The Nucleus **53**, No. 3 (2016) 171-179

www.thenucleuspak.org.pk

The Nucleus
ISSN 0029-5698 (Print)
ISSN 2306-6539 (Online)

The Gumbel Mixed Model for Flood Frequency Analysis of Tarbela

S. Naz[1], M.J. Iqbal[2*], S.M. Akhter[3] and I. Hussain[4]

[1]Department of Mathematics, University of Karachi, Karachi, Pakistan
[2]Institute of Space and Planetary Astrophysics, University of Karachi, Karachi, Pakistan
[3]Department of Applied Physics, University of Karachi, Karachi, Pakistan
[4]Department of Statistics, University of Karachi, Karachi, Pakistan
sanaz@uok.edu.pk; javiqbal@uok.edu.pk; mamnoon.akhtar@uok.edu.pk; insia2003@hotmail.com

ARTICLE INFO

Article history :
Received : 08 September, 2015
Revised : 10 August, 2016
Accepted : 18 August, 2016

Keywords :
Flood volume, peak, duration
Bivariate distribution
Joint distribution
Marginal distribution
Conditional distribution

ABSTRACT

Pakistan is considered a great natural reservoir of fresh water from the River Indus. When the river is in flood it causes immense damage. Flood frequency analysis provides guidance related to the behavior of anticipated flood flows using historical flow records. Flood frequency analysis usually focuses on the annual maximum peak discharge (flood peak) values, which is insufficient to solve many problems related to hydrological engineering design, management and planning. Therefore, the values of flood peak are not considered the only flood characteristic to assess the flood event. Flood durations and volumes are also associated with the flood frequency distributions as the characteristics of a flood event for authenticated results. The application of Gumbel mixed model on the recorded flood flows is proposed in this paper to analyze the joint probability distribution of flood volumes and peaks as well as flood durations and volumes, which are mutually correlated. Gumbel mixed model is a bivariate extreme value distribution with marginal distributions of two random variables by which the joint probability distributions, the conditional probability functions and the related return periods are obtained. Application of the suggested model on the observed data of Tarbela dam in Pakistan reveals that the model is appropriate for representing the joint probability distribution of flood volumes and peaks, and the joint probability distribution of flood durations and volumes. Hence, it is concluded that a bivariate probability distribution provides detailed information regarding future floods, whereas the univariate probability distribution is insufficient in providing extensive information regarding future flows.

1. Introduction

Flood is a major natural catastrophe. Pakistan being a South Asian country is a flood prone area. Floods in the Indus River of Pakistan and its tributaries have frequently affected regions of Pakistan. Floods of 1928, 1929, 1955, 1957, 1959, 1973, 1976, 1988, 1992, 1995, 1996, 1997 and 2010 are the most severe incidents that resulted in loss of lives and infrastructure. Floods have been continuously recorded since the establishment of flood warning and forecasting mechanism in 1947. It is therefore crucial to gauge the flood risk in the flood-affected areas. Flood risks are predicted through the probability of event occurrence and the related consequences [1]. It can be demonstrated that the natural irregularities of geographical system and variation in the complex socioeconomic features are the root cause of risk and uncertainty in water resources [2].

Meteorological and hydrological parameters are often used for flood risk analysis. However, the risk of flood is also estimated by GIS (Geographical Information System) technique [3, 4]. Recently a study was conducted to estimate the flood risk along the River Indus in Pakistan by applying suitable probabilistic distributions (i.e. Weibul distribution, Pearson type-3 analysis) to the flood peak values of the observed data by which the associated return periods have been obtained for various Pakistani dams [5].

Flood frequency analysis used to concentrate only on the flood peak values for the analysis of a flood event, which is insufficient to acquire extensive knowledge about future flows. Exhaustive information regarding the causes of a flood event, other aspects of flood such as volume, peak, duration and shape of hydrograph are also required to obtain authenticated results for solving many hydrological problems. Some researchers have conducted such type of studies in which they have tested the flood event as a multivariate event by deriving the relationship among flood volume, peak and duration using several methodologies and distributions [6]. The Multivariate Partial Duration Series method (MPDS) is also applied to infer that the joint distribution of flood durations and peaks are assumed to be based on assumptions that; (i) both flood durations and peaks are exponentially

* Corresponding author

distributed and (ii) the conditional distribution of flood peaks and flood durations are normal [7].

Gumbel distribution is a statistical approach that is mostly used to predict extreme events such as flood [8-11]. The Gumbel mixed model has been applied to the recorded flood data of the Ashuapmushuan River basin located in the province of Quebec, Canada. [12, 13]. Therefore, this article also suggests the application of bivariate extreme value distribution such as the Gumbel mixed model with the marginal of Gumbel for representing the joint probability distribution of flood volumes and peaks along with flood durations and volumes. The joint distributions, the conditional probability functions and the related return periods can readily be obtained by the marginal distributions of the associated random variables from the Gumbel mixed model.

