CHAPTER III

THEORETICAL CONSIDERATION

3.1 Relationship between Runoff and Catchment Area

In general runoff is a function of many factors such as intensity, duration of precipitation, catchment area, vegetation, geology, and slope of catchment area, etc., so

$$Q = (A, R, S, T, ...)$$

where
- Q = runoff
- A = catchment area
- R = rainfall
- S = land slope
- T = return period in case of flood and drought

However, the exponential forms of this function are popularly expressed as shown below

1. $Q = K_1 A^{n_1}$ ....(1)
2. $Q = K_2 A^{n_1} R^{n_2}$ ....(2)
3. $Q = K_3 A^{n_1} R^{n_2} S^{n_3}$..........(3)

The above three equations can be applied for average annual flood, average annual runoff, average annual drought , and average monthly drought; so

Q = average annual flood, average annual runoff , average annual and monthly droughts in cms.

A	=	catchment area in sq.km.
R	=	average maximum annual rainfall, average annual rainfall, average minimum annual and monthly rainfall of catchment area in m.m.
S	=	average slope of catchment area
n_1, n_2, n_3	=	constants derived by fitting regression line to the available data
K_1, K_2, K_3	=	constants for a particular area depended on soil, rainfall, slope, etc., derived by fitting regression line to the available data

3.2 Relationship between Flood Magnitude and Return Period

In general only annual floods are concerned, so the data for the analysis are the maximum flood peaks of each calendar or water year. The methods for finding this relation are as follows,

1. By Plotting Positions formulas

For n years record, there are n number s of the flood peak (one value from one year only). By arranging in descending order of the flood magnitude and assigning the order number m, the return period of each flood magnitude of the order number m can be found as follows,

1. California formula (1923)

$$T = \frac{n}{m} \qquad \text{............} \qquad (4)$$

2. Harzen formula (1930)

$$T = \frac{2n}{2m-1} \qquad \text{............} \qquad (5)$$

3. Weibull formula (1939)

$$T = \frac{n + 1}{m} \quad \ldots\ldots\ldots\ldots\ldots \quad (6)$$

4. Beard formula (1943)

$$T = \frac{1}{1 - 0.5^{1/n}} \quad \ldots\ldots\ldots\ldots\ldots \quad (7)$$ [1]

5. Chegodayev formula (1955)

$$T = \frac{n + 0.4}{m - 0.3} \quad \ldots\ldots\ldots\ldots\ldots \quad (8)$$

6. Blom formula (1958)

$$T = \frac{n + 0.25}{m - 0.375} \quad \ldots\ldots\ldots\ldots\ldots \quad (9)$$

7. Tu key formula (1962)

$$T = \frac{3n + 1}{3m - 1} \quad \ldots\ldots\ldots\ldots\ldots \quad (10)$$

8. Gringorten formula (1963)

$$T = \frac{n + 0.12}{m - 0.44} \quad \ldots\ldots\ldots\ldots\ldots \quad (11)$$

---

1 this formula applied only to m = 1

where	T	=	return period in years
	n	=	the number of years of record
	m	=	rank of event or flood ($m = 1$ for max. and $m = n$ for min. event or flood)

From one of equation (4) to (11) the return period of each flood can be calculated and plotted against each flood magnitude on a probability paper. After plotting, a curve may be fitted to the plotted points, and this curve will show the relation between T and Q_T.

2. <u>By Gumbel's formula</u>

If $X_1, X_2, X_3 \ldots\ldots\ldots X_n$ are the extreme values observed in n sample of equal size N and if X is an unlimited exponentially distributed variable, from the theory of extreme values, as n and N approach infinity, the cumulative probability (P_X) of an extreme being equal to or less than X is represented by,

$$P_X = e^{-e^{-y}} \qquad \ldots\ldots\ldots(12)$$

where	e	=	the base of Napierian logarithms,
	y	=	the reduced variate
		=	$a(X - X_f)$,
and	a, X_f	=	statistical parameters which are functions of the arithmetic mean, $\overline{X}$, and the standard deviation, σ_X

Also, the theory of extreme values for an infinitely large sample gives

X_f = the mode of distribution

= $\overline{X} - 0.45005\, \sigma_x$ ....(13)

a = the disper sion parameter

= $\dfrac{1.28255}{\sigma_x} > 0$ ....(14)

In the case of annual floods, $X_1, X_2, X_3 \ldots\ldots\ldots X_n$ are maximum flows for each water year (N = 1 year).

From equation (12) the return period, T_x (or the recurrence interval) in years of which a flood flow is equal to or greater than X is expressed by

$$T_x = \frac{1}{1 - P_x} \qquad \text{..........(15)}$$

Equation (12) and (15) yield

$$X = X_f - \frac{1}{a} \ln \left[-\ln \left(1 - \frac{1}{T_x}\right) \right] \qquad \text{......(16)}$$

Expanding equation (16) leads to

$$X = X_f + \frac{1}{a} \left[\ln T_x - \frac{1}{2T_x} - \frac{5}{24T_x^2} - \frac{1}{8T_x^3} - \right] \text{..(17)}$$

For $T_x > 20$ years, the terms after the first term $\ln T_x$ in bracket can be neglected with an error of less than 0.7 percent. The equation (17), then, becomes

$$X = X_f + \frac{1}{a} \ln T_x \qquad \text{.........(18)}$$

From equation (12), (13), (14), (15), (16) and (18) the return period, T_x of the observed flood, the magnitude of an annual flood at any return period can be calculated. Moreover, these equations can be applied to calculate the monthly flood by replacing the maximum annual flow with the maximum monthly flow.

