrence interval. These curves are named "design-probability curves" to distinguish them from the conventional cumulative frequency curve. It should be noted that the middle curve of Fig. 6, that for P = 0.50, is essentially the same as the curve obtained by the Beard[2] method.

TABLE 2.—PROBABILITY OF NOT EXCEEDING IN n YEARS

Design Period, n, in years (1)	0.25		0.75	
	p (2)	R.I. (3)	p (4)	R.I. (5)
2	.5	2.0	.134	7.5
3	.37	2.7	.092	10.9
5	.242	4.1	.056	17.8
10	.130	7.7	.028	35.7
20	.067	14.9	.0141	71.0
30	.045	22.2	.0095	105.0
40	.034	29.4	.0072	139.0
50	.0273	36.6	.0057	175.0
100	.0138	72.5	.0029	345.0
1,000	.00138	725	.000283	3,500.0

To further interpret the curves, consider a 20-yr design period on Fig. 6. The probability is 0.5 that a flood of 26,700 cfs will not be exceeded in the 20-yr period. Likewise, the probability is 0.25 that a flood of 23,000 cfs, and 0.75 that one of 31,000 cfs will not be exceeded in the 20-yr period. The latter discharge is based on an extension of the frequency curve (Fig. 3). Design-probability curves for small probabilities of exceedance are thus seen to be limited in extent by the length of the frequency curve from which they are obtained. If the frequency curve of Fig. 3 were plotted on Fig. 6 it would correspond to a probability of about 0.36.

This type of plot (Fig. 6) does furnish many answers needed by the designer. Further, it makes clear that no matter what design period is used, there is still an appreciable probability of experiencing an extremely large event in that period. This tends to be overlooked in interpreting the conventional frequency curve.

Although the probability of experiencing no exceedances is of primary interest, the probabilities associated with other outcomes help to complete the picture. These additional probabilities, computed as shown in the previous section, are given for several sizes of events and for 5 design periods in Table 3. Minor inconsistencies in the table are due to the limited number of significant

