CHAPTER - VII

WATER LEVEL ANALYSIS - EXCEEDENCES

7.1 Introduction

The common approach to the problem of flood prediction is mainly motivated by operational objectives and therefore the selection of a flood frequency distribution function is based on a criterion of 'best curve fit' to the observed largest flood discharge values at a lapse of fixed time. Distribution functions that are most commonly used include log-normal, log Pearson type 3 , two parameter gamma and Gumbel's extreme value distribution. Application of the first three distributions to the flood frequency analysis does not seem to be appropriate because of the assumptions involved. Also, the application of Gumbel's extreme value distribution is based on following two assumptions.

(i) the sequence of daily water discharge values of whole year form a sequence of independent and identically distributed random variables and

(ii) these random variables are assumed to have distributions with an exponential tail.

For the limitations of physical basis of these assumptions we quote Gumbel

" It must be admitted that the good fit cannot be foreseen from the theory, which is based on three assumptions

(i) the distribution of daily discharges is of exponential type

(ii) n = 365 is sufficiently large,

(iii) the daily observations are independent ".

Assumption(i) can not be checked since the analytical form of the distribution of daily water discharge values is unknown. n = 365 is not sufficiently large number for the convergence of a distribution function to the Gumbel extreme value distribution. The third assumption definitely does not hold as the daily observations are in general not independent.

These remarks by Gumbel indicate that the mathematical assumptions underlying the classical extreme value theory may not always be applicable to a water discharge series because it only provides asymptotic expressions. If the time interval of interest is less than a year which is often the case, then the use of an asymptotic expression can hardly be justified. Nonetheless, the fact that Gumbel provides a candid criticism of applying the extreme value theory to flood phenomena is of considerable significance because it points out the need to develop a theory that is physically more meaningful for flood frequency analysis than the asymptotic theory.

The first attempt to develop such a theory from the properties of stream flow (rather than to explain the properties of stream flow from such a theory) is that of Todorovic and his coworkers(1970,1971,1972). Their formulasim is based on the partial duration series of stream flow . The sequence of flows in such a series within fixed time interval is represented by random variables.

Todorovic(1971) used the above formulation along with mathematical assumptions of Todorovic and Zelenhasic(1970) to analyse another important feature of the extreme flood, namely, their time of occurrence within a selected time interval. The expression for the time of occurrence of extreme flood obtained by Todorovic(1971) is exact and tested on data for two rivers in the United States by Todorovic and Woolhiser(1972). It is worth mentioning that Zelenhasic(1970) has shown that the functional form of the distribution function of the largest exceedance derived by above approach, is similar to Gumbel's extreme value distribution.

In sections 7.2 and 7.3 of this chapter the work of Todorovic and his coworkers is reviewed. In section 7.4 the distribution function of number of exceedances in two variable is derived, while section 7.5 is concerned with the joint bivariate distribution function of supremum and infimum of a bivariate sample of magnitude of exceedances. Finally, sections 7.6 and 7.7 are devoted for the application of the results obtained in section 7.3 and 7.5 respectively. This is achieved by expressing the expected number of exceedances in a fixed time interval, for both the variables, by a Fourier expansion. The distribution function of magnitude of exceedances is assumed to be exponential. The dependence parameter of bivariate distribution function of supremum is estimated by relating it to the medial correlation coefficient. The above mentioned results have been derived for any bivariate population but to achieve the goal concerning the study of flood. The results are tested for secondary data. The data taken for 30 years from 1948 to 1977 is data on water discharge on the river Narmada at two stations(Mortakka and Garudeshwar) for a partial duration series. To demonstrate the goodness of fit achieved the theoretical and observed results are also presented graphically.

7.2 Exceedances in one dimension

The section deals with the problem of flood analysis based on the recent development in the theory of extreme values given by Todorovic(1970). In his approach Todorovic utilizes a stochastic model to describe and predict the behaviour of floods. He starts with a stochastic process X(t) defined as the highest magnitude of random variables in an interval of time (0 , t] . Since the number of flood peak discharges in (0 , t] exceeding a certain level x_0 and the magnitudes of these peaks are random variables, the foregoing model seems to confirm well to the flood phenomenon. An extension of this approach has been made to more than one variable which we shall describe in later parts of this chapter.

Let Q_i's be the peak values that exceeds a suitably choosen base level x_0 and $\xi_i = Q_i - x_0$. We call ξ_i as the magnitude of the i-th exceedance.

If the sequence of these peak discharge values is multiple peaked i.e. if the daily discharges exceed the fixed base discharge level x_0 for more than one day continuously without being less than x_0, then the practical aspects suggest that the entire peak discharges should not be conside-

-red but only the maximum peak discharge is taken as an exceedance.

According to Todorovic(1970) the distribution function of exceedances and their time of occurrence is derived as follows

We start reckoning the event of exceedance at a time designated as $t=0$ and define $\tau(k)$ as the time elapsed before the k-th exceedance occurs and ξ_k as its magnitude. We shall take $\tau(0)=0$ and $\xi_0=0$. Let $X_k = \sum_{\rho=1}^{k} \xi_\rho$ be the sum of the magnitude of k exceedances.

The time $\tau(k)$ satisfies the conditions

$$0 \leq \tau(k) \leq \tau(k+1), \quad k = 1,2, \ldots$$

and (7.2.1)

$$\lim_{k\to\infty} \tau(k) = \infty .$$

Assuming that the number of points $\tau(k)$ in the fixed time interval $(0,t]$ are the chance variables, then $\tau(k)$ are random variables as well. In addition it is also supposed that the $\tau(k)$ are continuous random variables.

In what follows we consider the distribution function of $\tau(k)$ and X_k .

Let two events E_k^t and G_k^x be defined as

$$E_k^t = [\tau(k) \leq t \leq \tau(k+1)] \tag{7.2.2}$$

$$G_k^x = [X_k \leq x \leq X_{k+1}] \tag{7.2.3}$$

then for all i and j (i,j=0 , 1,2 , ... ; i≠ j),

$$E_i^t \cap E_j^t = 0 \,, \qquad G_i^x \cap G_j^x = 0$$

and

$$\bigcup_{i=0}^{\infty} E_i^t = \bigcup_{i=0}^{\infty} G_i^x = \Omega$$

Where Ω denotes the whole sample space.

By virtue of (7.2.2) and (7.2.3), it follows that

$$P(E_k^t) = P(\tau(k) \leq t) - P(\tau(k+1) \leq t) \qquad (7.2.4)$$

$$P(G_k^x) = P(X_k \leq x) - P(X_{k+1} \leq x) \qquad (7.2.5)$$

Since $\tau(k)$ and X_k are strictly increasing sequences of random variables it follows that for all $t > 0$, $x > 0$, $P(E_k^t) > 0$ and $P(G_k^x) > 0$.

Finally, from (7.2.4) and (7.2.5), the respective distribution functions of $\tau(k)$ and X_k are seen to be

$$F_k(t) = 1 - \sum_{j=0}^{k-1} P(E_j^t) \qquad (7.2.6)$$

and

$$F_k(x) = 1 - \sum_{j=0}^{k-1} (P(G_j^x) \qquad (7.2.7)$$

. Let us now find an expression for the probabilities $P(E_k^t)$ and $P(G_k^x)$. $P(E_k^t)$ as defined in (7.2.4) is the probability of occurrence of k exceedances in time interval $(0,t]$. Let us denote by $\eta(t)$, the number of exceedances over x_0

in time interval $(0,t)$. Then $\eta(t)$ will assume values $0,1,2 \ldots$ and for all $t \geq 0$, $\Delta t \geq 0$; $\eta(t) = \eta(t+\Delta t)$ i.e. $\eta(t)$ is an nondecreasing function of time.

Todorovic and Zelenhasic(1970) assumed that these exceedances are governed by a Poisson process which has a time dependent intensity function. According to them, under the assumptions of Poisson process, the probabilities $P(E_k^t)$ satisfy the following system of stochastic differential equations.

$$\frac{d\,P(E_k^t)}{dt} = \lambda_{k-1}(t)\,P(E_{k-1}^t) - \lambda_k(t)\,P(E_k^t) \tag{7.2.8}$$

$$\frac{d\,P(E_o^t)}{dt} = -\lambda_o(t)\,P(E_o^t) \tag{7.2.9}$$

Where $\lambda_k(t) = \lim_{\Delta t \to 0} P(E_{k+1}^{t,t+\Delta t} \mid E_k^t)/\Delta t$ (7.2.10)

and

$$E_{k+1}^{t,t+\Delta t} = \eta(t+\Delta t) - \eta(t) = 1. \tag{7.2.11}$$

Equations(7.2.8) and (7.2.9) have the following solutions

$$P(E_o^t) = \exp(-\int_0^t \lambda_o(s)\,ds \tag{7.2.12}$$

and

$$P(E_k^t) = \exp\left[-\int_0^t \lambda_k(s)\,ds\right] \int_0^t \lambda_{k-1}(t)$$

$$\cdot \exp \int_0^t \left[\lambda_k(s) - \lambda_{k-1}(s)\right] ds$$

$$\cdot \int_0^{t_1} \ldots \int_0^{t_{k-1}} \lambda_o(t_k)$$

$$\cdot \exp \int_0^{t_k} \left[\lambda_1(s) - \lambda_o(s)\right] ds \; dt_k\, dt_{k-1} \ldots dt_1$$