Flood frequency analysis of Tarbela dam is carried out in this groundbreaking piece of research using historical data to verify the suggested model. Tarbela Dam on the Indus River is situated in the province of Khyber Pakhtunkhwa, Pakistan. The joint probability distributions of flood volumes and peaks, and the joint probability distributions of flood durations and volumes are analyzed using the annual maximum upstream values.

2. Materials and Methods

Flood frequency distributions may have many forms related to the equation used for performing statistical analysis. Gumbel's distribution is one of the statistical approaches that are mostly used to analyze flood data. The description is as follows.

2.1 Gumbel Mixed Model

Gumbel [14] originally proposed the Gumbel mixed model with the marginal of Gumbel. This model offers the joint cumulative distribution function as follows:

$$F(y,z) = F(y).F(z)\exp\left\{-\theta.\left[\frac{1}{lnF(y)}+\frac{1}{lnF(z)}\right]^{-1}\right\} \quad (0\leq\theta\leq1) \qquad (1)$$

Where F(y) and F (z) represents the marginal distribution functions of the two random variables Y and Z [y for peak discharge and z for volume second time y for volume and z for duration] and are as follows:

$$F(y) = \exp[-\exp(-y)] \qquad (2a)$$

$$F(z) = \exp[-\exp(-z)] \qquad (2b)$$

In addition, θ (0 ≤ θ ≤ 1) is a parameter, which describes the relationship between Y and Z. The formula for θ is introduced for bivariate extremes [15, 16].

$$\theta = 2\left[1-cos\left(\pi\sqrt{\frac{\rho}{6}}\right)\right] \quad \text{for } 0\leq\rho\leq2/3 \qquad (3)$$

Where ρ is a Pearson's correlation coefficient and is given by

$$\rho = \frac{E[(Y-\mu_y)(Z-\mu_z)]}{\sigma_y\sigma_z} \qquad (4)$$

Here (μ_y, σ_y) and (μ_z, σ_z) are mean and Standard deviations of random variables Y & Z. When ρ=0, the related parameter θ becomes zero. This is an independent case and the bivariate distribution then decomposes into the multiple of two marginal distributions as follows:

$$F(y,z) = F(y).F(z) \qquad (5)$$

When ρ= 2/3, the related parameter θ approaches its maximum limit i.e. equal to one. The model is not valid when ρ> 2/3 i.e. the model is employed when the correlation coefficient of the joint distribution range is 0≤ρ≤2/3.

The Gumbel form or the (Extreme value type 1 distribution) is obtained by setting the marginal distribution of the two random variables, we have:

$$F(y) = \exp\left[-\exp\left(-\frac{y-u_y}{\lambda_y}\right)\right] \qquad (6a)$$

$$F(z) = \exp\left[-\exp\left(-\frac{z-u_z}{\lambda_z}\right)\right] \qquad (6b)$$

Here u and λ are scale and location parameters of Gumbel distribution. Eq. (1) is used to derive the joint probability density function (pdf) as follows:

$$f(y,z) = \frac{\partial^2 F(y,z)}{\partial y\partial z} = \frac{1}{\lambda_y\lambda_z}F(y,z)e^{-a}$$

$$\times\left\{1-\theta\frac{e^{\frac{2(y-u_y)}{\alpha_y}}+e^{\frac{2(z-u_z)}{\alpha_z}}}{b^2}+2\theta\frac{e^{2a}}{b^3}+\theta^2\frac{e^{2a}}{b^4}\right\}$$

$$(0\leq\theta\leq1) \qquad (7)$$

Where,

$$a = \frac{y-u_y}{\lambda_y}+\frac{z-u_z}{\lambda_z} \qquad (8a)$$

$$b = e^{\frac{y-u_y}{\lambda_y}}+e^{\frac{z-u_z}{\lambda_z}} \qquad (8b)$$

Since Eq. (1) is the cumulative distribution function (cdf) of Y and Z by which the conditional probability distribution of Y given Z can also be obtained and represented as below:

$$F(y|Z\leq z) = \frac{F(y,z)}{F(z)}$$

$$F(y|Z\leq z) = F(y)exp\left\{-\theta\left[\frac{1}{lnF(y)}+\frac{1}{lnF(z)}\right]^{-1}\right\} \qquad (9a)$$

Similarly, the equivalent formula can express the conditional probability distribution of Z given Y as follows:

$$F(z|Y \le y) = \frac{F(y,z)}{F(y)}$$

$$F(z|Y \le y) = F(z)exp\left\{-\theta\left[\frac{1}{lnF(z)} + \frac{1}{lnF(y)}\right]^{-1}\right\} \quad (9b)$$

The associated return periods T_y and T_z which exceed the particular values of Y and Z are represented as below:

$$T_y = \frac{1}{1-F(y)} \qquad (F(y) = \Pr[Y \le y]) \quad (10a)$$

$$T_z = \frac{1}{1-F(z)} \qquad (F(z) = \Pr[Z \le z]) \quad (10b)$$

Here Pr shows probability.