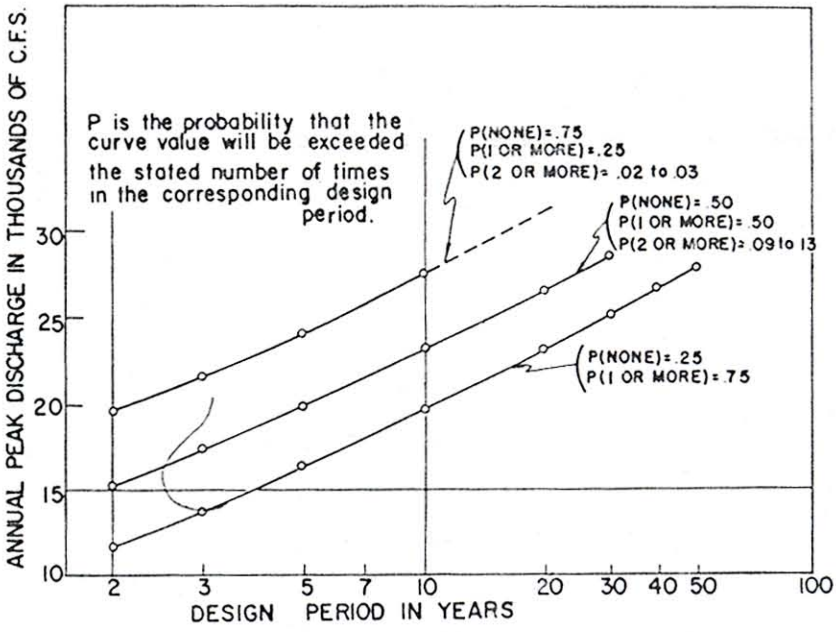

FIG. 6.—DESIGN-PROBABILITY CURVES BASED ON THE CURVE OF FIG. 3

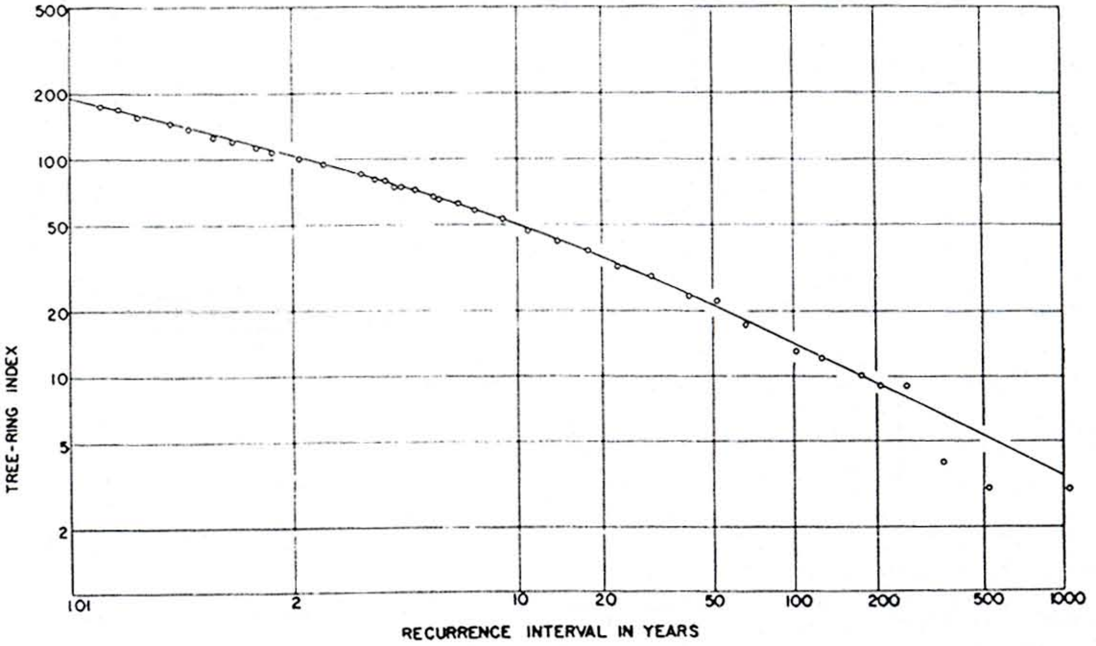

FIG. 7.—CUMULATIVE FREQUENCY CURVE OF TREE-RING INDEXES FOR RIO GRANDE AREA, NEW MEXICO, 908-1930 A. D.

figures used. For a given probability of no exceedances, the probability of one or more exceedances is fixed (because the sum of the two must equal one), but the probability of two or more exceedances varies somewhat with the length of the design period. It is for this reason that a range rather than a unique value is shown for this condition in Fig. 6.

APPLICATION OF DESIGN-PROBABILITY CURVES TO A LONG RECORD OF NATURAL EVENTS

A record of 1,023 tree-ring width indexes for the Rio Grande area, New Mexico, for the period 908 A. D. to 1930 A. D. has been compiled by Smiley,

TABLE 3.—PROBABILITY OF EXCEEDANCE OF VARIOUS EVENTS FOR FIVE DESIGN PERIODS

Design period, in years	R. I. of event from frequency curve	Probability of exceedance in one year	Number of exceedances in design period							
			0	1	2	3	4	1 or more	2 or more	3 or more
2	2.0	.500	.25	.50	.250	...	...	0.75	0.25	...
2	3.4	.293	.50	.41	.086	...	...	.50	.09	...
2	7.5	.134	.75	.23	.018	...	...	.25	.02	...
5	4.13	.242	.25	.40	.255	.082	.0130	.75	.35	.095
5	5.0	.200	.33	.41	.205	.051	.0064	.67	.26	.055
5	7.8	.129	.50	.37	.110	.016	.0013	.50	.13	.020
5	17.8	.056	.75	.22	.026	.002	.0001	.25	.03	.004
10	7.7	.130	.25	.37	.250	.100	.0273	.75	.38	.130
10	10.0	.100	.35	.39	.194	.057	.0112	.65	.26	.066
10	14.9	.067	.50	.36	.116	.022	.0028	.50	.14	.024
10	35.7	.028	.75	.22	.028	.002	.0001	.25	.03	.002
20	14.9	.067	.25	.36	.245	.105	.0319	.75	.39	.145
20	20.0	.050	.36	.38	.189	.059	.0133	.64	.26	.071
20	29.4	.034	.50	.35	.118	.025	.0028	.50	.15	.032
20	71.0	.014	.75	.21	.030	.003	.0002	.25	.04	.010
40	29.4	.034	.25	.35	.242	.110	.0352	.75	.40	.158
40	40.0	.025	.36	.37	.185	.060	.0143	.64	.27	.085
40	59.0	.017	.50	.34	.118	.026	.0041	.50	.16	.042
40	139.0	.007	.75	.21	.030	.003	.0002	.25	.04	.010