(7.2.13)

In general a closed expression for $P(E_k^t)$ in terms of $\lambda_k(t)$ is not possible, however, we believe that in the relevant cases of flood analysis, the intensity function is independent of k , i.e. , we can take

$$\lambda_k(t) = \lambda(t)$$

Hence (7.2.12) and (7.2.13) can be written respectively as

$$P(E_0^t) = \exp - \int_0^t \lambda(s)\, ds \qquad (7.2.14)$$

$$P(E_k^t) = \frac{\exp\left[-\int_0^t \lambda(s)\, ds\right]\left[\int_0^t \lambda(s)\, ds\right]^k}{k!} \qquad (7.2.15)$$

Similarly, the expression for the probabilities $P(G_k^x)$ can be obtained. The probabilities $P(G_k^x)$ satisfy the system of following stochastic differential equations.

$$\frac{d\, P(G_k^x)}{dx} = \mu_{k-1}(x)\, P(G_{k-1}^x) - \mu_k(x)\, P(G_k^x) \qquad (7.2.16)$$

and

$$\frac{d\, P(G_0^x)}{dx} = - \mu_0(x)\, P(G_0^x) \qquad (7.2.17)$$

where $\mu_k(x) = \lim_{\Delta x \to 0} \frac{P(G_{k+1}^{x,\, x+\Delta x} \;;\; G_k^x)}{\Delta x}$

and

$$G_{k+1}^{x,\, x+\Delta x} = (\ell(x+\Delta x) - \ell(x) = 1\,)$$

$$\ell(x) = \sup\{k : X_k \leq x\} \qquad (7.2.18)$$

In general the closed expression for the probabilities

$P(G_k^x)$ in terms of $\mu_k(x)$ is not possible. Again considering the simplified assumption of $\mu_k(x) = \mu(x)$, the expression for $P(G_k^x)$ and $P(G_o^x)$ are obtained as

$$P(G_o^x) = \exp\left[-\int_0^x \mu(s)\, ds\right] \qquad (7.2.19)$$

and

$$P(G_k^x) = \frac{\exp\left[-\int_0^x \mu(s)\, ds\right]\left[\int_0^x \mu(s)\, ds\right]^k}{k!} \qquad (7.2.20)$$

The distribution function of the time of occurrence of the k-th exceedance can be written from (7.2.6) and (7.2.15) as follows

$$F_k(t) = 1 - \sum_{j=0}^{k-1} \frac{\exp\left[-\int_0^t \lambda(s)\, ds\right]\left[\int_0^t \lambda(s)\, ds\right]^j}{j!} \qquad (7.2.21)$$

with its density function

$$f_k(t) = \frac{\Lambda(t)}{k!}\left[\int_0^t \lambda(s)\, ds\right]^{k-1} \exp\left[-\int_0^t \lambda(s)\, ds\right] \qquad (7.2.22)$$

where $\Lambda(t) = \int_0^t \lambda(s)\, ds$.

Similarly using (7.2.20) and (7.2.7), the expression for distribution function of X_k is given by

$$F_k(x) = 1 - \sum_{j=0}^{k-1} \frac{\exp\left[-\int_0^x \mu(s)\, ds\right]\left[\int_0^x \mu(s)\, ds\right]^j}{j!} . \qquad (7.2.23)$$

Using the expression of $P(E_k^t)$, the distribution functions of $\sup_{\tau(k)\le t} \xi_k$ and $\inf_{\tau(k)\le t} \xi_k$ are derived in the next section.

7.3 Distribution function of $\sup_{\tau(k)\leq t} \xi_k$ and $\inf_{\tau(k)\leq t} \xi_k$

Once the above terminology and initial structure of $P(E_k^t)$ have been developed, then we can easily find the distribution function of $\sup_{\tau(k)\leq t} \xi_k$ and $\inf_{\tau(k)\leq t} \xi_k$ in the following manner.

Let us consider the sequence of magnitude of exceedances

$$\xi_0(=0)\,,\xi_1\,,\xi_2\,,\ \ldots\ldots$$

and $n(t) = \sup\{k;\ t(k) \leq t\}$, $P(\eta(t) = k) = P(E_k^t)$, then as observed by Todorovic(1970), the following theorem gives the distribution function of $\sup_{\tau(k)\leq t} \xi_k$ and $\inf_{\tau(k)\leq t} \xi_k$ under the following assumptions

(i) the sequence $\xi_1, \xi_2, \ldots$ is a sequence of independent random variables with common distribution function $F(x)$, and

(ii) the number of exceedances $\eta(t)$ are independent of their magnitude ξ_k.

<u>Theorem 7.3.1</u> :- The distribution function $F_{t_s}(x)$ and $F_{t_i}(x)$ of $\sup_{\tau(k)\leq t} \xi_k$ and $\inf_{\tau(k)\leq t} \xi_k$ are given by

$$F_{t_s}(x) = \exp[-\Lambda(t)\,(1-F(x))] \qquad (7.3.1)$$

and

$$F_{t_i}(x) = 1 + \exp[\Lambda(t)] - \exp[-\Lambda(t)\,F(x)] \qquad (7.3.2)$$

Proof:- The distribution functions (7.3.1) and (7.3.2) are derived as the mathematical exectation of the following conditional probabilities.

$$P\left[\sup_{\tau(k)\leq t} \xi_k \leq x \Big/ \eta(t)\right] \tag{7.3.3}$$

$$P\left[\inf_{\tau(k)\leq t} \xi_k \leq x \Big/ \eta(t)\right] \tag{7.3.4}$$

The distribution functions can be written as

$$F_{t_s}(x) = E\,P\left[\sup_{\tau(k)=t} \xi_k \leq x \Big/ \eta(t)\right] \tag{7.3.5}$$

$$F_{t_i}(x) = E\,P\left[\inf_{\tau(k)\leq t} \xi_k \leq x \Big/ \eta(t)\right] \tag{7.3.6}$$

Taking expection we have

$$F_{t_s}(x) = \sum_{n=0}^{\infty} P\left[\sup_{0\leq k\leq n} \xi_k \leq x \ \cap\ E_n^t\right] \tag{7.3.7}$$

and

$$F_{t_i}(x) = 1 - \sum_{n=1}^{\infty} P\left[\inf_{0\leq k\leq n} \xi_k \leq x \ \cap\ E_n^t\right] \tag{7.3.8}$$

Under the assumptions mentioned earlier the distribution function (7.3.7) and (7.3.8) can be expressed by

$$F_{t_s}(x) = \sum_{n=0}^{\infty} (\ F(x)\ )^n\ P(E_n^t) \tag{7.3.9}$$

and

$$F_{t_i}(x) = 1 - \sum_{n=1}^{\infty} (\ 1-F(x)\ )^n\ P(E_n^t) \tag{7.3.10}$$

Substituting $P(E_n^t)$ from (7.2.15) in (7.3.9) and (7.3.10) the relevant form of the distribution functions of $\sup_{\tau(k)\leq t} \xi_k$ and $\inf_{\tau(k)\leq t} \xi_k$ are given in the form of following expressions

$$F_{t_s}(x) = \exp[-\Lambda(t)\ (1-F(x))]$$

and

$$F_{t_i}(x) = 1 + \exp[\Lambda(t)] - \exp[-\Lambda(t)\ F(x)].$$

Obviously (7.3.1) and (7.3.2) are nondecreasing functions of x such that

$$F_{t_s}(0) = F_{t_i}(0) = \exp\ (-\Lambda(t)) = P(E_0^t)$$

$$F_{t_s}(\infty) = F_{t_i}(\infty) = 1$$

Hence (7.3.1) and (7.3.2) satisfying all the properties of a distribution function.

If the exceedances are exponentially distributed with distribution function

$$F(x) = 1 - e^{-\alpha x} \qquad \alpha > 0, \quad x \geq 0 \qquad (7.3.11)$$

then the distribution function of $\sup_{\tau(k)\leq t} \xi_k$ is given by

$$F_{t_s}(x) = \exp[-\Lambda(t)\ e^{-\alpha x}] \qquad x \geq 0, \quad \alpha > 0 \qquad (7.3.12)$$

It may be noted that the functional form of the above distribution function is similar to Gumbel's extreme value distribution function. However, as indicated earlier, the Gumbel's distribution is an asymptotic expression for the distribution of the maximum in an independent identically distributed sequence of exponential random variables, whereas the expression (7.3.12) represents an exact (non-asymptotic) expression for the distribution of the maximum in a random sequence (governed by Poisson process) of independent identically distributed exponential random variables. The distribution function (7.2.12) will be used for illustration purposes in the section 7.6 . In the next section we will derive the distribution function of number of exceedances in two dimension.

<u>7.4 Exceedances in two dimensions</u>

Whereas, lot of work has been done in the theory of extreme values for scalar random variables, not much work is reported in literature for two dimensional variables. In problems of applied nature, it is always possible to get information on some other variables which are dependent on the variable of interest without much efforts and cost. Such situations provides opportunity for the study of multidimensional random variables. Therefore, in the sequal the study of extreme order statistics in bivariate distribution is taken up. This happens an extension of univariate distribution function of extremes given by Todorovic and his associates.

