DESIGN STORMS, FLOWS AND FLOODS

Introduction

Engineers design for extreme events. Dams must be built high enough to contain extreme floods, while bridges must be built high enough to remain above the high-water mark. The challenge is how to design against these extreme events which are rare and in the absence of comprehensive data.

Design storm:

A design storm is a precipitation pattern defined for use in design of a hydrologic system. It *serves as a system input*, and the resulting rates of flow through the system can be calculated using R-R relationships or flow routing procedures. *Design storms can be based on historical rainfall data at a site or can be constructed using general rainfall characteristics in the surrounding area.* Their application ranges from use of point rainfall values in the rational formula to determine peak flow rates for storm sewers and highway culverts design, to use of storm hyetographs as inputs in R-R analysis of urban detention basins or for design of spillways in large reservoir projects.

For efficient / economic design of flood control structures it is important that floods be *estimated accurately.* Design of culverts, road and rail bridges, drainage works and irrigation diversion works require ***accurate flood estimation*** at the concerned site. Hydrologists are also required to provide an estimate of the maximum flood that would occur at a site during a specified period of time in future. But because runoff is a random process it can only be estimated with a certain *degree of probability.* The decision as to what ***flood magnitude*** the structure is ultimately designed for depends on the client taking into consideration economy, safety and all other factors into consideration. A higher value increases the cost of the hydraulic structure, while a low value increases risk to the structure and to the population below. ***Design Flood***, is the flood adopted for the design of a structure and could be;

- The entire flood hydrograph.
- The peak discharge of the flood hydrograph.

Design flood may be MPF or SPF or a flood of any desired *recurrence interval* depending on the degree of flood protection required and the cost of constructing the structures to the desired flood stage. It is usually selected after *cost-benefit analysis* and application of sound engineering judgement. Design flood is related to project features e.g. a design flood adopted for *spillway design may be higher than that adopted for a control reservoir or temporary coffer dam*. The most economical design flood is often found after consideration of various options and is in most cases less than MPF which is uneconomical.

When a structure is designed for a value lower than MPF, there is a certain degree of flood risk to the structure. It is also uneconomical to design for 100% flood protection so a balance must be struck between economy and safety. Protection against high but rare floods is also uneconomical because of the large investment. In estimating design flood, reference is made to *two classes:*

Maximum probable flood (MPF) is the flood that may be expected from the most severe combination of critical meteorological and hydrological conditions reasonably possible in an area. It is generally a very large flood rarely used except in spillway design. It is computed using Probable Maximum Storm (PMS) which is an estimate of the physical upper limit of storm rainfall over the catchment. **MPF** is obtained from records of all storms that have occurred in an area and maximizing them for the most critical atmospheric condition. It differs from SPF as it includes extremely rare and catastrophic floods. It is confined to the design of spillways for very high dams. For a basin, SPF is usually about 80% of MPF.

Standard Project Flood (SPF): Is the flood that results from the most severe combination of meteorological and hydrological conditions considered reasonably characteristic of an area. It is computed from Standard Project Storm (SPS) in the catchment under consideration and may be taken as the heaviest storm observed in the basin. It is not maximized for the most critical atmospheric conditions but may be transposed from an adjacent basin (meteorologically similar) to the catchment under consideration in the absence of storm records.

Flood of a specific return period: is estimated through frequency analysis of annual flood values for a sufficient length of time. When flood data is *inadequate,* frequency analysis of available storm data is made and the storm of a particular frequency applied to the unit hydrograph to derive the design flood.

Methods to estimate design floods include but are not limited to:

i. Empirical formulae- location specific and should be used as a last resort. Developed for specific conditions and only usable under those conditions and for structures of less importance.
ii. Increasing observed maximum flood by a certain percentage. Used only in the absence of data.
iii. Rational method.
iv. Unit Hydrograph method
v. Statistical methods (***Frequency analysis***)
vi. Physical indications of past floods — flood marks on river reaches and local history together with manning's equation can be used to estimate flood discharge (include FoS of at least 1.5)

For an economic design consistent with safety, it is recommended to use complete flood hydrograph rather than only the peak flood. This is best achieved through *the UH approach.* In frequency analysis the flood hydrograph may be constructed *by making its peak discharge equal to the estimated design* flood, and other features then added to resemble the observed hydrographs constructed from available records. The method adopted depends on the type and amount of data available.

Rational method

This assumes that *maximum flood flow is produced by a certain rainfall intensity lasting for a time equal to or greater than the time of concentration (t_c).* When the storm lasts longer than t_c every part of the catchment would be contributing to runoff at the outlet representing the condition of peak runoff. The runoff rate corresponding to this condition is then given by:

Q = CIA *Where*
A is catchment area, **I** is rainfall intensity and **C** is runoff coefficient that accounts for rainfall abstractions. In SI units if A is in Km^2 and I in cm/hr then Q is given by:

$$Q = 2.778CIA$$

Maximum rainfall intensity depends on duration and frequency (***Refer IDF curves***). The intensity used in the above equation should therefore correspond to the duration equal to the time of concentration which can be obtained from Kirpich' equation (1940):

$$t_c = 0.0195 L^{0.77} S^{-0.385}$$ Where

t_c is in minutes, L is the maximum. length of travel of water along the water course (m) and S is the slope expressed as the ratio of the difference in elevation (H) between the remotest point on the catchment and the outlet to the length L (S = H/L). t_c is taken to be equal to storm duration.

Once the time of concentration has been determined, rainfall intensity can then be determined using IDF relationship applicable to the catchment (**figure 1**). The Return Period adopted may vary from 5 to 50 years depending on the type and importance of the structure. Once rainfall intensity and catchment area are obtained, a runoff coefficient C applicable to the design condition is selected and peak flood estimated using the equation. Typical values of runoff coefficients are as given in table 1: A given catchment may have different characteristics requiring use of different coefficients for different subareas in which case $Q = 2.778I\sum(C_i A_i)$ where C_i is the runoff coefficient of *i*th sub-catchment and A_i is the drainage area of *i*th sub-catchment.