Same principle is followed to represent the joint return period $T_{(y,z)}$ of Y and Z related to the event (Y>y, Z>z or Y>y and Z>z i.e. at least one value of y and z is exceeded) as follows:

$$T_{(y,z)} = \frac{1}{1-F(y,z)} \quad (F(y,z) = \Pr[Y \le y, Z \le z]) \quad (10c)$$

The conditional return periods of Y given Z ≤ z and Z given Y ≤ y are, respectively, expressed as follows:

$$T_{(y|z)} = \frac{1}{1-F(y|z)} \quad (F(y|z) = \Pr[Y \le y|Z \le z]) \quad (10d)$$

$$T_{(z|y)} = \frac{1}{1-F(z|y)} \quad (F(z|y) = \Pr[Z \le z|Y \le y]) \quad (10e)$$

3. Results and Discussion

3.1 Description of Tarbela Dam

Tarbela Dam on the River Indus is considered the largest, in the category of world's largest earth filled dam and second largest as a structural volume. It is situated in the district of Haripur, Hazara Division, province of Khyber Pakhtunkhuan, approximately 50 km northwest of Islamabad, Pakistan. The dam is 148 m high above the riverbed. The dam forms the Tarbela Reservoir, which is 8.5 km long with the surface area of 250 square kilometers and holds 14.3 cubic kilometers of water. The dam was designed for storing water from the River Indus. Sources of the River Indus are rainwater and melted water from glacier through the Himalayas. High flow season is in the Kharif season due to snow melt and heavy rainfall to the river runoff. Therefore, the annual maximum upstream flow has been recorded in the Kharif season (i.e. 6 months from April to September) for 35 years (1977-2012).

3.2 Flood Event Characteristics

The fundamental flood event characteristics are flood volume (*V*), flood peak (*Q*) and flood duration (*D*). The components of hydrograph play a pivotal role to estimate the characteristics of a flood event. The two main components of hydrograph are base flow and direct runoff as shown in Fig. 1.

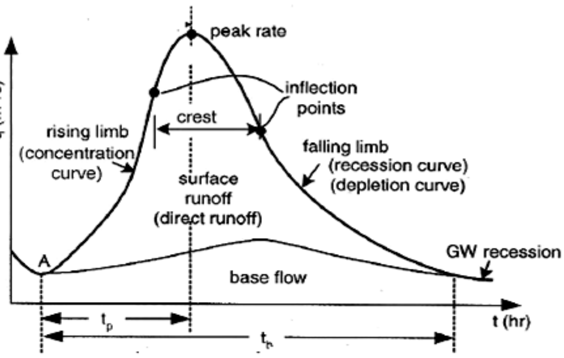

Fig. 1: Flood event characteristics

Where q is discharge flow, t_p is time to peak, t_b is base time, and GW is ground water. Base flow is the amount of water in a stream that comes from ground water discharge or seepage of large lakes whereas direct runoff is the amount of water due to rainfall, snowmelt or other sources that flow over the ground surface directly into the streams, rivers and lakes. Since there is a difference in the properties of direct runoff and base flow, a slope of hydrograph changes significantly due to the transition from direct runoff to base flow. The start of increasing limb and the end of decreasing limb are generally the time limits of a flood. Therefore, the flood duration (*D*) is obtained by identifying the initial and final days of flood runoff. The day when there is an instant rise in hydrograph is said to be an initial day of flood runoff and the day when recession limb of hydrograph starts to flatten is considered as a final day of flood runoff. These criteria are used to estimate the initial day (ID_k) and the final day (FD_k) of flood runoff for the k^{th} year to generate the series of flood duration D_k as follows:

$$D_k = FD_k - ID_k \quad (11)$$

The series of flood volume V_k is computed using the formula

$$V_k = \sum_{l=ID_k}^{FD_k} q_{kl} - \frac{1}{2}(q_{ki} + q_{kf}) \quad (12)$$

Where, q_{kl} is the recorded value of daily upstream flow on l^{th} day in k^{th} year, q_{ki} and q_{kf} are respectively the recorded values of daily upstream flow on the initial day (ID_k) and the final day (FD_k) of flood runoff in *k*th year. The series of flood peak Q_k is generated as follows:

$$Q_k = q_{mkl} \quad (13)$$

Where, q_{mkl} is the maximum-recorded value of daily upstream flow on l^{th} day in k^{th} year.

3.3 Marginal of Flood Event Characteristics for Empirical Probabilities

It has been revealed by the work cited in literature [17-20] that the Grigorten formula for Extreme Value1 quintile estimation is not biased. The Grigorten formula is used to estimate the probability of non-exceedance. The works by Gringorten (1963) and Cunnane (1978), and Guo (1990) have demonstrated that the

$$P_i = \frac{i-0.44}{N+0.12} \qquad (14)$$

Where, P_i represents the cumulative frequency which is the probability that a given value is smaller than the k^{th} smallest value among N observations. The computer package (M.S Excel) is used to execute the chi-square test for the Gumbel distribution to test the goodness of fit. The test shows that the Gumbel distribution can represent all three variables of a flood event characteristics i.e. flood volume, duration and peak. Figs. 2, 3 & 4 illustrate the results of the test.