Let ($x(t)$, $y(t)$) be an observation in two variables at a time t . The nature of time t for practical reasons will formally be discrete. We denote by $\tau(k)$ and $\tau(\ell)$ the time that elapsed before the k-th and ℓ-th exceedance occured in x and y respectively. We shall take $\tau(0) = 0$ for both the events. The time $\tau(k)$ and $\tau(\ell)$ satisfy the conditions

$$0 \leq \tau(k) \leq \tau(k+1)$$

$$0 \leq \tau(\ell) \leq \tau(\ell+1)$$

We shall denote by ξ_k and θ_ℓ the magnitude of the k-th and ℓ-th exceedances respectively in x and y and also assume $\xi_0 = 0$, $\theta_0 = 0$. The time $\tau(k)$ and $\tau(\ell)$ in $(0, t]$ are random variables. In addition these are continuous and dependent on each other. Here we are interested in finding out the joint distribution function of

$$\left[\inf_{\tau(k) \leq t} \xi_k \; , \; \inf_{\tau(\ell) \leq t} \theta_\ell\right]$$

and (7.4.1)

$$\left[\sup_{\tau(k) \leq t} \xi_k \; , \; \sup_{\tau(\ell) \leq t} \theta_\ell\right]$$

Before deriving an expression for the joint distribution function of ($\sup \xi_k$, $\sup \theta_\ell$) and ($\inf \xi_k$, $\inf \theta_\ell$) we will derive an expression for the distribution function of time of occurrence of k-th exceedance for x and ℓ-th exceedances for y in this section.

We define an event

$$E_{k\,\ell}^{t,t} = \left[\tau(k) \leq t \leq \tau(k+1) \,, \tau(\ell) \leq t \leq \tau(\ell+1)\right] \qquad (7.4.2)$$

Obviously $E_{k,\ell}^{t,t} \cap E_{m,n}^{t,t} = \phi$ for $k, \ell \neq m, n$

$$\bigcup_{k\,\ell} E_{k,\ell}^{t,t} = \Omega$$

By virtue of (7.4.2) we have

$$P(E^{t,t}_{k,\ell}) = P[\tau(k) \leq t\ ,\ \tau(\ell) \leq t] - P[\tau(k) \leq t\ ,\ \tau(\ell+1) \leq t] - P[\tau(k+1) \leq t, \tau(\ell) \leq t] + P[\tau(k+1) \leq t\ ,\ \tau(\ell+1) \leq t] \qquad (7.4.3)$$

Adding (7.4.3) over all values of k and ℓ, the distribution function of time of occurrence becomes

$$F_{k\ell}(t,t) = \sum_{i=k}^{\infty} \sum_{j=\ell}^{\infty} P(E^{t,t}_{i,j}) \qquad (7.4.4)$$

In order to derive an expression for the probabilities $P(E^{t,t}_{n,m})$, we start with x_o and y_o as critical points for two variables as defined in section 7.2. Let $\eta(t)$ and $\beta(t)$ be the number of exceedances in the time interval$(o\ ,\ t]$ for x and y respectively. Then for all $t > 0$ and $\Delta t > 0$

$$\eta(t) = \eta(t+\Delta t)$$
$$\beta(t) = \beta(t+\Delta t) \qquad (7.4.5)$$

which implies that the number of exceedances for both the variables are nondecreasing function of time.

The possibilities accounting for exceedances in two variables are enumerated below

(i) $[X(t) > x_o\ ,\ Y(t) < y_o]$

(ii) $[X(t) < x_o\ ,\ Y(t) > y_o]$ (7.4.6)

(iii) $[X(t) > x_o\ ,\ Y(t) > y_o]$

The corresponding probabilities can be derived as follows. Suppose that in time interval $(0, t]$, k and ℓ exceedances had already occured in x and y respectively, we denote by $\mu_{k\ell}(t)$, $\gamma_{k\ell}(t)$ and $\lambda_{k\ell}(t)$ the probabilities corresponding to three events of (7.4.6) that in short interval $(t, t+\Delta t)$ an exceedance will occur either in $x(t)$ or in $y(t)$ or in both.

$$\mu_{k\ell}(t) = \lim_{\Delta t \to 0} \frac{P\left[E^{t,t+\Delta t;\ t,t+\Delta t}_{k+1,\ell} \;\middle|\; E^{t,t}_{k,\ell}\right]}{\Delta t} \tag{7.4.7}$$

where $E^{t,t+\Delta t;\ t,t+\Delta t}_{k+1,\ell} = \left[\eta(t+\Delta t) - \eta(t) = 1,\ \beta(t+\Delta t) - \beta(t) = 0\right]$

i.e. one exceedance occurs in $x(t)$ with no exceedance in $y(t)$. Similarly for the second event we have probability

$$\gamma_{k\ell}(t) = \lim_{\Delta t \to 0} \frac{P\left[E^{t,t+\Delta t;\ t,t+\Delta t}_{k,\ell+1} \;\middle|\; E^{t,t}_{k,\ell}\right]}{\Delta t}$$

where $E^{t,t+\Delta t;\ t,t+\Delta t}_{k,\ell+1} = \left[\eta(t+\Delta t) - \eta(t) = 0,\ \beta(t+\Delta t) - \beta(t) = 1\right].$ (7.4.8)

and finally the probability for the third event is

$$\lambda_{k\ell}(t) = \lim_{\Delta t \to 0} \frac{P\left[E^{t,t+\Delta t;\ t,t+\Delta t}_{k+1,\ell+1} \;\middle|\; E^{t,t}_{k,\ell}\right]}{\Delta t} \tag{7.4.9}$$

where $E^{t,t+\Delta t;\ t,t+\Delta t}_{k+1,\ell+1} = \left[\eta(t+\Delta t) - \eta(t) = 1,\ \beta(t+\Delta t) - \beta(t) = 1\right]$

The probabilities $P(E^{t,t}_{k,\ell})$ satisfy a system of differential equations which may be derived by considering the events and their associated conditional probabilities. Using the arguments of a stochastic process we get the following differential equation

$$P(E_{k,\ell}^{t,t+\Delta t;\ t,t+\Delta t}) = \lambda_{k\ell}(t)\,\Delta(t)\,P(E_{k-1,\ell-1}^{t,t}) \tag{7.4.10}$$
$$+ \mu_{k\ell}(t)\,\Delta(t)\,P(E_{k-1,\ell}^{t,t})$$
$$+ \gamma_{k\ell}(t)\,\Delta(t)\,P(E_{k,\ell-1}^{t,t})$$
$$+ \; 1-[\lambda_{k\ell}(t) + \mu_{k\ell}(t) + \gamma_{k\ell}(t)]\,\Delta t \;\; P(E_{k,\ell}^{t,t})$$

On rearrangement the above equation may be written as

$$\frac{d\,P(E_{k,\ell}^{t,t})}{dt} = \lambda_{k\ell}(t)\;P(E_{k-1,\ell-1}^{t,t}) \tag{7.4.11}$$
$$+ \mu_{k\ell}(t)\,P(E_{k-1,\ell}^{t,t})$$
$$+ \gamma_{k\ell}(t)\,P(E_{k,\ell-1}^{t,t})$$
$$- [\lambda_{k\ell}(t) + \mu_{k\ell}(t) + \gamma_{k\ell}(t)]\,P(E_{k,\ell}^{t,t})$$

This equation in general can not be solved in closed form in terms of $\lambda_{k\ell}(t)$, $\mu_{k\ell}(t)$ and $\gamma_{k\ell}(t)$, however, we believe that in the relevant case of flood analysis, the intensity functions are indeoendent of k and ℓ. We take

$$\lambda_{k\ell}(t) = \lambda(t)$$
$$\mu_{k\ell}(t) = \mu(t)$$
and
$$\gamma_{k\ell}(t) = \gamma(t).$$

Consequently, the equation(7.4.11) can be expressed as

$$\frac{d\,P(E^{t,t}_{k,\ell})}{dt} = \lambda(t)\,P(E^{t,t}_{k-1,\ell-1}) + \mu(t)\,P(E^{t,t}_{k-1,\ell}) + \gamma(t)\,P(E^{t,t}_{k,\ell-1}) - [\lambda(t)+\mu(t)+\gamma(t)]\,P(E^{t,t}_{k,\ell}) \tag{7.4.12}$$

The above equation can be solved with the help of probability generating function. The equation in the form of generating function is

$$\sum\sum s_1^k\, s_2^\ell\, \frac{dP(E^{t,t}_{k,\ell})}{dt} = \lambda(t)\sum\sum s_1^k\, s_2^\ell\, P(E^{t,t}_{k-1,\ell-1}) + \mu(t)\sum\sum s_1^k\, s_2^\ell\, P(E^{t,t}_{k-1,\ell}) + \gamma(t)\sum\sum s_1^k\, s_2^\ell\, P(E^{t,t}_{k,\ell-1}) - [\lambda(t)+\mu(t)+\gamma(t)]\sum\sum s_1^k\, s_2^\ell\, P(E^{t,t}_{k,\ell})$$

or

$$\frac{dP(s_1,s_2,t)}{dt} = \lambda(t)\, s_1\, s_2\, P(s_1,s_2,t) + \mu(t)\, s_1\, P(s_1,s_2,t) + \gamma(t)\, s_2\, P(s_1,s_2,t) - [\lambda(t)+\mu(t)+\gamma(t)]\, P(s_1,s_2,t) \tag{7.4.13}$$

where $P(s_1,s_2,t) = \sum\sum s_1^k\, s_2\, P(E^{t,t}_{k,\ell})$.