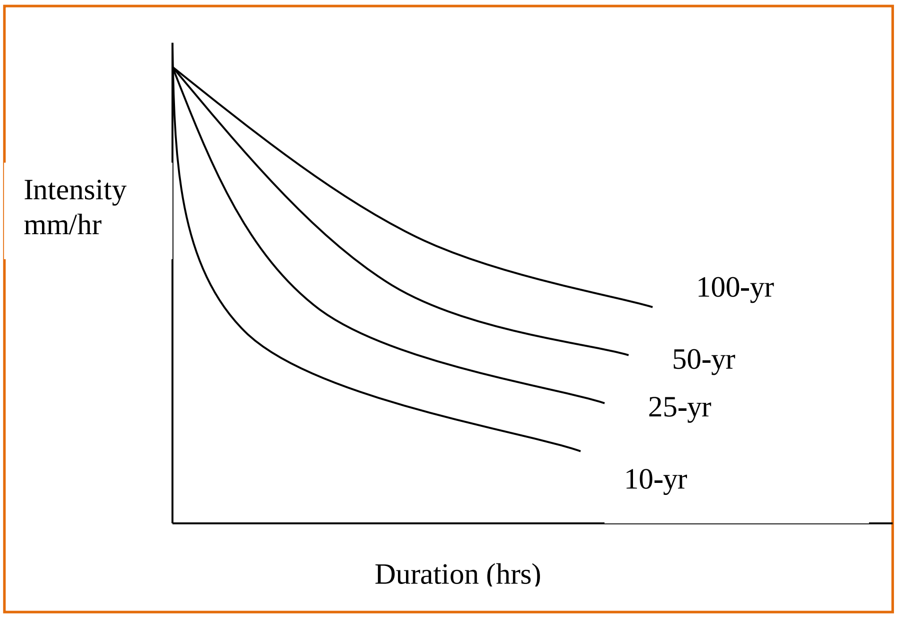


Figure 1: Sample IDF curves

Table 1: Typical runoff coefficients for various areas

Type of terrain	*Value of C*
Flat residential areas	*0.4*
Moderately steep residential area	*0.6*
Built up areas—impervious	*0.8*
Rolling lands and clay-loam soils	*0.5*
Hilly areas, forests, clay and loamy soils	*0.5*
Flat cultivated lands and sandy soils	*0.2*

The rational formula gives good results for catchments below 50km^2 and is used for estimating peak floods in the design of urban drainage systems (storm sewers, culverts and bridges). Once the design discharge Q_p entering a sewer has been calculated, the pipe diameter D required to carry this discharge can then be determined from Manning's equation.

$$Q_p = A\frac{1}{n}m^{\frac{2}{3}}S^{\frac{1}{2}}$$

Limitations of the rational formula in estimation of flood magnitudes

i. It gives only peak runoff but does not provide the complete hydrograph.
ii. Rainfall intensity must be constant over the whole watershed during the time of concentration.
iii. Assumption of rainfall duration ≥ time of concentration for maximum flow is not always true.
iv. The formula is best suited to small catchments with areas less than 50 km^2.
v. Coefficient **C** is assumed same for all storms meaning losses are constant for all storms.

Example:

An outlet is to be designed for a rural watershed of area 12 km^2 of which 30% is cultivated area (C = 0.20), 50% is under forest (C = 0.10) and 20% under grass cover (C =0.35). The watercourse is 1.8 km long with a drop of 22 m. The IDF relation for the area is given by:

I = $80T_r^{0.2}\ \frac{80T_r^{0.2}}{(t+13)^{0.46}}$ where T_r is in years and t is in minutes.
Estimate the peak runoff with a return period of 25 years.

Length of water course = 1.8 km = 1800m
Drop in elevation = 22m

$\therefore$ slope S = $\frac{22}{1800} = 0.01222$

$t_c = 0.0195 L^{0.77} S^{-0.385}$ = $t_c = 0.0195(1800)^{0.77}(0.01222)^{-0.385}$ = 34 minutes

$I \frac{80(25)^{0.2}}{(34+13)^{0.46}} = 25.9 cm/hr$

A_1 = 3.6 C_1 = 0.20
A_2 = 6.0 C_2 = 0.10
A_3 = 2.4 C_3 = 0.35

Substituting these values in the equation we get peak runoff rate
Q = 2.778 x 25.9 (3.6 x 0.2) + (6.0 x 0.1) +(2.4 x 0.35) **= 155.41m³/s**

Estimation of design flows in small ungauged watersheds.

Various methods are available for estimating peak flow rates required for application in **small urban and rural watersheds** with no adequate runoff data. Some are based on a rational analysis of the R-R process, whereas others are completely empirical or correlative in that they predict peak-runoff rates by correlating flow rates with simple drainage basin characteristics, such as area or slope. These methods are limited in a number of ways, for example, application of the rational formula is difficult unless return periods for rainfall and runoff are assumed to equal. For empirical formulas, estimates of coefficients required in the formula are subjective.

Estimation of design flows in gauged watersheds

Rational and empirical methods are appropriate for small *ungauged* catchments. For large catchments with sufficient hydrologic data, a calculated risk can be taken by designing hydraulic structures **for floods less than the most severe flood**. The appropriate return period for use in such design is selected on the basis of economic analyses, policy considerations and acceptable risk. Methods used are: Flood frequency analysis and estimation of design flood from design storms. ***A design storm is an estimate of the rainfall amount and distribution over a particular drainage area, accepted for use in determining the design flood***. Depending on importance and acceptable level of risk, design floods may be taken as Probable Maximum Flood, Standard Project Flood or as specified by national policy (existing guidelines). Based on availability of data peak flow from a given gauged catchment can be estimated from two approaches:

i. When peak flood discharges are available
ii. When complete runoff hydrograph, rainfall data and PMP for the watershed available.

Approach (i) uses frequency analysis while approach (ii) uses the UH method

Probable Maximum Flood (PMF): Is the greatest flood that can be expected assuming ***complete coincidence*** of all factors that would produce *the heaviest rainfall and maximum runoff.* It is derived from PMP; hence its frequency cannot be determined. Economically it is *prohibitive* to design a structure for PMF, except for very large spillways whose failure could lead to excessive *damage and loss of life*. A pragmatic approach for many design situations is to define *design flood as an estimated limiting value, but scale it downwards by a percentage depending on the type of structure and potential hazard if it fails.* For this purpose, the flood event actually used in design is often the greatest flood that may reasonably be expected, taking into account all pertinent conditions of location, meteorology, hydrology, and topography. This design flood may be determined analytically as the flood caused by a *transposed* historically largest storm and magnitude may be a fraction of the *ELV*. In practice, design flood (*standard project flood)* is estimated using R-R modelling by applying the UH method to the standard project storm, which is the greatest storm that may be reasonably expected. SPS is derived from a detailed analysis of storm patterns and transposition of storms to a position that gives maximum runoff. Past estimates have indicated that SPS magnitudes and SPF discharges are generally in the range of 40 to 60% of ELV (estimated limiting values) for the same basins. *SPF estimates are usually made only for major and intermediate structures since they require considerable effort in preparation.*

FREQUENCY ANALYSIS

Determination of the frequency of occurrence of extreme hydrological events e.g. floods, droughts etc. is important in water resources planning and management as there exists a definite relationship between frequency of occurrence and magnitude with ordinary events occurring regularly. **Probability distribution** helps to relate the magnitude of these extreme events with their number of occurrences such that their chance of occurrence with time can be predicted accurately.

Frequency analysis uses past observed data to predict future flood events, their probabilities and return periods. It is based on the assumption that combination of factors that produce floods are a ***matter of pure chance*** and are subject to analysis through probability theory. When stream flow peaks are arranged in a **descending order** of magnitude they constitute a statistical array whose distribution can be expressed in terms of frequency of occurrence.

There are two methods of compiling peak flood data-***annual series and partial duration series.*** In annual series, *only the highest flood* in each year is used while ignoring the next highest in any year, which may sometimes exceed many of the annual maximum. In partial duration series, all floods above a *set minimum* are included in analysis, regardless of the time-interval. They do not furnish a proper frequency series to enable *reasonable statistical analysis*. All large floods are used in analysis, which is an advantage.