Stubbs, and Bannister.[10] The record is a composite each part of which has been adjusted for growth trend. The annual growth of a tree (as measured by the tree-ring width index) in certain locations in the arid West is very sensi-

10 "A Foundation for the Dating of Some Late Archeological Sites in the Rio Grande Area, New Mexico," Bulletin, University of Arizona, Vol. XXIV, No. 3 (Laboratory of Tree-Ring Research Bulletin, No. 6).

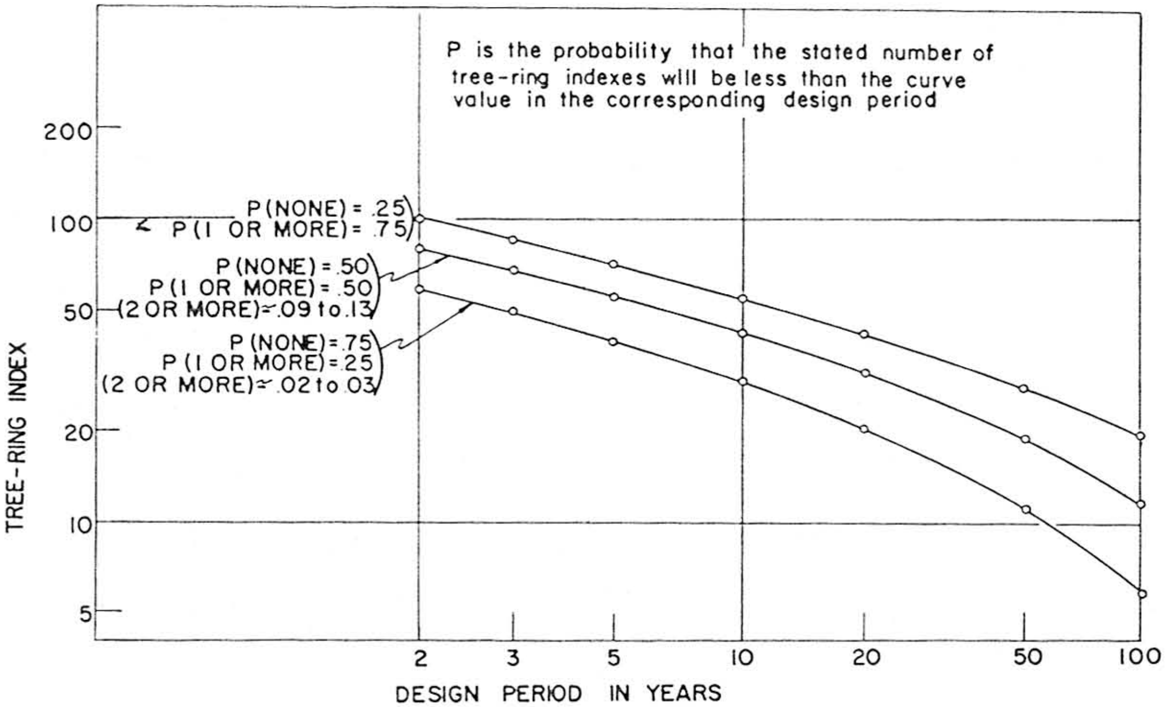

FIG. 8.—DESIGN PROBABILITY CURVES OF TREE-RING INDEXES (BASED ON CURVE OF FIG. 7)