3.4. Estimation of Gumbel Parameters

The moment method (MM) is used for estimating the Gumbel parameters as follows :

$$\lambda = \frac{\sqrt{6}}{\pi} S \qquad (15a)$$

$$u = M - 0.577\lambda \qquad (15b)$$

Where, S and M represents the sample standard deviation and sample mean, respectively. The estimated standard deviation and mean of flood characteristics through sample data as well as the estimated Gumbel parameters are given in Table 1.

Table 1: Statistics and parameters of flood peak (Q), volume (V) and duration (D)

Parameters	Statistics		Parameters of Gumbel	
	M	S	λ	u
Q(m^3/s)	10323	2276	1774.9	9299.1
V(day m^3/s)	584220	120650	94070	529950
D(days)	103.1765	19.4086	15.1329	94.4448

3.5 Correlation Between Flood Event Characteristics

The correlation coefficients ρ_{VQ} between flood volumes and peaks, ρ_{DV} between flood durations and volumes and ρ_{DQ} between flood durations and peaks are computed using Eq. (4) and the values are $\rho_{VQ} = 0.1961$, $\rho_{DV} = 0.1812$ and $\rho_{DQ} = -0.5188$ respectively. Physically, a strong correlation is observed between flood volumes and peaks and between flood durations and volumes because the values of ρ_{VQ} and ρ_{DV} lie within the limit $0 \le \rho \le 2/3$ which verifies the validity of the proposed model for computing the joint distribution of two flood event characteristics.

Flood durations and peaks are assumed to be mutually exclusive events because the value of ρ_{DQ} does not lie within the limit $0 \le \rho \le 2/3$ i.e. the correlation between them is not close enough so that an inverse correlation could exist. Therefore, a bivariate extreme model can be employed for presenting the mutual behaviour of a flood event i.e. the joint distribution of flood volumes and peaks and the joint distribution of flood durations and volumes. Eq. (3) is used to compute the values of the associated parameter θ, which must occur within the range $0 \le \theta \le 1$ to describe the association between the pair of joint characteristics. The obtained values of "θ" are $\theta_{VQ} = 0.319$ between flood volumes and peaks, and $\theta_{DV} = 0.291$ between flood durations and volumes. If a strong correlation was observed between flood durations and peaks then a multivariate distribution would be employed for representing the event of a flood.

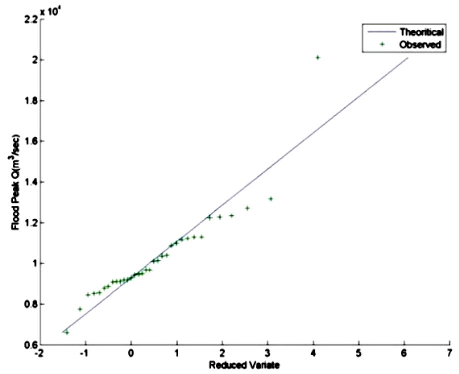

Fig. 2: Flood peak distribution

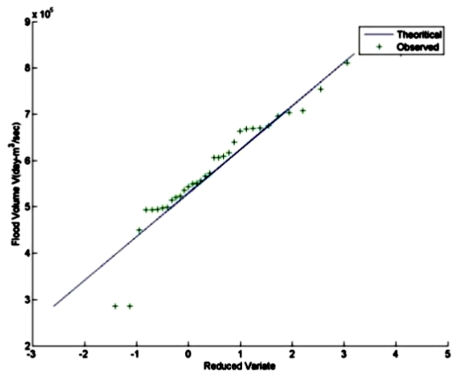

Fig. 3: Flood volume distribution

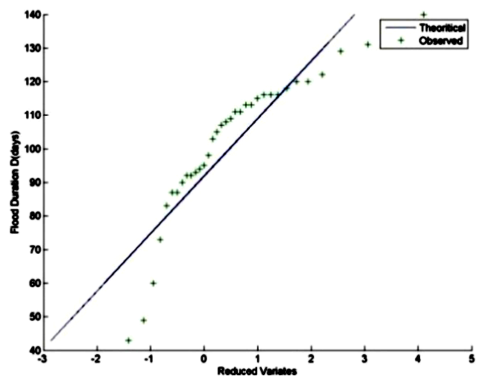

Fig. 4: Flood duration distribution

Table 2: Empirical and theoretical non-exceedance joint probabilities of flood peak (Q) and volumes (V)