In annual series some big floods are omitted because they were not the highest in any year.

Hydrological processes such as floods are a result of a number of component parameters (*catchment characteristics, rainfall and antecedent conditions,* each one of which depends on a host of constituent parameters thus very difficult to model them analytically. An alternate approach to flood flow prediction is ***frequency analysis*** which guides judgement in determining the capacity of bridges / coffer dams and as a means of estimating the probable flood damage that can be prevented by a system of flood protection works. Values of annual maximum flood from a given catchment for a large number of successive years constitute an **annual series.** When data in the series is arranged in a descending order of magnitude, the probability **P** of each event being ***equalled*** / ***exceeded*** can be calculated by the plotting position formula:

$P = \frac{m}{N+1}$ or return period or frequency by: $T = \frac{N+1}{m}$ Weibull method (1939):

Where m = order number of the event (m = 1 for highest flood) and N = total number of events in the data series (years of record). From the above Recurrence interval is also given by:

$T = \frac{1}{P}$ or the percent chance of its occurrence in any one year, *i.e.,* frequency (*F*) is $F = \frac{I}{T} x100$ while the probability that it will ***not*** occur in a given year, *i.e.* the probability of non-exceedance (P'), is P' = $1 - P$ or $1 - \frac{1}{T}$. Other formulae commonly used include:

Allen Hazen method (1930): $T = \frac{n}{m - \frac{1}{2}} = \frac{2n}{2m-1}$ *and California method* $(1923): T = \frac{n}{m}$

Weibull is the most commonly used approach.

Return period (T) is the average length of time between occurrences of an event or a greater one (Not the actual time interval). Return periods of annual series and *partial duration series* have different meanings. In annual series it means the average number of years between occurrences of an event of a given magnitude as annual maximum but in partial duration series it carries no implication of annual maximum.

For ***small return periods*** or where limited extrapolation is required, *a simple best fitting curve through plotted points is used as the probability distribution* with a logarithmic scale for T being advantageous. However, when large extrapolations of T are required, ***theoretical probability*** distributions are used. According to Chow the general equation of hydrologic frequency analysis is given by**:**

$$X_T = \bar{x} + K\sigma \quad \textbf{where}$$

x_T = value of the variate x of a random hydrologic series with a return period T, $\bar{x}$ = mean, σ = standard deviation, K = frequency factor which depends on return period and the *assumed frequency distribution.*

Commonly used frequency distribution functions used for predicting extreme flood events are: (i) **Gumbel's** extreme value distribution, (ii) **Log-Pearson type III** distribution and (iii) **Log normal** distribution. In flood frequency analysis, the usual problem is to predict the magnitude of extreme flood events and for this, *specific frequency distribution functions are assumed* and the required statistical parameters calculated from available data. From these parameters, *flood magnitudes of specified return periods are estimated*. Accuracy of frequency analysis depend on length of data used. Minimum number of years of record required to obtain reasonable estimates is *30 years*.

Data selection

Observed data is a sample of the total population of floods that have occurred and that may be expected to occur in future. If sample size is too small predictions of future floods is unreliable. Data used should be homogenous so that it represents a common set of catchment conditions (natural / modified). It may consist of annual series or partial duration series with choice depending on purpose of study. If extreme floods are the main concern, where flood magnitudes with exceedance probability of less than 0.5 are required, annual series is preferred. Where estimates of very frequent events with return periods less ***than 5 years*** are required (e.g. in urban drainage design) partial duration series is preferred.

Probability Plotting

After compiling the series, data is arranged in descending order of magnitude and assigned a rank (***m)*** with 1 representing the highest value. Return period for each event is then calculated from:

$$T = \frac{N+1}{m}$$ Where T is return period in years, m is the order and N is years of record.

If a graph of flood magnitude against its return period is plotted on simple plane coordinates, the resulting plot (probability distribution) may be extrapolated to obtain design flood *of any return period*. Results obtained this way are approximate as the extrapolation may be influenced by a few outlier points (fig. 1). The graph may also be drawn with a linear scale on the y-axis and a logarithmic scale on the x-axis. This approach is simple but not very accurate. In general probability plots can be used to find flood magnitudes with return periods of less than $\frac{n}{5}$. For longer return periods it is best to fit a theoretical distribution to the data using either **Gumbel or Log Pearson type III distributions.** *Prediction of flood peaks are reliable when analysis is done for return periods less than the data length.*

Example 1

Observed annual flood peaks for a stream from 1941 to 1980 is as given below in m^3/s: 395, 619,766, 422, 282, 900, 705, 528, 520,436, 697, 624, 496, 589, 598, 359, 686, 726, 527, 310, 408, 721, 814, 459, 440, 632, 343, 634, 464, 373, 289, 371, 522, 342,446, 366, 699, 560, 450 and 610. Construct a probability plot for the data on both ordinary **and semi-log graph papers.** From the plots determine the magnitude of a flood with a *return period of 100 years. C*ompare results from the two plots.

Solution

Year	Flood peak (m^3/s)	Flood peaks in descending order (m^3/s)	Order (m)	Return period $\left(T_r = \frac{41}{m}\right)$	Reduced variate $y = -InIn\frac{T_r}{T_r - 1}$
1	**2**	**3**	**4**	**5**	**6**
1941	395	900	1	41	3.7013
42	619	814	2	20.5	2.9955
43	766	766	3	13.67	2.5775
44	422	726	4	10.25	2.2764
45	282	721	5	8.20	2.0398
46	990	705	6	6.83	1.8432
47	705	699	7	5.85	1.6742
48	528	697	8	5.01	1.5022
49	520	686	9	4.55	1.3936
1950	436	634	10	4.10	1.2744
51	697	632	11	3.73	1.1644
52	624	624	12	3.4	1.0547
53	496	619	13	3.15	0.9625
54	589	610	14	2.93	0.8735
55	598	598	15	2.73	0.7849
56	359	589	16	2.56	0.7025
57	686	560	17	2.41	0.6235
58	726	528	18	2.28	0.5494
59	527	527	19	2.16	0.4753
1960	310	522	20	2.05	0.4019
61	408	520	21	1.95	0.3297
62	721	496	22	1.86	0.2595
63	814	464	23	1.78	0.1923
64	459	459	24	1.71	0.1290
65	440	450	25	1.64	0.0608
66	632	444	26	1.58	0.0022
67	343	440	27	1.52	-0.0701
68	634	436	28	1.46	-0.1441
69	464	422	29	1.41	-0.2112
1970	373	408	30	1.37	-0.2693
71	289	395	31	1.32	-0.3486
72	371	373	32	1.28	-0.4186
73	522	371	33	1.24	-0.4960
74	342	366	34	1.21	-0.5603
75	446	359	35	1.17	-0.6570
76	366	343	36	1.14	-0.7406
77	699	342	37	1.11	-0.8380
78	560	310	38	1.08	-0.9565
79	450	289	39	1.05	-1.1133
1980	610	282	40	1.03	-1.2630

NB: Information highlighted is for use with example 2

Probability plot of flood magnitudes (column **3)** against return periods (column **5)** on an ordinary graph paper is shown in fig. 1. From the plot, flood magnitude with a return period of 100 years is **1270 m^3/s.** Probability plot of flood discharge *on a linear scale* against its return period on *a logarithmic scale* (figure 2) gives **1060 m^3/s showing the** two results to be slightly different.