Rearranging (7.4.13) we have

$$\frac{1}{P(s_1,s_2,t)} \cdot \frac{dP(s_1,s_2,t)}{dt} = \lambda(t)\,(s_1 s_2-1) + \mu(t)(s_1-1) + \gamma(t)\,(s_2-1)$$

or

$$\frac{d\,\log P(s_1,s_2,t)}{dt} = \lambda(t)\,(s_1 s_2-1) + \mu(t)\,(s_1-1) + \gamma(t)(s_2-1)$$

$$\log P(s_1,s_2,t) = \int_0^t \left[\lambda(t)\,(s_1 s_2-1) + \mu(t)\,(s_1-1) + \gamma(t)\,(s_2-1)\right]dt + C . \qquad (7.4.14)$$

Where C is a constant of integration. $t = 0$ and $P(s_1,s_2,0)= 1$ give $C = 0$ from initial conditions. Then from (7.4.14) we have

$$\log P(s_1,s_2,t) = \int_0^t \left[\lambda(t)\,(s_1 s_2-1)+ \mu(t)\,(s_1-1)+\gamma(t)\,(s_2-1)\right]dt . \qquad (7.4.15)$$

or

$$\log P(s_1,s_2,t) = (s_1\,s_2-1)\int_0^t \lambda(t)\,dt + (s_1-1)\int_0^t \mu(t)\,dt + (s_2-1)\int_0^t \gamma(t)\,dt . \qquad (7.4.16)$$

Further putting

$$\int_0^t \lambda(t)\,dt = W(t) , \quad \int_0^t \mu(t)\,dt = U(t)$$

and

$$\int_0^t \gamma(t)\,dt = V(t)$$

we get

$$P(s_1,s_2,t) = \exp\left[(s_1 s_2-1)W(t)+ (s_1-1)U(t)+ (s_2-1)V(t)\right] \qquad (7.4.17)$$

Expression(7.4.17) is identified by Martin(1952) as the probability generating function of bivariate Poisson distribution with parameter $W(t)$, $U(t)$ and $V(t)$. The probability of Poisson distribution corresponding to generating function(7.4.17) is seen to be

$$P(E^{t,t}_{k,\ell}) = \exp[-(U(t)+V(t)+W(t))]\left(\frac{a^k}{k!}\right)\left(\frac{b^\ell}{\ell!}\right) \cdot \sum_{r=0}^{s} \frac{k^{(r)}\ \ (\ell)^{(r)}}{a^r\ b^r}\ \frac{[W(t)]^r}{r!} \tag{7.4.18}$$

where

$$a = [U(t) - W(t)] > 0$$

$$b = [V(t) - W(t)] > 0$$

$$k^{(r)} = k(k-1)\ \ldots\ (k-r+1)$$

and

$$s = \min(k\ ,\ \ell\ ) .$$

7.5 Distribution function of ($\operatorname{Sup}\xi_k$, $\operatorname{Sup}\theta_\ell$) and ($\operatorname{Inf}\xi_k$, $\operatorname{Inf}\theta_\ell$):

We now proceed to derive the joint distribution function of two extreme of two dependent sequences of exceedances ξ_k and θ_ℓ. The number of these exceedanees for the sequences ξ_k and θ_ℓ will be taken n and m respectively. These exceedances can occur in the following three ways.

(i) There may be an exceedance for the sequence ξ_k but not for the sequence θ_ℓ in the time $(0,t]$.

(ii) There may be an exceedance for the sequence θ_ℓ but not for sequence ξ_k in the time $(0,t]$.

(iii) There is an exceedance for both the sequence ξ_k and θ_ℓ in time $(0,t]$.

First, we prove two theorems for the joint distribution functions satisfying the assumption 'A' given in 7.3 ,i.e.,

(i) the sequences $\xi_1, \xi_2, \ldots, \xi_n$ and $\theta_1, \theta_2, \ldots, \theta_m$ of random variables have their common bivariate distribution function $F(x,y)$ with $F_1(x)$ and $F_2(y)$ as marginal distribution function, and

(ii) the number of exceedances $\eta(t)$ and $\beta(t)$ in both the variables are independent of their magnitude ξ_k and θ_l.

Theorem 7.5.1 :- If the sequence of random variables satisfy the assumption 'A', then the joint distribution function of maximum of two dependent sequences is given by

$$F_{ts}(x,y) = \exp\Big[- U(t) - V(t) + W(t) + W(t)\ F(x,y) + (U(t) - W(t))\ F(x) + (V(t) - W(t))\ F(y)\Big]. \tag{7.5.1}$$

Proof:- The joint distribution function $F_{ts}(x,y)$ of the maximum of two dependent sequences is derived as the mathematical expectation of the following conditional probability

$$P\left[\sup_{\tau(k)\le t} \xi_k \le x\ ,\ \sup_{\tau(l)\le t} \theta_l \Big/ \eta(t),\ \beta(t)\right] \tag{7.5.2}$$

The expression for $F_{ts}(x,y)$ is

$$F_{ts}(x,y) = P\left[\sup_{\tau(k)\le t} \xi_k \le x,\ \sup_{\tau(l)\le t} \theta_l \le y\right]$$

$$= EP\left[\sup_{\tau(k)\le t} \xi_k \le x,\ \sup_{\tau(l)\le t} \theta_l \le y \Big/ \eta(t)\ ,\ \beta(t)\right]$$

$$= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} P\left[\sup_{0\le k\le n} \xi_k \le x,\ \sup_{0\le l\le m} \theta_l \le y\ \cap\ E^{t,t}_{n,m}\right] \tag{7.5.3}$$

Under assumption 'A', the expression(7.5.3) can be written as

$$= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} P\left[\sup_{0 \le k \le n} \xi_k \le x, \sup_{0 \le \ell \le m} \theta_\ell \le y\right] P(E^{t,t}_{n,m}) \qquad (7.5.4)$$

Taking into the consideration the three situations in which three exceedances in two variables occur, the expression for

$$P\left[\sup_{0 \le k \le n} \xi_k \le x, \sup_{0 \le \ell \le m} \theta_\ell \le y\right]$$

is given by

$$\sum_{r=0}^{\min(m,n)} [F(x,y)]^r \, [F(x)]^{n-r} \, [F(y)]^{m-r} \qquad (7.5.5)$$

where r is the number of exceedances when the exceedances in both the sequences occur.

Substituting (7.5.5) in (7.5.4), we obtain

$$F_{t_s}(x,y) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \sum_{r=0}^{\min(m,n)} [F(x,y)]^r [(F(x)]^{n-r} [F(y)]^{m-r} P(E^{t,t}_{n,m})$$

(7.5.6)

$$= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \sum_{r=0}^{\min(m,n)} [F(x,y)]^r [F(x)]^{n-r} [F(y)]^{m-r}$$

$$\exp[-U(t) - V(t) + W(t)] \frac{[U(t) - W(t)]^n}{n!} \frac{[V(t) - W(t)]^m}{m!} .$$

$$\cdot \sum_{r=0}^{\min(m,n)} \frac{[W(t)]^r}{r!} \frac{n^{(r)}}{[U(t) - W(t)]^r} \frac{m^{(r)}}{[V(t) - W(t)]^r}$$

(7.5.7)

after $P(E^{t,t}_{n,m})$ has been replaced by (7.4.18).

Interchanging and rearranging the summations and the limits in

(7.5.7) we obtain

$$F_{t_s}(x,y) = \exp\left[-U(t) - V(t) + W(t)\right] \sum_{r=0}^{\infty} \sum_{n=r}^{\infty} \sum_{m=r}^{\infty} \frac{\left[W(t)\,F(x,y)\right]^{r}}{r!}$$

$$\frac{\left[(U(t) - W(t))F(x)\right]^{n-r}}{(n-r)!} \quad \frac{\left[(V(t) - W(t))\;F(y)\right]^{m-r}}{(m-r)!} \qquad (7.5.8)$$

$$= \exp\left[-U(t) - V(t) + W(t)\right] \sum_{r=0}^{\infty} \frac{\left[W(t)\;F(x,y)\right]^{r}}{r!}$$

$$\sum_{n=r}^{\infty} \frac{\left[(U(t) - W(t))\;F(x)\right]^{n-r}}{(n-r)!} \sum_{m=r}^{\infty} \frac{\left[\left[V(t) - W(t)\right]F(y)\right]^{m-r}}{(m-r)!} \qquad (7.5.9)$$

$$= \exp\left[-U(t) - V(t) + W(t)\right] \exp\left[W(t)\;F(x,y)\right]$$

$$\exp\left[\left[U(t) - W(t)\right]F(x)\right] \exp\left[\left[U(t) - W(t)\right]F(y)\right] \qquad (7.5.10)$$

$$= \exp\left[-U(t) - V(t) + W(t) + W(t)\;F(x,y) + \left[U(t) - W(t)\right]F(x) + \left[V(t) - W(t)\right]F(y)\right] \qquad (7.5.11)$$

The distribution function(7.5.11) satisfies the properties of distribution function and we can get marginals from it by substituting $x = \infty$, $y = \infty$ as follows

$$F_{t_s}'(x,\infty) = \exp\left[-U(t) - V(t) + W(t) + F(x,\infty)\;W(t) + \left[U(t) - W(t)\right]F(x) + \left[V(t) - W(t)\right]F(\infty)\right]$$

$$= \exp\left[-U(t)\;(1 - F(x))\right]$$

and

$$F_{t_s}(\infty,y) = \exp\left[-V(t)\;(1 - F(y))\right].$$

These forms are same as given(7.3.1) for one variable case.