3.3 Flood flow in terms of Return-Period and Catchment Area

From equation (18) for the flood magnitude,

$$X = X_f\left(1 + \frac{1}{ax_f} \ln T_x\right),$$

it is the reasonable to find the maximum flood flow from Fuller's formula,

$$Q_T = \overline{Q}_a (1 + C \ln T_x) \quad \ldots\ldots\ldots\ldots (19)$$

where Q_T = flood flow, in cms. or cfs. at return period T_x

$\overline{Q}_a$ = the average annual flow in cms. of cfs.,

T_x = return period in years

C = a coefficient

Equation (19), in view of Equation (1), yields

$$Q_T = K_1 (1 + C \ln T_x) A^{n_1} \quad \ldots\ldots\ldots (20)$$

3.4 Relationship between Drought Magnitude and Return Period

For annual droughts

From n-year records of flows, the minimum flows of each year are used for frequency analysis by the methods as follows:-

1. By Plotting Positions formulas

The basis of this method is the same as in the case of annual flood, but in this case the n values of minimum flow data are used instead of maximum ones. Giving m = 1 for the minimum event, m = n for the maximum event. Then, equation (4) to equation (11) can be used for finding the return period.

2. By Gumbel's formula

If X_1, X_2, X_3 ....X_n are minimum flows for of each water year of n-year records. The probability P_{x_1} of the minimum flow (smallest value) being equal to or larger than X is calculated by,

$$P_{X_1} = e^{-e^{y}} \qquad \text{.............(21)}$$

where e = the base of Napierian logarithms

y = reduced variable

= $a\,(X - X_f)$

and X_f = the mode of distribution

$$= \overline{X} - 0.45005\sigma_x \qquad \text{.....(22)}$$

I16668911

a = the dispersion parameter

$$= \frac{1.28255}{\sigma_x} > 0 \qquad \text{...........(23)}$$

From (21) the return period T_{x_1}, in years, of which the flow or drought is equal to or less than X is expressed by,

$$T_{x_1} = \frac{1}{1 - P_{x_1}} \qquad \text{............(24)}$$

Equation (21) and (24) yield

$$X = X_f + \frac{1}{a} \ln \left[-\ln \left(1 - \frac{1}{T_{x_1}}\right) \right] \qquad \text{......(25)}$$

Expanding equation (25) leads to

$$X = X_f - \frac{1}{a} \left[\ln T_{x_1} - \frac{1}{2T_{x_1}} - \frac{5}{24T_{x_1}^2} - \frac{1}{8T_{x_1}^3} \right] \qquad \text{......(26)}$$

for $T_{x_1} > 20$, the terms after first terms $\ln T_{x_1}$ in bracket can be neglected. Therefore, the equation (26) becomes,

$$X = X_f - \frac{1}{a} \ln T_{x_1} \qquad \text{........(27)}$$

Equation (21), (22), (23), (24), (25) and (27) are used to find the return period T_{x_1} of the observed drought and the magnitude of an annual drought at any return period, T_{x_1}.

For Monthly Droughts

Equations (4) to (11) and (21) to (27) for this case can be applied and minimum annual flow datas replaced by minimum monthly flow datas.

3.5 Correlation

In general, in a basin area there are always many stations which have incompleted records of stream flows but the values for the missing years can be estimated by correlating the flow values of the incomplete records with the complete records of another station.

Moreover, the correlation method can be used to determine the predicted flow at an ungaged station (Incomplete records) from the recorded flow at the gaged station (complete records) if the correlation between these two stations are well. This means that if two stations show good correlation. It is possible to estimate the flow of one station from the record of the other stations.

In general, the correlation may be expressed by an arbitary analytic curve. In many cases, however, it is advantageous and time-saving if it can be expressed by a straight line of a linear correlation or regression.

First, let the two series of observed discharges at stations X and Y are

$$X_1, X_1, X_3, \ldots\ldots\ldots, X_i, \ldots\ldots\ldots\ldots, X_N$$

$$Y_1, Y_2, Y_3, \ldots\ldots\ldots\ldots Y_i, \ldots\ldots\ldots, Y_N$$

in which X_i and Y_i are correspond for each i.

Second, assume that there is a linear correlation between those corresponding points of (X_i, Y_i) then the linear regression analysis can be used.