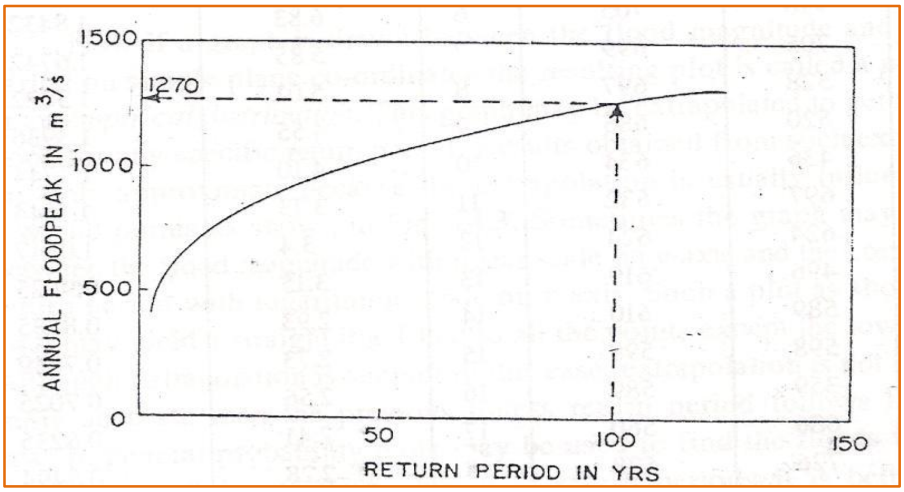


Figure 1: Ordinary probability plot

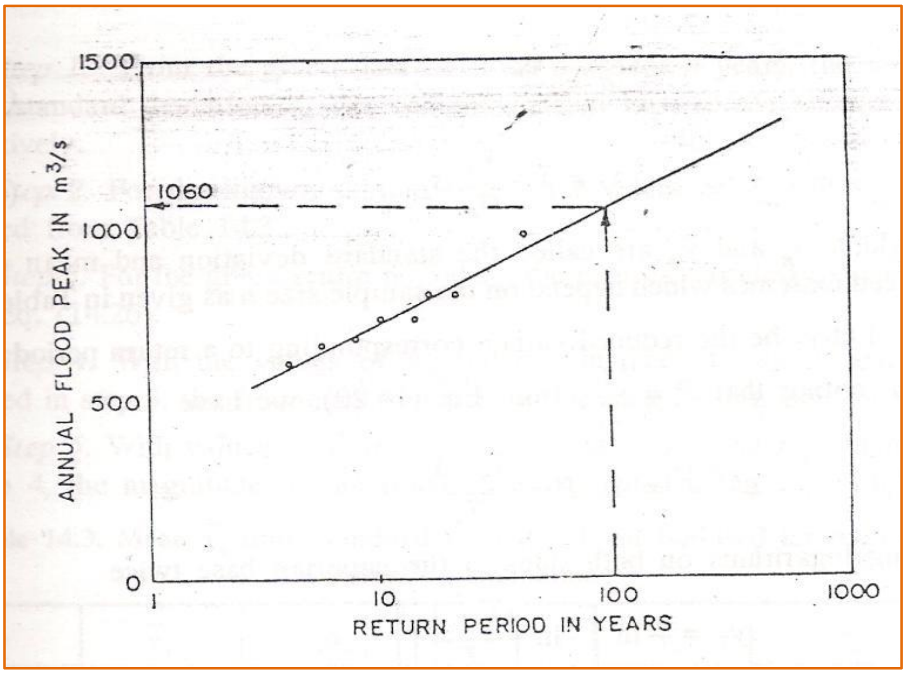


Figure 2: Logarithmic probability plot

Gumbel's extreme value distribution is one of the most widely used probability-distribution functions for extreme values in hydrological / meteorological studies. It is useful for predicting flood peaks, maximum rainfall, wind speed etc.

Use of this distribution function to describe annual peak discharge is based on the following assumptions:

i. Daily discharge follows an exponential type of distribution
ii. Original number of elements from which the peak value is selected is sufficiently large (365 daily discharges).
iii. Daily discharges are independent of each other.

Gumbel defined a flood *as the largest of the 365 daily flows and the annual series of flood flows constitute a series of largest values of flows.* According to his theory of extreme events, the probability of occurrence of an event equal to or larger than a value x_0 is given by:

$P(x \geq x_0) = 1 - e^{-e^{-y}}$ Where y is called the reduced variate given by $y = a(x - x_f)$

Where a and x_f *are parameters of distribution obtained from sample statistics through the method of moments.* For large sample sizes say n > 200, they are given by:

$a = \frac{1.2825}{s_x}$, $x_f = \bar{x} - 0.45005 s_x$ (x_f ***is also referred to as*** β**,** S_x **as** σ **and** $\bar{x}$ **as** μ)

Where $\bar{X}$ and S_x are the mean and standard deviations of the sample as computed from relevant equations. For ***small samples***, a and x_f are estimated using Gumbel's equations as follows:

$a = \frac{\sigma_n}{S_x}$ and $x_f = \bar{x} - \frac{\overline{y_n}}{\sigma_n} S_x$ where σ_n and $\overline{y_n}$ are standard deviation and mean of the *reduced extremes* which depend on the sample size n as given in **Table 1.**

Let y_T be the reduced variate corresponding to a return period T_r. Then since $P = \frac{1}{T_r}$ from the equation

$P = 1 - e^{-e^{-y}} = \frac{1}{T_r}$ we have $e^{-e^{-yT}} = 1 - P = \frac{T_r - 1}{T_r}$.

Taking logarithms on both sides to the natural base twice we have:

$$y_T = -In\left[-In\left(\frac{T_r - 1}{T_r}\right)\right] \text{ Or } y_T = -In\left[In\left(\frac{T_r}{T_r - 1}\right)\right]$$

Alternatively converting the logarithms to base 10, we have:

$$y_T = -\left[0.834 + 2.303 \log\log\left(\frac{T_r}{T_r - 1}\right)\right]$$

Let x_T denote the magnitude of a flood with a return period of T_r years. Then from the equation y = $a(x - x_f)$ we have $y_T = a(x_T - x_f)$. Substituting equations $a = \frac{\sigma_n}{S_x}$ and $x_f = \bar{x} - \frac{\overline{y_n}}{\sigma_n} S_x$ into equation $y_T = a(x_T - x_f)$ and simplifying, we get $x_T = \bar{x} + \left(\frac{y_T - \overline{y_n}}{\sigma_n}\right) x S_x$ or $x_T = \bar{x} + K_T x S_x$.

K_T known as ***frequency factor*** depends on return period and type of distribution and is given by $K_T = \left(\frac{y_T - \overline{y_n}}{\sigma_n}\right)$. $\overline{y_n}$ and σ_n are obtained from tables provided by Gumbel (**see table** 1)

$\overline{y_n}$ = reduced mean is a function of sample size N.