Year	Q (m^3/s)	V ($daym^3/s$)	Joint probabilities		Bias	(Em−Th)/ Em (%)
			Empirical	Theoretical	Em-Th	
2004	6603	494864	0.016413	0.003039	0.013374	81.48
2001	7767	514989	0.045721	0.033913	0.011808	25.83
1980	8466	572993	0.075029	0.117675	-0.04265	-56.84
2008	8523	566450	0.075029	0.1187	-0.04367	-58.21
1985	8557	549639	0.075029	0.10829	-0.03326	-44.33
2009	8790	663366	0.162954	0.215942	-0.05299	-32.52
1984	8863	703627	0.192263	0.244675	-0.05241	-27.26
1993	9084	497279	0.045721	0.089218	-0.0435	-95.14
1979	9118	285498	0.016413	5.89E-07	0.016412	100.00
1987	9121	285498	0.045721	5.90E-07	0.04572	100.00
1981	9194	523063	0.162954	0.131523	0.031431	19.29
2007	9197	499849	0.133646	0.098782	0.034864	26.09
1991	9277	707153	0.368113	0.320549	0.047563	12.92
1997	9444	494060	0.075029	0.103173	-0.02814	-37.51
1982	9458	520536	0.221571	0.146887	0.074684	33.71
2003	9512	609261	0.368113	0.284068	0.084045	22.83
1992	9684	606395	0.397421	0.30401	0.093411	23.50
2002	9684	606395	0.397421	0.30401	0.093411	23.50
1998	10089	810745	0.543962	0.505693	0.038269	7.04
1986	10132	556956	0.309496	0.270713	0.038783	12.53
2000	10361	449897	0.075029	0.060729	0.014301	19.06
1999	10389	668821	0.543962	0.478526	0.065436	12.03
1983	10874	669252	0.573271	0.543633	0.029638	5.17
2005	10990	616747	0.514654	0.475473	0.039181	7.61
1977	11157	551234	0.338804	0.333165	0.00564	1.66
1990	11230	754634	0.719812	0.660961	0.058851	8.18
1978	11298	675399	0.661196	0.599754	0.061442	9.29
1979	11301	830409	0.807737	0.699539	0.108198	13.40
1996	12239	639673	0.573271	0.619832	-0.04656	-8.12
1998	12278	695860	0.719812	0.711796	0.008017	1.11
2006	12335	669930	0.690504	0.679617	0.010887	1.58
1989	12703	536147	0.309496	0.347363	-0.03787	-12.24
1995	13162	542947	0.338804	0.381359	-0.04255	-12.56
2010	20108	494060	0.133646	0.230787	-0.09714	-72.69

3.6 Joint Distribution Statistics of Flood Volume (V) and Peak (Q)

3.6.1 The proposed model validity

Principle of a univariate probability is followed to compute the empirical joint probabilities of two random variables. First construct a two dimensional table and arrange the variables V and Q in ascending order in it. Elements of the table in k^{th} row and l^{th} column represent the joint frequency function for two random variables and are obtained by the approach given in reference [12] as follows:

$$f(v_k, q_l) = P(V = v_k, Q = q_l) = \frac{n_{kl}}{N+0.12} \qquad (16)$$

Where, N is the number of total observed values (i.e. N = 35) and n_{kl} is the number of mutual occurrence of v_k and q_l. The theoretical joint probabilities of two random variables are obtained using Eq. (7). The empirical non-exceeding joint probabilities (cumulative joint frequencies) of flood volumes and peaks are analyzed by the formula similar to Eq. (13) for managing the format of marginal distribution according to reference [12] and is represented as:

$$F(v,q) = P(V \le v_k, Q \le q_l) = \frac{\sum_{i=1}^{k}\sum_{j=1}^{l} n_{ij}}{N+0.12} \quad (17)$$

Eq. (1) is used to compute the theoretical joint probabilities of flood volume and peak discharge. The recorded values of flood peak (in ascending order), corresponding flood volumes, and their related theoretical and empirical joint probabilities are listed in Table 2. It was observed that the difference between empirical and theoretical values is not significant and when the kolmogorov-Smirnov test was applied to test the goodness of fit of the empirical joint probabilities to the theoretical distribution at the significance level α=0.05, the critical kolmogorov-Smirnov value of D33(0.05)=0.121 was obtained. As the maximum difference between the empirical joint probabilities and the respective theoretical ones is 0.108, thus the conclusion is made that the proposed methodology is appropriate for representing joint distribution of correlated flood characteristics (i.e. volume and peak).

3.6.2 *Pdf $f(v,q)$, cdf $F(v,q)$, graphs and return period $T(v,q)$*

The joint pdf $f(v,q)$and cdf $F(v,q)$ of flood volumes and peaks are obtained using Eqs. (7) and (1) and are shown in Figs. 5(a) and (b) respectively. The joint return period $T(v,q)$ of flood volume and peak are obtained by Eqs.(1) and (10c) respectively. This parameter is pivotal for urban planners and civil engineers to determine the best locations and construction practice for new development.

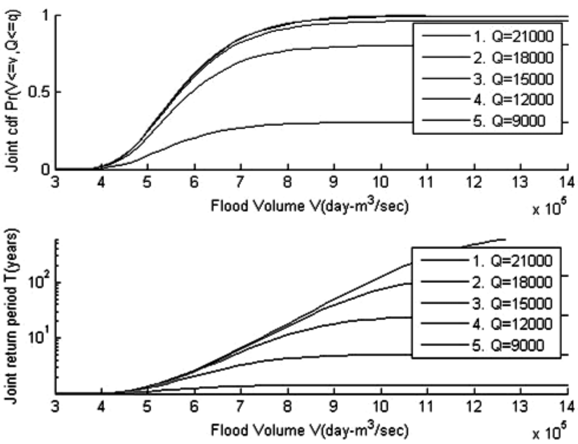

Fig. 5: (a) & (b) Joint cdf and return period of flood volume (V) and peak (Q)

3.6.3 *Conditional return period $T_{(v|q)}$ and $T_{(q|v)}$*

Mathematically, the conditional return period $T_{(v|q)}$ of flood volume given peak is obtained using Eqs. (9a) and (10d) and is graphically represented in Fig. 6. The conditional return period $T_{(q|v)}$ of flood peak given volume is obtained using Eqs. (9b), (10e), and graphically displayed in Fig. 7.