Theorem 7.5.2 :- The joint distribution function $F_{t1}(x,y)$ of minimum observations of two dependent sequences is given by

$$F_{t1}(x,y) = 1+\exp[-U(t)]-\exp[-U(t)\ (1-P(x))]+\exp[-V(t)]$$
$$-\exp[-V(t)\ (1-P(y))]-\exp[-U(t)-V(t)+W(t)]$$
$$+\exp\Big[-U(t)-V(t)+W(t)+W(t)\ P(x,y)+P(y)\ [V(t)-W(t)]$$
$$+P(x)[U(t)-W(t)]\Big] \qquad (7.5.12)$$

where

$$P(x) = 1-F(x)$$
$$P(y) = 1-F(y)$$
$$P(x,y) = 1-F(x)-F(y)+F(x,y) \qquad (7.5.13)$$

Proof:- The bivariate distribution function for the infimum of two dependent sequences ξ_k and θ_ℓ of independent random variables is derived as the mathematical expectation of the conditional probability

$$P\left[\inf_{T(k)\le t}\xi_k \le x,\ \inf_{T(\ell)\le t}\theta_\ell \le y \,\Big/\, \eta(t),\ \beta(t)\right] \qquad (7.5.14)$$

as follows

$$F_{t1}(x,y) = P\left[\inf_{T(k)\le t}\xi_k \le x,\ \inf_{T(\ell)\le t}\theta_\ell \le y\right]$$

$$= E\ P\left[\inf_{T(k)\le t}\xi_k \le x,\ \inf_{T(\ell)\le t}\theta_\ell \le y \,\Big/\, \eta(t),\ \beta(t)\right]$$

$$= 1- E\ P\left[\inf_{T(k)\le t}\xi_k > x \,\Big/\, \eta(t),\ \beta(t)\right]$$

$$- E\ P\left[\inf_{T(\ell)\le t}\theta_\ell > y \,\Big/\, \eta(t),\ \beta(t)\right]$$

$$+ E\ P\left[\inf_{T(k)\le t}\xi_k > x,\ \inf_{T(\ell)\le t}\theta_\ell > y \,\Big/\, \eta(t),\ \beta(t)\right]$$

$$= 1 - \sum_{n=1}^{\infty} P\left[\inf_{0 \le k \le n} \xi_k > x \cap E_n^t\right]$$

$$- \sum_{m=1}^{\infty} P\left[\inf_{0 \le \ell \le m} \theta_\ell > y \cap E_m^t\right]$$

$$+ \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} P\left[\inf_{0 \le k \le n} \xi_k > x, \inf_{0 \le \ell \le m} \theta_\ell > y \cap E_{n,m}^{t,t}\right] \qquad (7.5.15)$$

Under the assumptions which are given for maximum case the bivariate distribution function(7.5.15) can be written as

$$F_{t1}(x,y) = 1 - \sum_{n=1}^{\infty} P\left[\inf_{0 \le k \le n} \xi_k > x\right] P(E_n^t)$$

$$- \sum_{m=1}^{\infty} P\left[\inf_{0 \le \ell \le m} \theta_\ell > y\right] P\left[E_m^t\right]$$

$$+ \sum_{m=1}^{\infty} \sum_{m=1}^{\infty} P\left[\inf_{0 \le k \le n} \xi_k > x, \inf_{0 \le \ell \le m} \theta_\ell > y\right] P\left[E_{n,m}^{t,t}\right] \qquad (7.5.16)$$

Taking into account the three possibilities of the events

$$P\left[\inf_{0 \le k \le n} \xi_k > x, \inf_{0 \le \ell \le m} \theta_\ell > y\right]$$

is given by

$$\sum_{r=0}^{\min(m,n)} \left[1-F(x)\right]^{n-r} \left[1-F(y)\right]^{m-r} \left[1-F(x)-F(y)+F(x,y)\right]^r$$

or

$$\sum_{r=0}^{\min(m,n)} \left[P(x,t)\right]^r \left[P(x)\right]^{n-r} \left[P(y)\right]^{m-r} \qquad (7.5.17)$$

Substituting the expression for the three probabilities in (7.5.16) we get

$$F_{t_1}(x,y) = 1 - \sum_{n=1}^{\infty} [P(x)]^n \; P[E_n^t] - \sum_{m=1}^{\infty} [P(y)]^m \; P[E_m^t]$$

$$+ \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \sum_{r=0}^{\min(m,n)} [P(x,y)]^r [P(x)]^{n-r} [P(y)]^{m-r} \; P[E_{n,m}^{t,t}]$$

$$= 1 - \sum_{n=1}^{\infty} [P(x)]^n \frac{\exp[-U(t)][U(t)]^n}{n!}$$

$$- \sum_{m=1}^{\infty} P(y)^m \frac{\exp[-V(t)][V(t)]^m}{m!}$$

$$+ \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \sum_{r=0}^{\min(m,n)} [P(x,y)]^r [P(x)]^{n-r} [P(y)]^{m-r}$$

$$\exp[-U(t)-V(t)+W(t)] \frac{[U(t)-W(t)]^n}{n!} \frac{[V(t)-W(t)]^m}{m!}$$

$$\sum_{r=0}^{\min(m,n)} \frac{[W(t)]^r}{r!} \frac{n^{(r)}}{[U(t)-W(t)]^r} \frac{m^{(r)}}{[V(t)-W(t)]^r}$$

(7.5.18)

when $P(E_n^t)$, $P(E_m^t)$ and $P(E_{n,m}^{t,t})$ are replaced with their values from (7.2.15) and (7.4.18).

Rearranging and interchanging the summations for last expression in (7.5.18) we get

$$F_{t1}(x,y) = 1 + \exp\left[-U(t)\right] - \exp\left[-U(t)\,(1-P(x))\right]$$
$$+ \exp\left[-V(t)\right] - \exp\left[-V(t)\,(1-P(y))\right]$$
$$- \exp\left[-U(t) - V(t) + W(t)\right] + \exp\left[-U(t) - V(t)\right.$$
$$\left. + W(t) + W(t)P(x,y) + P(y)(V(t) - W(t)) + P(x)(U(t) - W(t))\right]$$
$$(7.3.19)$$

In the next sections the results obtained in section 7.2 to section 7.5 are used for predicting recurrences over a fixed danger level.

7.6 Application of one dimensional theory:

In this section results obtained in section 7.3 are applied to the flood frequency analysis of river Narmada. The results of section 7.2 to section 7.5 of this chapter is of immence practical applications. We note that the variables under study are based on very general conditions given as assumption 'A' in section 7.3. We shall illustrate below that the study of flood phenomenon can be carried using the above results. It is natural to have the data on flood level at two different places for such study. We also verify that the assumption 'A' are satisfied by 'flood levels'. Therefore all the results derived in the section 7.2 to section section 7.5 can be applied to this variable and it will be possible to find the probability distribution of the exceedances when the flood level at a certain point of a river exceeds some preassigned mark. This may be helpful for planning purposes.

For illustration we have taken the data on the river Narmada at two points i.e. Mortakka and Gardeshwar. However, it has not been possible for us to obtain the data on flood level but the data of flood discharge was easily made available from 'central water commission'. Since two variables flood level and flood discharge are very highly correlated we have used the flood discharge as a proxy variable for the flood level. It may be recalled that we are using only a partial duration data which is taken for 30 years viz. from 1948 to 1977* as indicated earlier. The time interval of interest has been taken from July Ist (which is starting point for rainy season of every year) to October 28 for both the stations. The expression(7.3.12) indicates that the estimation and prediction of floods is possible through it, if one has the estimates of the probabilities $P(E_k^t)$.

It is clear from equation(7.2.15), in order to estimate $P(E_k^t)$, we have to evaluate the function $\Lambda(t)$ (the average number of exceedances in $(0, t)$). The function $\Lambda(t)$ for the stations Mortakka and Gardeshwar is denoted by $U(t)$ and $V(t)$ respectively and determined in the following manner:

---

* The author acknowledges greatfully Dr.M.N.Das and Central Water Commission for providing the required data.