For a large sample (say n $\rightarrow \infty$, $\overline{y_n}$ equals 0.5772 and σ_n equals 1.2825) (***Table 1***)

According to Chow, most frequency distribution functions applicable in hydrologic studies can be expressed in the form $\bar{x} + K_T * S_x$.

Procedure for estimating the design flood of any return period using Gumbel's distribution is as follows:

- ❖ From the given flood peaks data for n years, the mean $\bar{x}$ and standard deviation s_x are computed using relevant equations.
- ❖ For the known sample size n, values of $\overline{y_n}$ and σ_n are obtained from **table 1**
- ❖ For the given return period T_r, the reduced variate y_T is calculated from equations $y_T = -In\left[In\left(\frac{T_r}{T_r - 1}\right)\right]$ or $y_T = -In\left[-In\left(\frac{T_r - 1}{T_r}\right)\right]$

- With values of $\overline{y_n}$ and σ_n obtained from table 1 and the y_T computed above, the frequency factor K_T is calculated using the equation $K_T = \left(\frac{y_T - \overline{y_n}}{\sigma_n} \right)$.
- With values of $\bar{x}$ and s_x obtained in the first step and K_T obtained above, the magnitude of the flood x_T with the given return period T_r is computed from the equation:

$$X_T = \bar{x} + K_T * S_x$$

Checking for Gumbel's fit

A simple graphical test to determine whether observed data follows Gumbel's distribution is as follows:

- After arranging the observed data in descending order of magnitude and assigning return period for each magnitude, the reduced variate corresponding to each return period is computed from:

$$y_T = -In\left[-In\left(\frac{T_r - 1}{T_r} \right) \right]$$

- A plot of *flood magnitude* and its *reduced variate* corresponding to each return period is then plotted on an ordinary graph paper. If *an eye* fit shows a straight line then, Gumbel's distribution *is a good fit for the data since x and y are linearly related through the equation* $y = a(x - x_f)$.

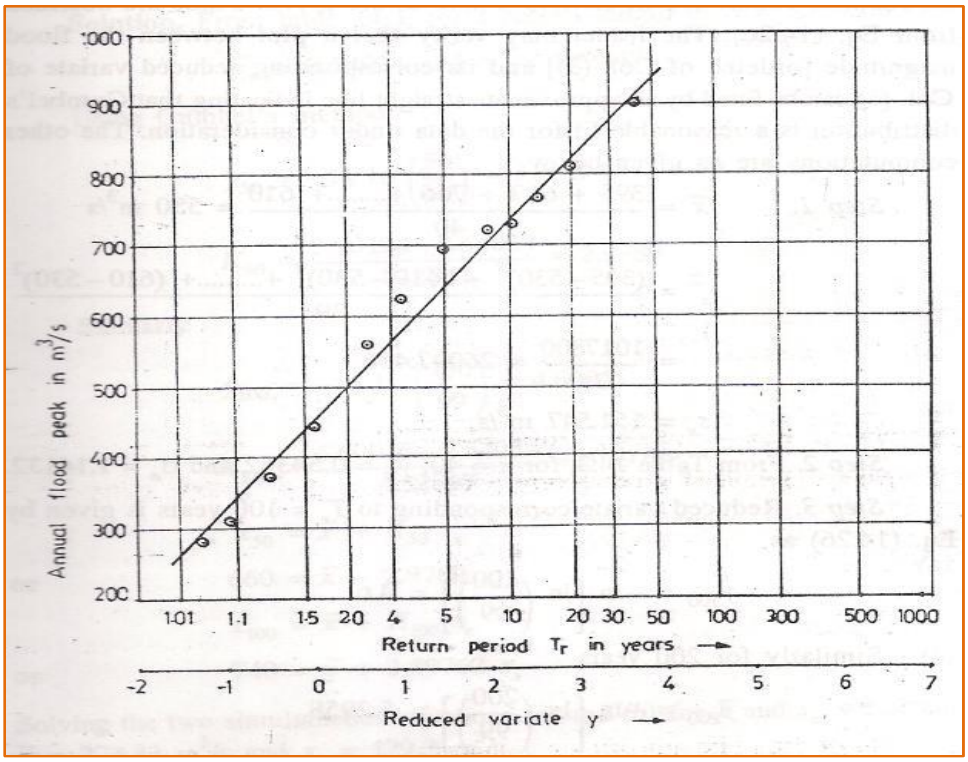
Extreme Value probability paper

A probability paper which plots Gumbel's extreme value distribution as a straight line is constructed as follows:

If any random variable x follows Gumbel's extreme value distribution, then from equation $y = a(x - x_f)$ *the relationship between the reduced variate and variable x is linear and a* plot of y against x on an ordinary graph paper will be a straight line.

Now the scale of the axis of the reduced variate y is transformed into corresponding return period using the relationship

$y = -InIn(\frac{T_r}{T_r - 1})$. Thus for example, points corresponding to y = 0.37, 1.5, 2.25,3.2, 3.9 and 4.6 on the axis of reduced variate are marked as 2, 5, 10, 25, 50 and 100 years respectively. The other axis representing flood magnitude however, remains on an ordinary scale. The resulting probability paper (below) is also known as *extreme value probability paper*. Also shown on the figure are the points obtained by plotting flood peak (column ***3)*** in the previous example against return period (column ***5).*** These points indicate a linear trend and a straight line can be fitted by eye through these points.



Extreme value probability paper

Example 2

Assuming data in **example 1** fits Gumbel's distribution, estimate the 100 and 200 year floods.

Solution

Calculate the reduced variate for the given data using the equation $y_T = -In In(\frac{T_r}{T_r - 1})$

- Calculate the mean $\bar{x}$ (530) and standard deviation S_x (161.547) using relevant equations
- From **table 1** for n = 40 years, $\bar{y}_n$ = 0.5436 and σ_n = 1.1413
- Reduced variate corresponding to $T_r = 100$ years is then given by $y_T = -In\left[In\left(\frac{T_r}{T_r - 1}\right)\right]$

  i.e. $y_{100} = -In\left[In\left(\frac{100}{99}\right)\right]$ = 4.6 and $y_{200} = -In\left[In\left(\frac{200}{199}\right)\right] = 5.2958$

- From the equation $K_T = \left(\frac{y_T - \bar{y}_n}{\sigma_n}\right)$, $K_{100} = \left(\frac{4.6 - 0.54362}{1.14132}\right) = 3.554$ and for 200 year flood $K_{200} = \left(\frac{5.2958 - 0.54362}{1.14132}\right) = 4.164$
- From equation x_T = $\bar{x} + K_T * S_x$, the floods are computed as follows:

  100-year flood, x_{100} = $530 + 3.554x161.547$= **1,104.14 m³/s**

  200-year flood x_{200} = $530 + 4.164x161.547$= **1,202.68 m³/s**

Example 3

It is known that the annual peak discharges of Ndarugu River follow **GUMBEL'S** extreme value type I distribution. If its mean and standard deviations are 10,000 m³/s and 3,000 m³/s respectively, find the probability that the annual peak discharge will exceed 15,000 m³/s. Also determine the magnitude of a flood with an exceedance probability of 0.1.