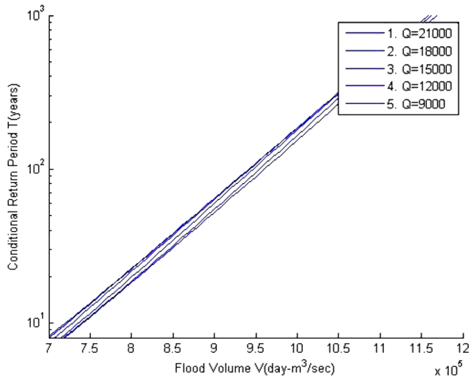

Fig. 6: Conditional return period of flood volume(V) given flood peak(Q)

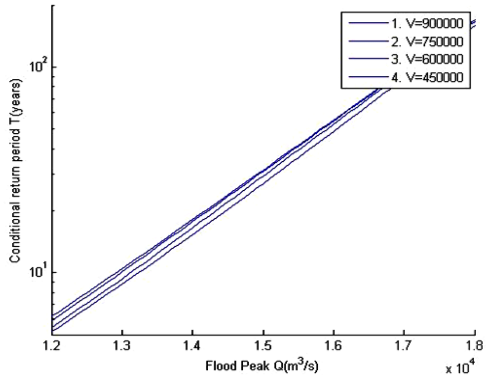

Fig. 7: Conditional return period of flood peak(Q) given flood volume(V)

It is observed from Figs. 5, 6 and 7 that the suggested model can be beneficial to solve various problems related to hydrologic engineering management, planning and design for which the univariate flood frequency analysis is insufficient. Consider an example; if a return period of a flood event is given, it is certain to have several combinations of occurrence related to flood volumes and peaks, and vice versa. These several scenarios could be beneficial to assess and analyze the risk related to various problems of hydrology such as design of spillways and flood controlling devices. Information regarding the probabilities of occurrence of flood volume given flood peak and vice versa is obtained by the proposed model.

3.7 *Joint Distribution of Flood Durations (D) and Volumes (V)*

3.7.1 *The proposed model validity*

Procedure of section 3.6.1 is followed to obtain the theoretical and empirical joint probabilities of flood volumes and durations, which are given in Table 3. It is observed from Table 3 that the differences between

Table 3: Empirical and theoretical non-exceedance joint probabilities of flood volumes (*V*) and duration (D)

Year	*D* (day)	*V* (day m^3/s)	Joint Probabilities		Bias	% Bias
			Empirical	Theoretical	Em-Th	(Em−Th)*100 /Em
2006	49	669930	0.016413	0.000000	0.016413	100.00
2009	60	663366	0.016413	0.000051	0.016362	99.69
1979	73	285498	0.016413	0.000000	0.016413	100.00
1995	83	542947	0.045721	0.063386	-0.01767	-38.64
1977	87	551234	0.045721	0.027345	0.018376	40.19
2000	87	449897	0.104338	0.108340	-0.00400	-3.84
1996	90	639673	0.133646	0.211483	-0.07784	-58.24
1978	92	675399	0.104338	0.092206	0.012132	11.63
1997	92	494060	0.280188	0.267787	0.0124	4.43
2010	92	494060	0.104338	0.092206	0.012132	11.63
2007	93	499849	0.133646	0.106772	0.026874	20.11
2002	94	606395	0.221571	0.258793	-0.03722	-16.80
1998	95	695860	0.368113	0.340197	0.027916	7.58
1981	98	523063	0.162954	0.184969	-0.02201	-13.51
2003	103	609261	0.280188	0.405613	-0.12543	-44.76
2005	105	616747	0.309496	0.445538	-0.13604	-43.96
1993	107	497279	0.133646	0.178986	-0.04534	-33.93
1989	108	536147	0.221571	0.291458	-0.06989	-31.54
1987	109	285498	0.045721	0.000001	0.04572	100.00
1985	111	549639	0.309496	0.348920	-0.03942	-12.74
1986	111	556956	0.368113	0.369959	-0.00185	-0.50
1980	113	572993	0.397421	0.428489	-0.03107	-7.82
1991	113	707153	0.661196	0.666237	-0.00504	-0.76
2008	115	566450	0.397421	0.422209	-0.02479	-6.24
1979	116	830409	0.162954	0.199573	-0.03662	-22.47
1984	116	703627	0.719812	0.696949	0.022864	3.18
2004	116	494864	0.778429	0.764984	0.013445	1.73
2001	118	514989	0.250879	0.269008	-0.01813	-7.23
1990	120	754634	0.543962	0.561470	-0.01751	-3.22
1992	120	606395	0.837046	0.776901	0.060145	7.19
1999	122	668821	0.690504	0.702468	-0.01196	-1.73
1983	129	669252	0.719812	0.739539	-0.01973	-2.74
1982	131	520536	0.280188	0.312839	-0.03265	-11.65
1998	140	810745	0.954279	0.913917	0.040362	4.23