The partial duration series of water discharge values is divided into eight intervals of fifteen days periods for each year. As mentioned earlier, the base level is taken as 150,000 c.f.s.(cubic feet per second) for both the stations. The number and magnitude of exceedances have been given in appendix.

Zelenhasic(1970) has expressed the function $\Lambda(t)$ as a finite Fourier series based on the analysis of data. Using the same method, the expression for U(t) and V(t) for the exceedances given in appendix (Table 1 and 2) are given as follows

$$\begin{aligned} U(t) = & - 5.2142856 + 14.0396824\, t + 6.5357143\, t^2 \\ & - 0.6944444\, t^3 - 0.0181075 \cos 45 t \\ & - 0.0106729 \sin 45 t + 0.0554091 \cos 135 t \\ & + 0.0583746 \sin 135 t + 0.0238095 \cos 90 t \\ & + 0.0738095 \sin 90 t + 0.0111111 (-1)^t \end{aligned} \quad (7.6.1)$$

$$\begin{aligned} V(t) = & - 1.5000000 + 5.7871573\, t + 7.5259740\, t^2 \\ & - 0.7373737\, t^3 - 0.0157369 \cos 45 t \\ & - 0.0110213 \sin 45 t + 0.0113997 \cos 90 t \\ & + 0.0838383 \sin 90 t + 0.0648711 \cos 135 t \\ & + 0.0288776 \sin 135 t + 0.0096580 (-1)^t \end{aligned} \quad (7.6.2)$$

To demonstrate the appropriateness of the model observed and computed values of U(t) and V(t) are shown in Fig(7.6.1) and Fig(7.6.2) . This has come to be an excellent fit.

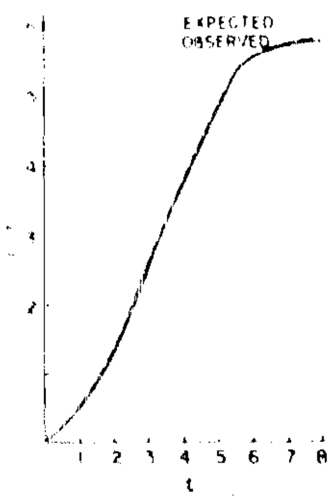

FIG. 7.6.1 OBSERVED U (t) AND FITTING FUNCTION. UNIT INTERVAL ON t AXIS STANDS FOR 15 DAYS PERIOD

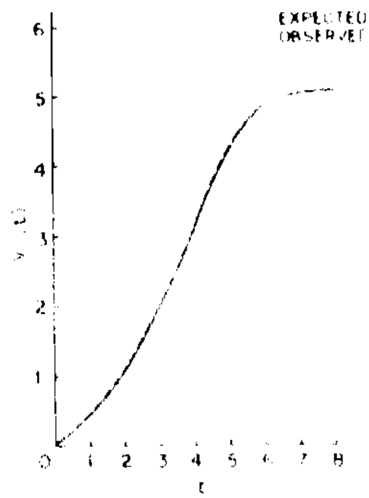

FIG. 7.6.2 OBSERVED V(t) AND FITTING FUNCTION. UNIT INTERVAL ON t AXIS STANDS FOR 15 DAYS PERIOD

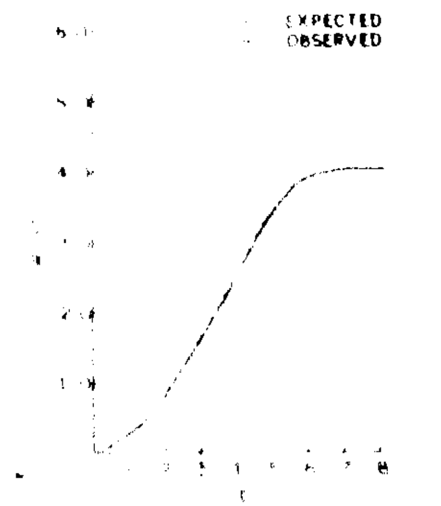

FIG. 7 [illegible] OBSERVED AND FITTING FUNCTION [illegible] UNIT INTERVAL ON t AXIS STANDS FOR 15 DAYS INTERVAL

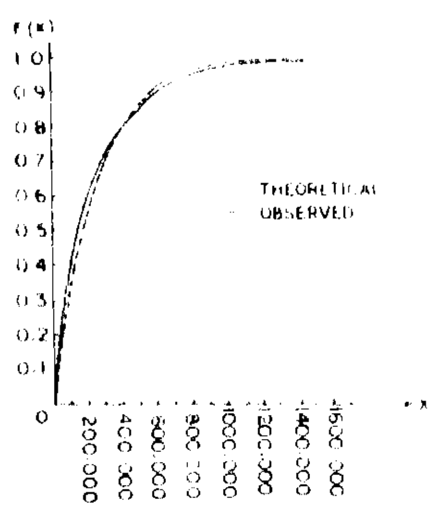

FIG. 7.6.3 OBSERVED AND THEORETICAL DISTRIBUTION FUNCTIONS OF EXCEEDANCES FOR THE NARMADA RIVER AT MORTAKKA FOR 120 DAYS PERIOD

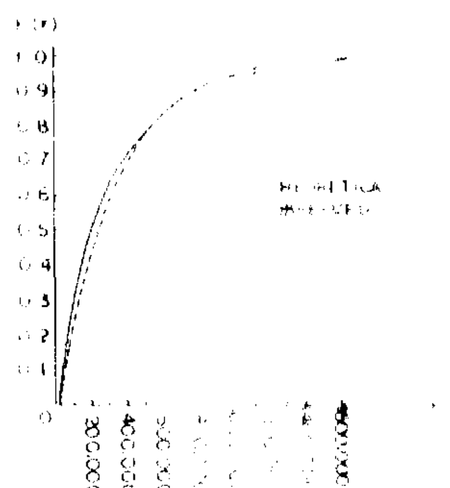

FIG. 7.6.4 OBSERVED AND THEORETICAL DISTRIBUTION FUNCTIONS OF EXCEEDANCES FOR THE NARMADA RIVER [illegible] FOR 120 DAYS PERIOD

The next step is to determine the distribution function of the flood peak exceedances. In consistancy with the assumptions of (7.3.12) , we assume that the exceedences during the given period are independent and identically distributed. The magnitude of these exceedances have a common distribution function(7.2.11) with an unknown parameter (α). For both the stations the value of this parameter of the distribution function(7.2.11) is estimated by the well known technique i.e. maximum likelihood method.

The values (α_1, α_2) of α for these two stations are presented in the following table(7.6.1)

Table 7.6.1

Station	$E(\xi_k)$	$\alpha = [E(\xi_k)]^{-1}$
Mortakka(α_1)	241037.93	4.1487246×10^{-6}
Gardeshwar(α_2)	343391.88	2.9121247×10^{-6}

The distribution function(7.2.11) can be calculated now for different values of x . The observed and corresponding theoretical distribution functions of the flood peak exceedances are represented in figure(7.6.3) and (7.6.4).
The distribution functions of the largest flood exceedances for both the stations can now be written using(7.2.12) , (7.6.1) and (7.6.2) as

$$F_{t_s}(x) = \exp(-\ 5.7333333\ e^{-4.1487246\ \times\ 10^{-6}\ x}) \quad (7.6.3)$$

at Mortakka , and

$$F_{t_s}(y) = \exp(-\ 5.0333333\ e^{-2.912124\ \times\ 10^{-6}\ y}) \quad (7.6.4)$$

at Gardeshwar, where the time of interest was from July Ist to October 28 .

Here x and y are measured in c.f.s. . Fig.(7.6.5) and (7.6.6) compare graphically the distribution functions (7.6.3) and (7.6.4) and corresponding observed distributions. A fairly good agreement between theoretical and observed results indicates that the assumptions underlying the derivation of the model are basically correct.

In the following the return period is calculated with the help of distribution functions (7.6.3) and (7.6.4).

<u>Return period</u>:- As defined in (1.3.2), the return period of a particular value x can be written as

$$T(x) = \frac{1}{1-F(x)} \quad (7.6.5)$$

Substituting in (7.6.5) $F_{t_s}(x)$ and $F_{t_s}(y)$ for $F(x)$ from (7.6.3) and (7.6.4) , the expected return periods for different values of x and y are calculated and given in the following table (7.6.2). The calculated values are also presented graphically in the Fig.(7.6.7) and (7.6.8) .