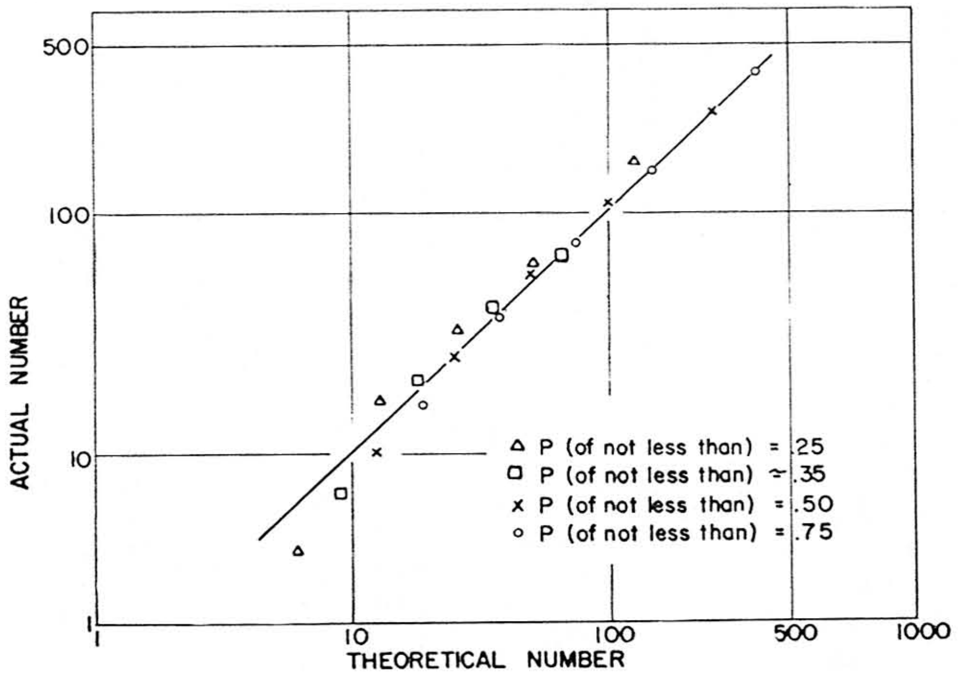

FIG. 9.—COMPARISON OF ACTUAL AND THEORETICAL NONEXCEEDANCES FROM ANALYSIS OF A 1,023-YEAR RECORD OF TREE-RING INDEXES

tive to weather conditions. As previously proposed by Douglass[11] the record of tree-ring width indexes has some of the characteristics of a weather record.

Fig. 7 shows the cumulative frequency curve obtained from the 1,023 indexes. From that curve the design-probability curves of Fig. 8 were obtained by the method previously described. These design-probability curves were verified by dividing the 1,023-yr record into n-year periods and comparing the actual and theoretical number of periods in which the appropriate index value was not exceeded.

From Tables 1 and 2 and Fig. 7, the values shown in Table 4 are obtained for a 20-yr period.

The values of tree-ring index in Table 4 are the ones that define the design-probability curves of Fig. 8 at a design period of 20 yr. Using these index values, the corresponding theoretical number of nonexceedances can be checked

TABLE 4

Probability of not being exceeded in 20 years (1)	Recurrence interval (2)	Tree-ring index (3)
0.25	14.9	41
.36	20.0	35
.50	29.4	29
.75	71	17

by counting the number of 20-yr periods in the original record in which these index values are not exceeded (not exceeded means "greater than" here because the cumulative frequency curve gives values of "less than"). There are 51 twenty-yr periods in the record. The number of 20-yr minimums greater than the index values are given in Table 5. The record was also divided into

TABLE 5.—TWENTY YEAR MINIMUMS GREATER THAN THE INDEX VALUES

Index Value	41	35	29	17
Theoretical Number	12.8	18.4	25.5	38.2
Actual number	17	20	25	38

2, 5, 10, and 40-yr periods and similarly studied. These results together with those obtained for 20-yr periods are shown in Fig. 9. Agreement between actual and theoretical numbers is good enough to substantiate applicability of the method to this type of natural event. For this example the cumulative frequency curve is known. In actual practice, a curve based on a small sample must be used and the reliability of the results obtained will depend on the sampling error.

11 "Climatic Cycles and Tree Growth: A Study of the Annual Rings of Trees in Relation to Climate and Solar Activity," by A. E. Douglass, Publication No. 2890, Carnegie Inst., Washington, D. C., 1928.