theoretical and empirical values are good. The kolmogorov-Smirnov test was applied to test the goodness of fit of the empirical joint probabilities to the theoretical distribution. At the significance level of α=0.05, the critical kolmogorov-Smirnov value is $D_{33}(0.05)=0.152$. The maximum difference between the empirical joint probabilities and the respective theoretical ones is 0.0601. Thus, the conclusion is made that the proposed methodology is appropriate for representing joint distribution of correlated flood characteristics (i.e. volume and duration).

3.7.2 *The pdf $f(v,d)$, cdf $F(v,d)$ graphs and return periods $T(v,d)$, $T_{(v|d)}$ and $T_{(d|v)}$*

Joint *pdf* $\mathrm{f}(\mathrm{v},\mathrm{d})$, *cdf* $\mathrm{F}(\mathrm{v},\mathrm{d})$ and associated return periods $\mathrm{T}(\mathrm{v},\mathrm{d})$ of flood volumes and durations were obtained by following the procedure outlined in section 3.6.2. Graphically, the joint return period of flood volumes and durations is illustrated in Fig. 8. The conditional return period $T_{(d|v)}$ of flood duration given flood volume and the conditional return period $T_{(v|d)}$ of flood volume given flood duration are graphically presented in Figs. 9 and 10, respectively. These results verify the benefits of the suggested model as concluded in section 3.6. It is essential to note that apart from flood peak, flood volume and flood duration are also the main characteristics of flood event on which damage caused by a flood depends. The results obtained by the proposed model can be used for calibrating the functions of flood damage, which are served as estimation tools to local agencies and insurance companies in the studies of pre-flood, or post flood as demonstrated in reference [21].

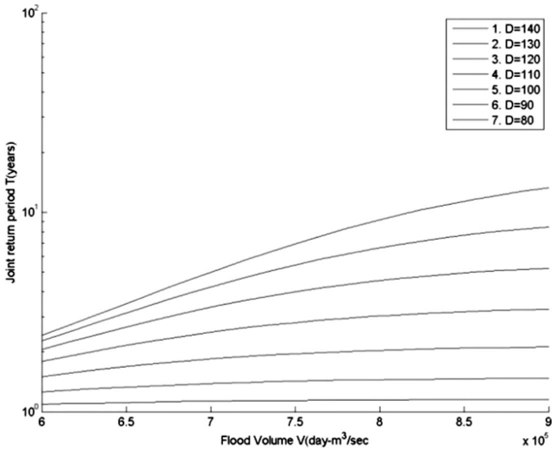

Fig. 8: Plot of Joint return period of flood volumes(V) and durations(D)

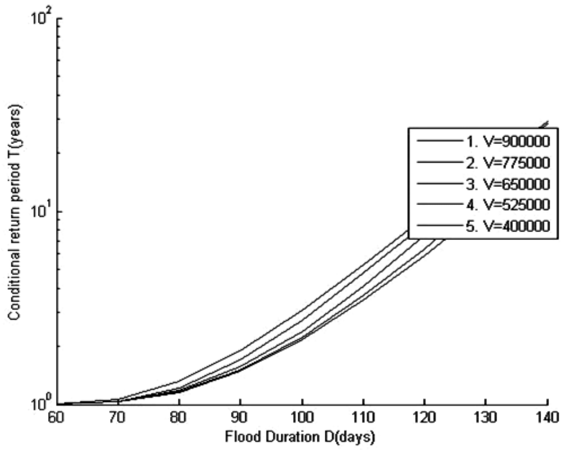

Fig. 9: Conditional return period of flood duration given volume (V)

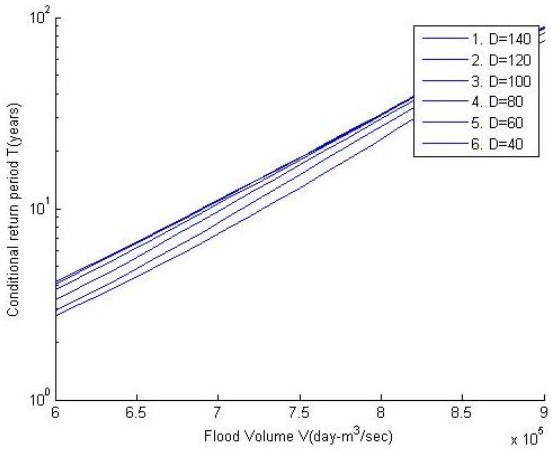

Fig. 10: Conditional return periods of flood volume given duration

4. Conclusion

According to Gumbel model, if this distribution can approximately represent the marginal distribution of the two random variables, the joint probability distributions, the conditional distribution and the related return periods of these variables. Parameters of the suggested model can easily be obtained from the sample data which are then associated to the marginal distributions of random variables. The model is tested and validated by Tarbela Dam data.