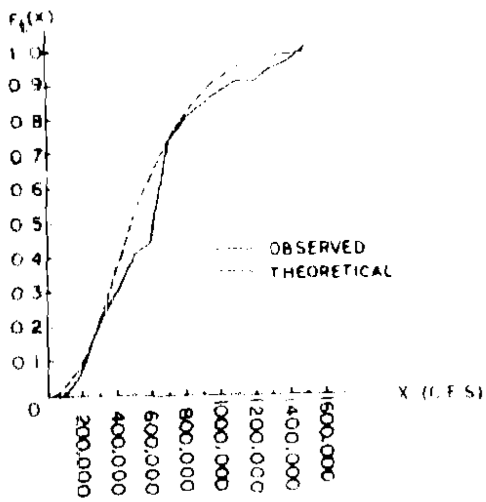

FIG 7.6.5 OBSERVED AND THEORETICAL DISTRIBUTION FUNCTION OF THE MAXIMUM FLOOD PEAK EXCEEDANCES FOR THE NARMADA RIVER AT MORTAKKA

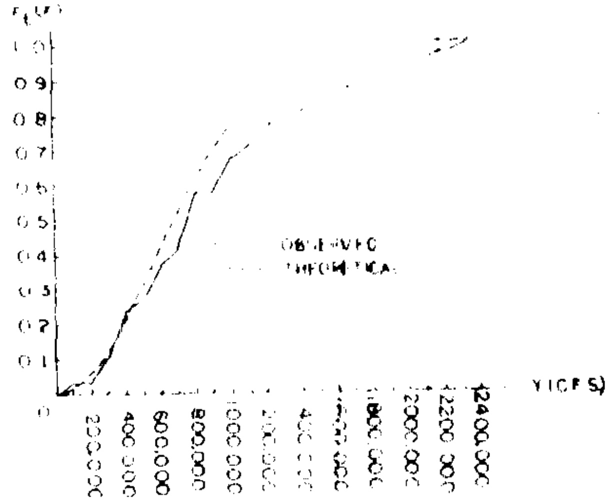

FIG 7.6.6 OBSERVED AND THEORETICAL DISTRIBUTION FUNCTION OF THE MAXIMUM FLOOD PEAK EXCEEDANCES FOR THE NARMADA RIVER AT GARDESHWAR

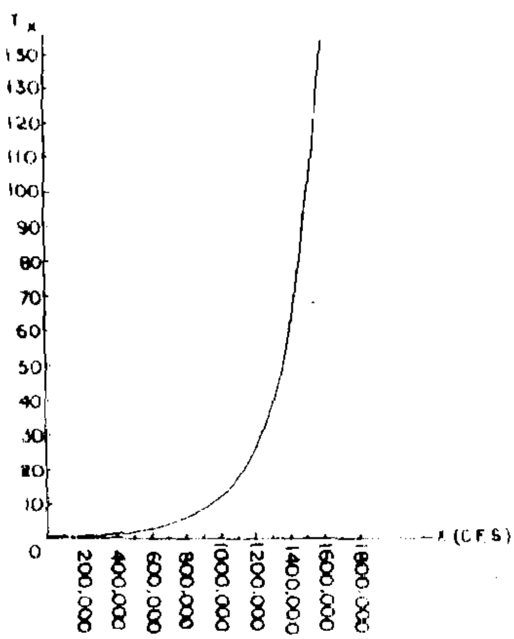

FIG 7.6.7 RETURN PERIOD AT MORTAKKA

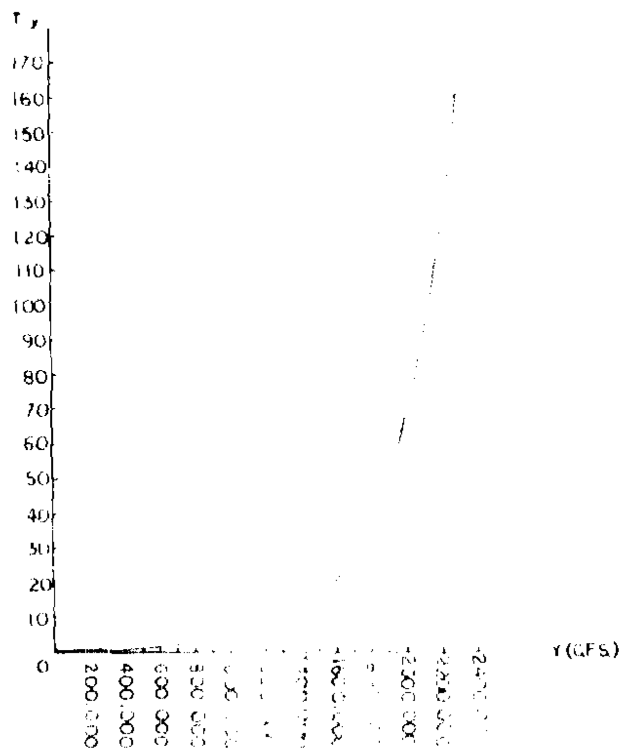

FIG 7.6.8 RETURN PERIOD AT GARDESHWAR

Table 7.6.2

Return period at Mortakka and Gardeshwar

Water discharge	Return period at Mortakka	Return period at Garceshwar
100,000	1.0232	1.0238
200,000	1.0894	1.0639
300,000	1.2373	1.1238
400,000	1.5059	1.2626
500,000	1.9478	1.4477
600,000	2.6416	1.7124
700,000	3.7091	2.0799
800,000	5.3369	2.5820
900,000	7.8092	3.2619
1000,000	11.5574	4.1775
1100,000	17.2364	5.4074
1200,000	25.8331	7.0562
1300,000	38.8500	9.2651
1400,000	58.5703	12.2221
1500,000	88.4885	16.1814
1600,000	133.6255	21.4792
1700,000	202.2040	28.5709
1800,000	305.8759	38.0507
1900,000	461.9084	50.7511
2000,000	700.2801	67.7598
2100,000	1064.7359	90.4454
2200,000	1613.6840	120.9029
2300,000	2440.2147	161.5117

7.7. Application of two dimensional theory:

Here in, the results obtained in sections 7.4 and 7.5 are applied to the same data which is used in previous section. It was seen in section 7.6 that ξ_i's and θ_i's are exponentially distributed with their respective distribution functions

$$F(x) = 1 - e^{-\alpha_1 x} \qquad x \geq 0 \, , \, \alpha_1 > 0$$

$$F(y) = 1 - e^{-\alpha_2 y} \qquad y \geq 0 \, , \, \alpha_2 > 0 \qquad (7.7.1)$$

It is also obvious that their joint bivariate distribution function should be of exponential type. As it is well known that infinite number of distribution functions can exist corresponding to given marginals, then the question arises, what should be the form of bivariate exponential distribution function $F(x,y)$ given that x and y themselves are exponentially distributed. The commonly used bivariate distribution functions which are having above given exponential marginals are given below.

(i) The Morgenstern distribution

$$F(x,y) = 1-e^{-x} - e^{-y} + e^{-x-y} \left[1+ a(1-e^{-x})(1-e^{-y})\right] \quad (7.7.2)$$

(ii) Gumbel's type I distribution

$$F(x,y) = 1- e^{-x} - e^{-y} + \exp(-x - y + \theta x y)$$

$$x \geq 0, \; y \geq 0, \; 0 \leq \theta \leq 1 \quad (7.7.3)$$

(iii) Gumbel's type II distribution

$$F(x,y) = 1 - e^{-x} - e^{-y} + \exp\left[-\left(x^m + y^m\right)^{\frac{1}{m}}\right] \qquad (7.7.4)$$

(iv) The Marshall-Olkin distribution

$$F(x,y) = 1 - e^{-x} - e^{-y} + \exp\left[-x - y - \lambda \max(x,y)\right]$$

$$x \geqslant 0, \quad y \geqslant 0, \lambda > 0 \qquad (7.7.5)$$

(v) Mardia distribution

$$F(x,y) = 1 - e^{-x} - e^{-y} + \left(e^{x} + e^{y}\right)^{-1}$$

$$x \geqslant 0, \quad y \geqslant 0 \qquad (7.7.6)$$

It will be noticed that all the above distribution functions except that of Mardia's distribution contain an unknown dependence parameter. The so called dependence parameter can be estimated by the methods given in section 5.2 . The selection of bivariate distribution function is depends much upon the estimation of dependence parameter. Due to complicacy of calculations it is not easy to find out the estimate for all the bivariate distribution function with the help of all the existing methods. It can only be estimated either by relating it to the different correlation coefficients or by the method of quadrants when they are used in (7.5.1) .

It is again noted that for the above given bivariate distribution functions the value of the estimate for dependence parameter ranges from 0.0 to 0.7 by maximum except Gumbel's type II, however, in practical situations the correlation may lies between - 1 to + 1 . Now we can say that the main criterion in selection of bivariate exponential distribution is the estimation of dependence parameter.

On the basis of the calculations for the estimation of dependence parameter, it is seen that Gumbel type II distribution is an appropriate form for $F(x,y)$ in (7.6.1) . Substituting the expression of $F(x)$, $F(y)$ and $F(x,y)$ from (7.7.1) and (7.7.4) in (7.6.1) we have

$$F_{t_s}(x,y) = \exp\Big[- U(t)\, e^{-\alpha_1 x} - V(t)\, e^{-\alpha_2 y} + W(t) \exp\Big[-((\alpha_1 x)^m + (\alpha_2 y)^m)^{\frac{1}{m}}\Big]\Big] \quad (7.7.7)$$

and its marginals are

$$F_{t_s}(x) = \exp\left[- U(t)\, e^{-\alpha_1 x}\right]$$
$$F_{t_s}(y) = \exp\left[- V(t)\, e^{-\alpha_2 y}\right] \quad (7.7.8)$$

It is worthwhile to mention, that the functional form of the distribution function (7.7.7) is similar to the asymptotic bivariate extreme value distribution given by Gumbel and Mustafi (1967) . However, the Gumbel and Mustafi's distribution is an asymptotic expression, where as (7.7.7) represents an exact (non-asymptotic) expression for the bivariate distribution.