$\mu = \text{Mean}$ $\quad \sigma$ standard deviation, $\alpha = a$ and $\beta = x_f$

Solution

μ = 10,000 and σ = 3,000, $\alpha = \dfrac{1.28255}{\sigma} = \dfrac{1.28255}{3000} = \dfrac{1}{2329}$

$\beta = \mu - 0.45005\sigma$ = 10000-0.45005x3000 = 8650

- When x = 15,000, y = $\alpha(x-\beta) = \dfrac{1}{2329}$ (15000 - 8650) = 2.7148

F(x) = P [X ≤] = $e^{-e^{-y}}$ = $e^{-e^{-2.7148}}$ = 0.9359

P[X>15,000] = 1-F(x) = 1- 0.9359 = 0.0641

Probability that the annual peak discharge in any year will be more than 15,000 m^3/s is ***<u>0.0641</u>***

- Exceedance probability = 0.1

  Therefore 1-F (x) = 0.1

  F(x) = 0.9 = $e^{-e^{-y}}$

Taking logarithms to natural base twice, y = - in [- in (0.9)] = 2.25037 hence $y = \dfrac{1}{2329}(x - 8650)$

x = 2329y + 8650 = 2329 * 2.25037 + 8650 = 13,913.6 m^3/s

The magnitude of annual peak discharge exceeded in any year with a probability of ***<u>0.1</u>*** is 13,913 m^3/s

Assignment

From the data given in table XX for Ndarugu sub-catchment, compute the 50 and 100 year floods using Gumbel's extreme value and Log-Pearson type III distributions and compare the results

Table xx:

Year	1950	51	52	53	54	55	56	57	58	59	60	61	62
Runoff (mm)	113	94.5	76.0	87.5	92.7	71.3	77.3	85.1	122.8	69.4	81.0	94.5	86.3
Year	63	64	65	66	67	68	69	70	71	72	73	74	75
Runoff (mm)	68.6	82.5	90.7	99.8	74.4	66.6	65.0	91.0	106.8	102.2	87.0	84.0	

Solution: The following are obtained for the series: Mean = 86.8mm, standard deviation =14.66mm.

Gumbel's distribution: K_{50}= 3.088 and K_{100}= 3.729 (from table zz)
Therefore, X_{50} = 86.8+(3.088x14.66) = ***132.0mm***
and X_{100} = 86.8+(3.729x14.66) = ***141.5mm***

Log Pearson Type-III Distribution:
For log transferred series: Mean = 1.9327, standard deviation = 0.072, coefficient of variation = 0.0372 and coefficient of skewness = 0.214
Frequency factors from table 2 are K_{50} = 2.167 and K_{100}= 2.485 (interpolation between values necessary)
y_{50} = 1.9327+(2.167x0.07198) = 2.0887 $\Rightarrow X_{50} = 10^{2.08872}$ = **122.6 mm**
y_{100} = 1.9327+(2.485X0.07189) = 2.11139 $\Rightarrow X_{100} = 10^{2.11139}$ = ***129.24 mm***

From the above Gumbel's approach gives higher values.

In Kenya the recommended one Nationally is Gumbel's extreme Value type I distribution

Table ZZ: Frequency factors for Gumbel's extreme value distribution

Sample size	*Return priod (T) in years*										
n	*5*	*10*	*15*	*20*	*25*	*30*	*50*	*60*	*75*	*100*	*1,000*
15	0.967	1.703	2.117	2.410	2.632	2.823	3.321	3.501	3.721	4.005	6.265
20	0.919	1.625	2.023	2.302	2.517	2.690	3.179	3.352	3.563	3.836	6.006
25	0.888	1.575	1.963	2.235	2.444	2.614	3.088	3.257	3.463	3.729	5.842
30	0.866	1.541	1.922	2.18	2.393	2.560	3.026	3.191	3.393	3.653	5.727
35	0.851	1.516	1.891	2.152	2.354	2.520	2.979	3.142	3.341	3.598	
40	0.838	1.495	1.866	2.126	2.326	2.489	2.943	3.104	3.301	3.554	5.576
45	0.829	1.478	1.847	2.104	2.303	2.464	2.913	3.078	3.268	3.520	
50	0.820	1.466	1.831	2.086	2.283	2.443	2.889	3.027	3.241	3.491	5.478
55	0.813	1.455	1.818	2.071	2.267	2.426	2.869	3.027	3.219	3.467	
60	0.807	1.455	1.818	2.071	2.267	2.426	2.869	3.008	3.219	3.467	
65	0.801	1.446	1.806	2.059	2.253	2.411	2.852	2.992	3.200	3.446	
70	0.797	1.437	1.796	2.048	2.241	2.398	2.837	2.979	3.183	3.429	
75	0.792	1.430	1.788	2.038	2.230	2.387	2.824	2.9.67	3.169	3.413	5.359
80	0.788	1.423	1.780	2.029	2.220	2.377	2.812	2.956	3.155	3.400	
85	0.785	1.413	1.767	2.020	2.212	2.368	2.802	2.946	3.145	3.387	
90	0.782	1.409	1.762	2.007	2.205	2.353	2.785	2.938	3.125	3.367	
95	0.780	1.405	1.757	2.002	2.193	2.347	2.777	2.930	3.116	3.357	
100	0.779	1.401	1.172	1.998	2.187	2.341	2.770	2.922	3.109	3.349	5.261

Table 1: Mean $\overline{y_n}$ and standard deviation S_n of reduced extremes of Gumbel's Distribution.

Size of sample n	$\overline{y}_n$	σ_n
10	0.4952	0.2457
15	0.5128	1.0206
20	0.5236	0.0628
25	0.5309	1.6915
30	0.5362	1.1124
35	0.5403	1.1283
40	0.5436	1.1413
45	0.5436	1.1518
50	0.5465	1.1607
55	0.5504	1.1681
60	0.5521	1.1747
65	0.5536	1.1803
70	0.5548	1.1854
75	0.5549	1.1898

80	0.5569	1.1938
85	0.5578	1.1973
90	0.5539	1.2007
95	0.5553	1.2038
100	0.5600	1.2065
200	0.5672	1.2359
500	0.5724	1.2588
1000	0.5745	1.2685

Example:

From analysis of available data on annual flood peaks of a small stream for period of 35 years, the 50 and 100 year floods were estimated at 600 and 740 m^3/s respectively using Gumbel's approach. From this information, estimate the 200-year flood.

Solution:

From table 1, for a record length of n = 35 years

$\overline{y_n}$ = 0.5403 and σ_n = 1.1283

Using Gumbel's method

$y_{50} = \dfrac{3.90194 - 0.54034}{1.12847}$ **=2.9789**

Similarly y_{100} = -In In (100/99) = 4.6002

$K_{100} = \dfrac{4.60015 - 0.54034}{1.12847}$ **= 3.5976**

$x_{50} = \bar{x} + K_{50} * S_x$ **or 660 =** $\bar{x} + \mathbf{2.9789} * S_x$ **-- (a)**

$x_{100} = \bar{x} + K_{100} * S_x$

740 = $\bar{x} + \mathbf{3.5976} * S_x$ **-- (b)**

Solving the two simultaneous equations (a) and (b) for $\bar{x}$ and S_x, we obtain:

$\bar{x}$ = 274.83 m^3/s and S_x = 129.3 m^3/s.