It is observed that the values of theoretical and empirical joint probabilities do not vary significantly which verified that the proposed model is valid for representing the joint distribution of correlated flood characteristics i.e. flood volumes and peaks, also flood durations and volumes. It is crucial to note that the proposed model, which is not possible, provides extensive information by univariate analysis such as joint and conditional return period of correlated flood characteristics.

The volume and peak discharge worked out here can be used to address issues related to engineering design of hydraulic structure and management. Therefore, if a return period of flood event is given, it is certain to acquire several-combined occurrence of flood volumes and peaks, and vice versa. The finding of this study suggests increasing the design capacity of dam and constructing new dams with greater return period.

Acknowledgement

Authors would like to acknowledge The Dean, Faculty of Science(University of Karachi) for the partial financial funding and thank Ms. Ayman Iqbal for her kind assistance.

References

[1] K, Smith, “Environmental hazard”, Rutledge, London, 1996.

[2] Z. Kaczmarek, “The impact of climate variability on flood risk in Poland”, Risk Anal., vol. 23, no. 3 , pp. 559-566, 2003.

[3] G. Khan, “Flood hazard assessment and mitigation along River Indus from Chashma Barrage to Sukkur Barrage using satellite image.” M. Phil thesis, Institute of Space and Planetary Astrophysics”, University of Karachi, 2007.

[4] T. Tahirkaili and F. Nawaz, “Role of GIS and RS for flood hazard management in Pakistan: A case study of Jhelum, Pakistan”, MAPAsia, (http://www.gisdevelopment.net/proceedings/mapasia/003/disaster/index.htm , 2003.

[5] B. Khan, M. J. Iqbal and M.A. Yousufzai, “Flood risk assessment of River Indus of Pakistan”, Arabian Journal of Geosciences, vol. 4, pp. 115-122, 2011.

[6] F. Ashkar, “Partial duration series models for flood analysis”. Ph.D. thesis, Ecole Poly technique of Montreal, Montrea l, Canada, 1980.

[7] F.N. Correia, “Multivariate partial duration series in flood risk analysis”, V.P. Singh (Ed.), Hydrologic Frequency Modeling, Reidel, Dordrecht, pp. 541–554, 1987.

[8] E.J. Gumbel, "Statistics of extremes", Columbia University Press, New York, 1958.

[9] P. Todorovic, "Stochastic models of floods", Water Resource Research, vol.4, no. 2 , pp. 345–356, 1978.

[10] E. Castillo, "Extreme value theory in engineering", 1st edn., Academic Press, New York, 1988.

[11] W.E. Watt, K.W. Lathem, C.R. Neill, T.L. Richard and J. Rousselle, "Hydrology of floods in Canada: A Guide to planning and design", National Research Council of Canada, 1989.

[12] S. Yue, B.M.J Ouarda, B. Bobée, P. Legendre and P. Bruneau, "The Gumbel mixed model for flood frequency analysis", Journal of Hydrology, vol. 226, pp. 88-100, 1999.

[13] S. Yue, "The Gumbel mixed model applied to storm frequency analysis", Water resources Management, vol. 14, no. 5, pp. 377-38, 2000.

[14] E.J. Gumbel, "Multivariate extreme distributions", Bulletin of the International Statistical Institute, vol. 39, no. 2, pp. 471–475, 1960.

[15] J.T.D. Oliveria, "Bivariate extremes: extensions", Bulletin of the International Statistical Institute, vol. 46 , no. 2, pp. 241-251, 1975.

[16] J.T.D. Oliveria, "Bivariate extremes: models and statistical decision", Technical Report No. 14, Center for Stochastic Processes, Department of Statistics, University of North Carolina, Chapel Hill, NC, USA, 1982.

[17] I.I. Gringorten, "A plotting rule for extreme probability", Journal of Geophysical Research vol. 68, no. 3, pp. 813–814, 1963.

[18] C. Cunnane, "Unbiased plotting positions: A Review", Journal of Hydrology, vol. 37, no. 3/4, pp. 205-222, 1978.

[19] S.L. Guo, "A discussion on unbiased plotting positions for the general extreme value distribution", Journal of Hydrology, vol.121, no.1, pp. 33-44, 1990.

[20] V. Fortin, Bernier, B. Bobée, Determination des crues de conception-Rapport final du project C3. INRS-Eau, Rapport de recherché confidential no. R-532, pp. 103, 1998.

[21] T.B.M.J. Ouarda, N.El-Jabi and F. Ashkar, "Flood damage estimation in the residential sector", Water Recourses and Environmental Hazards: Emphasis on Hydrologic and Cultural insight in the Pacific," AWRA Technical Publication series, TPS-95-2, pp. 73-82, 1995.

[22] E. J. Gumbel, Statistics of Extremes. Columbia University Press, New York, USA, 1958.