The estimation and prediction of the floods by the distribution function(7.7.7) is possible, if one has the estimates of the constants involved in (7.7.7). The constants $U(t)$, $V(t)$, α_1 and α_2 can be estimated with the help of marginals and these are given section 7.6. $M(t)$ (the expected number of exceedances in time (o , t] for both the variables) is expressed in the same way as a Fourier expansion in the following, using the data of exceedances given in Appendix (Table 3). The expression is given by

$$
\begin{aligned}
M(t) = & -5.2142856\, t + 14.0396824\, t \\
& + 6.5357143\, t^2 - 0.6944444\, t^3 \\
& - 0.0181075 \cos 45\, t - 0.0108728 \sin 45\, t \\
& + 0.0238095 \cos 90\, t + 0.0738095 \sin 90\, t \\
& + 0.0583746 \sin 135\, t + 0.0554091 \cos 135\, t \\
& + 0.0108301\, (-1)^t . \qquad (7.7.9)
\end{aligned}
$$

and presented in Fig.(7.7.1).

The dependence parameter 'm' may be estimated by relating it to the medial correlation coefficient. The medial correlation coefficient (γ) can be expressed in the form of bivariate distribution function $F(x,y)$ and its marginals $F(x)$ and $F(y)$ as follows

$$\gamma = 4\left[F(\tilde{x},\tilde{y}) - F(\tilde{x})\,F(\tilde{y})\right] \qquad (7.7.10)$$

where $\tilde{x}$ and $\tilde{y}$ are the medians.

For our case the bivariate and univariate distributions are given in (7.7.4) and (7.7.1) respectively and consequently

$$\gamma = 4\left[1 - e^{-\alpha_1 \tilde{x}} - e^{-\alpha_2 \tilde{y}} + \exp\left[-\{(x_1\, x)^m + (x_2\, y)^m\}^{\frac{1}{m}}\right]\right] - 1$$

(7.7.11)

Substituting the values of α_1, α_2, $\tilde{x}$ and $\tilde{y}$ in (7.7.11) we can find out a value of corresponding to each value of b ($= \frac{1}{m}$).

The values of α_1 and α_2 for the time interval from first July to 28th October are given by

$$\alpha_1 = 4.1487246 \times 10^{-6}$$

$$\alpha_2 = 2.9121247 \times 10^{-6}$$

and from Appendix (Table 4 and 5) the values of the medians $\tilde{x}$ and $\tilde{y}$ are

$$\tilde{x} = 173004.50$$

$$\tilde{y} = 243313.00$$

After estimating all the constants in (7.7.11) we give now a table for the value of γ corresponding to different values of b .

Table 7.7.1

The value of dependence parameter (b) $= \frac{1}{m}$	Medial correlation coefficient
1	0.039966
0.0	0.113625
0.8	0.234780
0.7	0.334973
0.6	0.436276
0.5	0.538144
0.4	0.640075
0.3	0.741542
0.2	0.842173
0.1	0.941486

The next step is to calculate the value of the medial correlation coefficient. Graphical procedure described in section 5.2 can be followed to estimate the value of γ which is given by

$$\gamma = \frac{2d}{n} - 1$$

where n = total number of points which are plotted in a scatter digram as given in Fig.(7.7.2)

d = the total number of points in positive quadrant and its opposite quadrant.

From Fig. (7.7.2) we note that

$$n = 118, \text{ and } d = 86$$

whence

$$\gamma = \frac{172}{118} - 1 = 1.457627 - 1 = 0.457627$$

Now table 7.7.1 can be used to interpolate b for the above value of γ . The corresponding 'b' comes to be 0.557 (approximately).

Return period in bivariate case - the return period in bivariate case is defind by

$$T = \frac{1}{P(X \geq x, Y \geq y)}$$

or alternatively by

$$\frac{1}{1-F_{t_s}(x) - F_{t_s}(y) + F_{t_s}(x,y)}$$

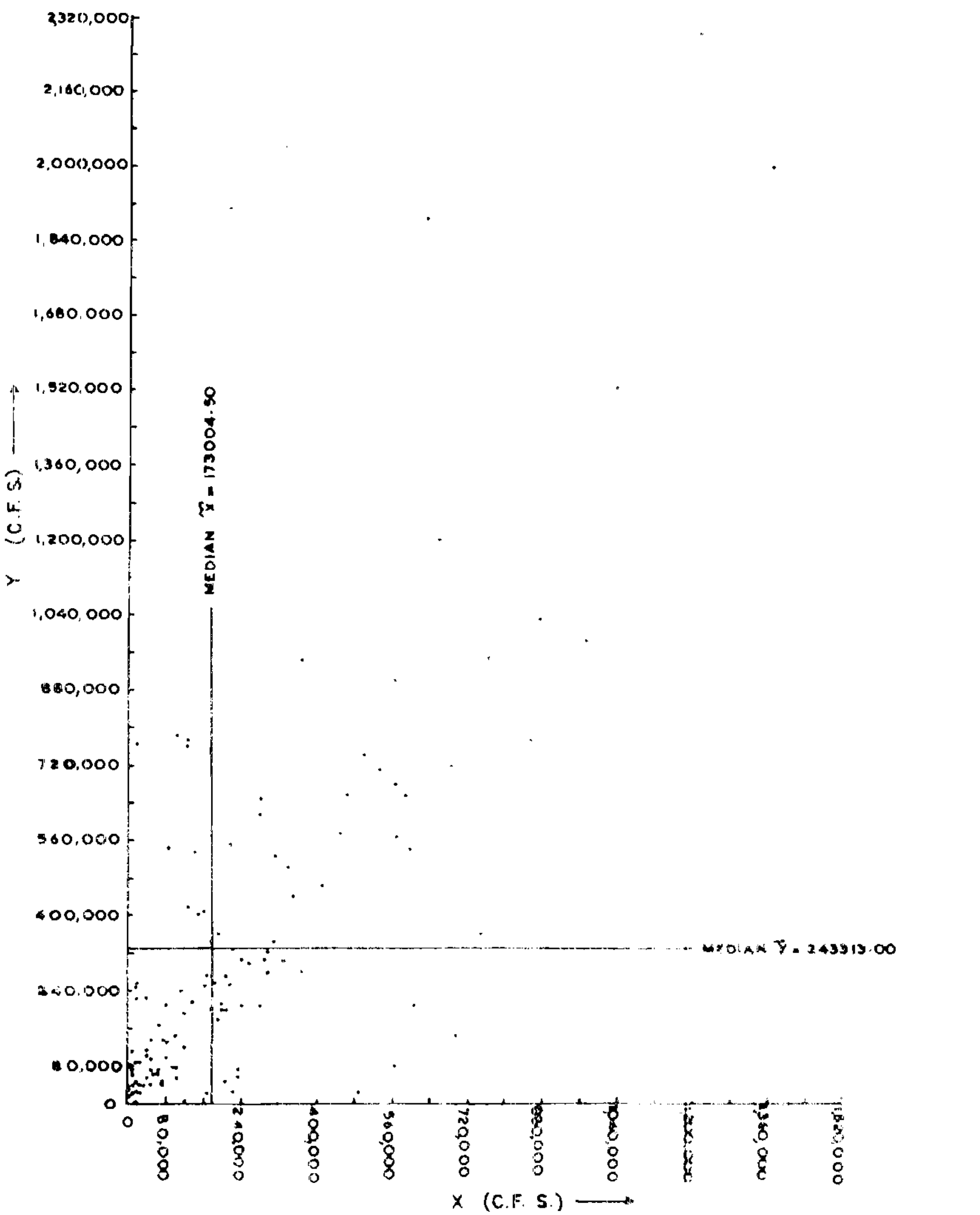

FIG. 7.7.2 SCATTER DIAGRAM OF PAIRED (X, Y)

Substituting the value of $F_{t_s}(x)$, $F_{t_s}(y)$ and $F_{t_s}(x,y)$ from (7.7.7) and (7.7.8) in the above expression we have

$$T_{x,y} = \frac{1}{1-\exp\left[-U(t)\,e^{-\alpha_1 x}\right]-\exp\left[-V(t)\,e^{-\alpha_2 y}\right]}$$

$$+\exp\Big[-U(t)\,e^{-\alpha_1 X}-V(t)\,e^{-\alpha_2 y}+W(t)$$

$$\exp\left[-\left(\left(\alpha_1 x\right)^m+\left(\alpha_2 y\right)^m\right)^{\frac{1}{m}}\right)\Big] \qquad (7.7.12)$$

Using the estimates of α_1, α_2 U(t), V(t), W(t) and m , the return period is calculated for some value of x and y and entered in table (7.7.2).

Table 7.7.2

Water discharge	Return period
100,000	1.0417
200,000	1.1373
300,000	1.3453
400,000	1.7340
500,000	2.4143
600,000	3.5785
700,000	5.5614
800,000	8.9438
900,000	14.7245
1000,000	24.6090
1100,000	41.5037
1200,000	70.3229
1300,000	119.2733
1400,000	205.9858