Now y_{200} **= -In In** $\left(\dfrac{200}{199}\right)$ **= 5.2958**

$K_{200} = \dfrac{5.29581 - 0.54034}{1.12847}$ **= 4.2141**

$x_{200} = \bar{x} + K_{200} * S_x$ = 274.83+ 4.2141*129.3 = <u>819.71 m^3/s.</u>

The 200-year flood for the stream is 820 m^3/s

Log-Pearson Type III Distribution

Although it has no theoretical basis, this distribution is widely used and it fits well to flood peaks in many cases. The procedure adopted to arrive at the flood discharge of any ***given return*** period is as follows:

1) The given flood discharge data series is converted into their logarithms

$$y_i = \log_{10}(x_i) \qquad (1)$$

2) The mean, standard deviation and skewness coefficient of the y series are then estimated using the following equations.

$$\bar{y} = \frac{1}{n}\sum y_i \quad . \qquad (2)$$

$$s_y = \left(\frac{\sum (y_i - \bar{y})^2}{(n-1)} \right)^{1/2} \quad (3)$$

$$g = \frac{n \sum \overline{(y_i - y)}^3}{(n-1)(n-2)s_y{}^3} \quad (4)$$

In all the above equations the summations run from 1 to n

3) For the given return period T_r and the estimated skewness coefficient g, the value of K_T is selected from standard tables based on **log-Pearson type III** distribution ***(see table 2).***

4) Knowing the mean, standard deviation and frequency factor, the logarithm of the design flood is computed from

$$y_T = \bar{y} + K_T * s_y \quad (5)$$

5) The design flood itself is then given by taking the anti-log of y_T as below:

$$x_T = antilog(y_T) = 10^{y_T} \quad (6)$$

Example:

The observed annual flood peaks of a stream for a period of 40 years from 1941 to 1980 in m^3/s is as given in below: 395, 619,766, 422, 282,990, 705, 528, 520,436, 697, 624, 496, 589, 598, 359, 686, 726, 527, 310, 408, 721, 814, 459, 440, 632, 343, 634, 464, 373, 289, 371, 522, 342,446, 366, 699, 560, 450 and 610. Determine the 100 year and 200 year floods using the Log-Pearson Type III distribution.

Solution

- The given flood discharges are converted into logarithms through the transformation $y_i = \log x_i$. The y series is then 2.5966, 2.7917 and 2.8842--------------------------------.
- Statistics for the y-series are
  $\bar{y}$ = 2.70507
  s_y = 0.1327
  g = + 0.023
- For T_r = 100 years and g = + 0.023, from given tables K_{100} = 2.343 and K_{200} = 2.5976
- From equation (5), the logarithm of 100-year flood is given by

y_{100} = 2.70507 + 2.343*0.1327 = **3.016** and for y_{200} = 2.70507 + 2.5976*0.1327 = **3.0498**

From equation (6) the 100-year design flood is given by

$x_{100} = (10)^{3.016}$ = **1037.53 m³/s** and

x_{200} = (10) 3.0498 = **1121.50 m³/s**

Selection of return Period

The return period established by frequency analysis indicates only the *average interval* between the occurrence of floods *equal to or greater than a given magnitude.* There is no implication for a T_r year flood to occur at the end of every T_r period. *In the terminology of probability theory,* $\frac{1}{T_r}$ *indicates the probability with which the* T_r *year flood may be equalled or exceeded in any one year*. Therefore, if it is desired to select a design flood which is unlikely to occur during the life time of the structure, it becomes necessary to use a return period *greater than the estimated useful life of the structure. This is very important in engineering designs.*

Table 2: Frequency factor K for use with Log-Pearson Type III Distribution

Co-efficient of skew g	Recurrence interval in years							
	2	5	10	25	50	100	200	1000
	Per cent chance of exceedence							
	50	20	10	4	2	1	0.5	0.1
3.0	−0.396	0.420	1.180	2.278	3.152	4.501	4.970	7.250
2.5	−0.360	0.518	1.250	2.262	3.048	3.845	4.652	6.600
2.2	−0.330	0.574	1.284	2.240	2.970	3.705	4.444	6.200
2.0	−0.307	0.609	1.302	2.219	2.912	3.605	4.298	5.910
1.8	−0.282	0.643	1.318	2.193	2.848	3.499	4.147	5.660
1.6	−0.254	0.675	1.329	2.163	-2.780	3.388	3.990	5.390
1.4	−0.225	0.705	1.337	2.128	2.706	3.271	3.828	5.110
1.2	−0.195	0.732	1.340	2.087	2.626	3.149	3.661	4.820
1.0	−0.164	0.758	1.340	2.043	2.542	3.022	3.489	4.540
0.9	−0.148	0.769	1.339	2.018	2.498	2.957	3.401	4.395
0.8	−0.132	0.780	1.336	1.998	2.453	2.891	3.312	4.250
0.7	−0.116	0.790	1.333	1.967	2.407	2.824	3.223	4.105
0.6	−0.099	0.800	1.328	1.939	2.359	2.755	3.132	3.960
0.5	−0.083	0.808	1.323	1.910	2.311	2.686	3.041	3.815
0.4	−0.066	0.816	1.317	1.880	2.261	2.615	2.949	3.670
0.3	−0.050	0.824	1.309	1.849	2.211	2.544	2.856	3.525
0.2	−0.033	0.830	1.301	1.818	2.159	2.472	2.763	3.380
0.1	−0.017	0.836	1.292	1.785	2.107	2.400	2.670	3.235
0	0	0.842	1.282	1.751	2.054	2.326	2.576	3.090
−0.1	0.017	0.836	1.270	1.716	2.000	2.252	2.482	2.950
−0.2	0.033	0.850	1.258	1.680	1.945	2.178	2.288	2.810
−0.3	0.050	0.853	1.245	1.643	1.890	2.104	2.394	2.675
−0.4	0.066	0.855	1.231	1.606	1.834	2.029	2.201	2.540
−0.5	0.083	0.856	1.216	1.567	1.777	1.955	2.108	2.400
−0.6	0.099	0.857	1.200	1.528	1.720	1.880	2.016	2.275
−0.7	0.116	0.857	1.183	1.488	1.663	1.806	1.926	2.150
−0.8	0.132	0.856	1.166	1.448	1.606	1.733	1.837	2.035
−0.9	0.148	0.854	1.147	1.407	1.549	1.660	1.749	1.910
−1.0	0.164	0.852	1.128	1.366	1.492	1.588	1.664	1.800
−1.2	0.195	0.844	1.086	1.282	1.379	1.449	1.501	1.625
−1.4	0.225	0.832	1.041	1.198	1.270	1.318	1.351	1.465
−1.6	0.254	0.817	0.994	1.116	1.166	1.197	1.216	1.280
−1.8	0.282	0.799	0.945	1.035	1.069	1.087	1.097	1.130
−2.0	0.307	0.777	0.895	0.959	1.980	0.990	0.995	1.000
−2.2	0.330	0.752	0.844	0.888	0.900	0.905	0.907	0.910
−2.5	0.360	0.711	0.771	0.793	0.798	0.799	0.800	0.802
−3.0	0.396	0.636	0.660	0.666	0.666	0.667	0.667	0.668

ENCOUNTER PROBABILITY

However, even if a flood of a long recurrence interval is chosen, there is always a possibility that the flood can be exceeded more than once during the interval. The probability of '*r*' events occurring in '*N*' possible events is given by:

$$P_{(N,r)} = \frac{N!}{r!(N-r)!} P^r (1-P)^{N-r}$$

Where p = probability of a single event.