The linear regression of the relation between X_i and Y_i are expressed by the equations of two least square regression lines as follows:-

1 $$Y = a_0 + a_1 X \qquad \ldots\ldots\ldots\ldots(28)$$

where a_0 and a_1 are obtained from the normal equations

$$\sum_{i=1}^{N} Y_i = a_0 N + a_1 \sum_{i=1}^{N} X_i$$

$$\sum_{i=1}^{N} X_i Y_i = a_0 \sum_{i=1}^{N} X_i + a_1 \sum_{i=1}^{N} X_i^2$$

hence
$$\left.\begin{aligned} a_0 &= \frac{\sum_{i=1}^{N} Y_i \sum_{i=1}^{N} X_i^2 - \sum_{i=1}^{N} X_i \sum_{i=1}^{N} X_i Y_i}{N\sum_{i=1}^{N} X_i^2 - \left(\sum_{i=1}^{N} X_i\right)^2} \\ a_1 &= \frac{N\sum_{i=1}^{N} X_i Y_i - \sum_{i=1}^{N} X_i \sum_{i=1}^{N} Y_i}{N\sum_{i=1}^{N} X_i^2 - \left(\sum_{i=1}^{N} X_i\right)^2} \end{aligned}\right\} \quad \ldots(29)$$

2 $$X_i = b_0 + b_1 Y_i \qquad \ldots\ldots(30)$$

where b_0 and b_1 are also obtained from the normal equations

$$\sum_{i=1}^{N} X_i = b_0 N + b_1 \sum_{i=1}^{N} Y_i$$

$$\sum_{i=1}^{N} X_i Y_i = b_0 \sum_{i=1}^{N} Y_i + b_1 \sum_{i=1}^{N} Y_i^2$$

hence

$$\left.\begin{aligned} b_0 &= \frac{\sum_{i=1}^{N} X_i \sum_{i=1}^{N} Y_i^2 - \sum_{i=1}^{N} Y_i \sum_{i=1}^{N} X_i Y_i}{N \sum_{i=1}^{N} Y_i^2 - \left(\sum_{i=1}^{N} Y_i\right)^2} \\ b_1 &= \frac{N \sum_{i=1}^{N} X_i Y_i - \sum_{i=1}^{N} X_i \sum_{i=1}^{N} Y_i}{N \sum_{i=1}^{N} Y_i^2 - \left(\sum_{i=1}^{N} Y_i\right)^2} \end{aligned}\right\} \quad \ldots\ (31)$$

The correlation coefficient $r = \pm\sqrt{a_1 b_1}$

so, from equation (29) and (31)

$$r = \sqrt{\frac{\left(N \sum_{i=1}^{N} X_i Y_i - \sum_{i=1}^{N} X_i \sum_{i=1}^{N} Y_i\right)\left(N \sum_{i=1}^{N} X_i Y_i - \sum_{i=1}^{N} X_i \sum_{i=1}^{N} Y_i\right)}{\left\{N \sum_{i=1}^{N} X_i^2 - \left(\sum_{i=1}^{N} X_i\right)^2\right\}\left\{N \sum_{i=1}^{N} Y_i^2 - \left(\sum_{i=1}^{N} Y_i\right)^2\right\}}}$$

$$= \frac{N \sum_{i=1}^{N} X_i Y_i - \sum_{i=1}^{N} X_i \sum_{i=1}^{N} Y_i}{\sqrt{\left\{N \sum_{i=1}^{N} X_i^2 - \left(\sum_{i=1}^{N} X_i\right)^2\right\}\left\{N \sum_{i=1}^{N} Y_i^2 - \left(\sum_{i=1}^{N} Y_i\right)^2\right\}}} \quad \ldots\ldots (32)$$

and equation (32) can be reduced to simple form

$$r = \frac{\overline{XY} - \bar{X}\bar{Y}}{\sigma_X \cdot \sigma_Y} \qquad \ldots \qquad (33)$$

where

$$\overline{XY} = \frac{1}{N}\sum_{i=1}^{N} X_i Y_i$$

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i$$

$$\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i$$

$$\sigma_X = \sqrt{\frac{\sum_{i=1}^{N}(X_i - \bar{X})^2}{N}} \qquad \text{for } N > 30$$

$$= \sqrt{\frac{\sum_{i=1}^{N}(X_i - \bar{X})^2}{N-1}} \qquad \text{for } N < 30$$

$$\sigma_Y = \sqrt{\frac{\sum_{i=1}^{N}(Y_i - \bar{Y})^2}{N}} \qquad \text{for } N > 30$$

$$= \sqrt{\frac{\sum_{i=1}^{N}(Y_i - \bar{Y})^2}{N-1}} \qquad \text{for } N < 30$$

N = the total number of elements in the series

The correlation coefficient varies between -1 and +1 and is equal to 1 if both lines are identical and merge. In hydrology, the following relation between flows may be ascertained on the basis of the values of the coefficient of correlation

$r = 1$ direct functional dependence,

$0.6 < r < 1$	good direct correlation,
$0 < r < 0.6$	insufficient direct correlation,
$r = 0$	no correlation,
$-0.6 < r < 0$	insufficient reciprocal correlation,
$-1 < r < 0.6$	good reciprocal correlation,
$r = -1$	reciprocal linear functional dependence