If r = 0, the flood will not be exceeded during "N" years, the useful life of the structure. The above equation then becomes:

Probability of non-exceedance $P_{(N,0)} = (1-P)^N$.

Suppose a structure is designed for T_r year flood and its useful life period is N years. The probability that the design flood is equalled or exceeded (***and hence the probability the structure will fail)*** in any year is given by $\frac{1}{T_r}$. Alternatively, probability that the structure does not fail in any year is given by $\left(1-\frac{1}{T_r}\right)$. Assuming that the annual flood peaks are independent events, the probability that the structure ***does not fail in the next N years*** is given by $\left(1-\frac{1}{T_r}\right)^N$. Hence the probability that the structure may fail in any one of the next N years which is the risk in the design is given by:

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

The equation is used to evaluate the risk involved in adopting a T_r year flood for a structure with a useful life of N years. Alternatively, it can be used to determine the required return period of the design flood for a given risk R and given life span of N years. For example, if a structure with a life span of 50 years is designed for a 50 year flood; the risk of failure from the equation $R = 1-\left(1-\frac{1}{T_r}\right)^N$ is 0.636. In other words, its chances of failure are very high at 63.6% and not 2% as usually thought. If there is need to reduce the risk to 0.10, the structure should be designed for a 475-year flood.

Example:

Analysis of a 30-year flood data at a point on river Ndarugu gave $\bar{x}$ **= 1200m³/s** and s_x = 650 m³/s. For what discharge are you required to design a structure (e.g. bridge) at this point in order to provide 95% assurance that the structure would not fail in the next 50 years?

Solution:

Assurance required = 95%

Therefore, allowable risk R = 100 - 95 = 5%

Life period = 50 years

From the risk equation $R = 1-\left(1-\frac{1}{T_r}\right)^N \Rightarrow 0.05 = 1-\left(1-\frac{1}{T_r}\right)^{50}$ = $1-\left(1-\frac{1}{T_r}\right)^{50}$ =

1- 0.05 = 0.095 i.e. $1-\frac{1}{T_r}$ = $(0.95)^{1/50}$ = 0.99897

$$\frac{1}{T_r} = 0.0010253 \Rightarrow T_r = \frac{1}{0.0010253} \text{ hence } T_r = 975.3 \text{ years}$$

Gumble'ls method can now be used to estimate the design flood with a return period of 975.3 years.

From tables for n = 30, we get $\bar{y}_n$ = 0.53622 and $\sigma_n = 1.11238$

From equation $y_T = -In\left[In\left(\frac{T_r-1}{T_r}\right)\right]$ y_T = -In In $\left(\frac{975.3}{974.3}\right)$ = 6.88223 and

$$K_T = \frac{6.88223-0.53622}{1.11238} = 5.705$$

The design flood is then given by 1200 + 5.705*650 = 4908.25 m³/s

The structure has therefore to be designed for a discharge of 4,909 m³/s

Example

A bridge has an expected life span of 25 years and is designed for a flood magnitude with a return period of 100 years. Determine the:

i. The Risk of the hydrologic design
ii. If a 10% risk is permissible, what return period should be adopted?

Solution

i. The risk $\overline{R} = \left[1-\frac{1}{T}\right]^n$. Hence n = 25 years and T = 100 years

$$\overline{R} = \left[1-\frac{1}{T}\right]^{25} = 0.222.$$ The inbuilt risk in this design is therefore 22.2 %

ii. If $\overline{R}$ = 10% = 0.10

$$0.10 = 1 - \left[1-\frac{1}{T}\right]^{25} \text{ hence } \left[1-\frac{1}{T}\right]^{25} = 0.90 \text{ and T = 238 years (say 240 years)}$$

If the design flood used has a return period of 100-years, the risk is 22.2%. For a risk of 10% you need to use a design flood with a return period of 240 years.

Low-Flow Frequency Analysis

Whereas high flows lead to floods, sustained low flows can lead to droughts. *A drought is defined as a lack of rainfall so great and continuing for so long as to affect the plant and animal life of a region adversely and to deplete domestic and industrial water supplies, especially in those regions where rainfall is normally sufficient for such purposes.*

In practice, a drought refers to a period of *unusually low water supplies, regardless of the water demand*. Regions most subject to droughts are those with the greatest variability in annual rainfall. Studies have shown that regions where the variance coefficient of annual rainfall exceeds 0.35 are more likely to have frequent droughts. Low annual rainfall and high annual rainfall variability are typical of *arid and semi-arid regions*. Therefore, these regions are more likely to be prone to droughts.

Studies of tree rings, which document long term trends of rainfall, show clear patterns of periods of wet and dry weather. While there is no apparent explanation for the cycles of wet and dry weather, the dry years must be considered in planning water resource projects. *Analysis of long records has shown that there is a tendency for dry years to group together.*

This indicates that the sequence of dry years is not random, with dry years tending to follow other dry years. It is therefore necessary to consider both these *severity and duration* of a drought period. The severity of droughts can be established by measuring (1) the deficiency in rainfall and runoff, (2) the decline of soil moisture, and (3) the decrease in groundwater levels. Alternatively, low-flow-frequency analysis can be used in the assessment of the probability of occurrence of droughts of different durations.

Methods of low-flow frequency analysis are based on an assumption of invariance of meteorological conditions. The absence of long records, however, imposes a stringent limitation on low-flow frequency analysis. When records of sufficient length are available, analysis begins with the identification of the low-flow series. Either the annual minima or the annual exceedance series are used. In a monthly analysis, the annual minima series is formed by the lowest monthly flow volumes in each year of record. If the annual exceedance method is chosen, the lowest monthly flow volumes in the record are selected, regardless of when they occurred. In the latter method, the number of values in the series need not be equal to the number of years of record.

A flow-duration curve can be used to give an indication of the severity of low flows. Such a curve, however, does not contain information on the sequence of low flows or the duration of possible droughts. The analysis is made more meaningful by abstracting the minimum flows over a period of several consecutive days. For instance, for each year, the 7- day period with minimum flow volume is abstracted, and the minimum flow is the average flow rate for that period. A frequency analysis on the low-flow series, using the Gumbel method, for instance, results in a function describing the probability of occurrence of low flows of certain duration.