Regional Flood Frequency Analysis Using Computer Simulations

Anusha M[1], Surendra H J[2]

[1]Asst. Professor, Dept. of Civil Engineering, Sri Venkateshwara College of Engineering, Bangalore, India
[2] Associate Professor, Dept. of Civil Engineering, ATRIA IT, Bangalore, India.

Abstract— *Different probability distribution methods were employed to determine the flood frequency analysis using computer simulations. Many probability distributions including Gumbel, lognormal, log-Pearson type iii, General Extreme value have been tried to fit the data. The length of record for most of the stations is over 10 years (chosen from 1956 0nwards). The data was procured from J.R.M.Hosking for various project sites. The common time period of 1956 onwards has been chosen only to avoid the effect of interception of basin due to construction of storage reservoir and was also subject of flood data. The best fitting distribution works out to be General Extreme Value Distribution. Gumbel's distribution ranks poorly among different probability distributions. A trial version of probability software is used to evaluate the best fit distribution and parameters of distribution.*

Keywords— Gumbel distribution, General Extreme value, Log normal, Log Pearson distribution.

I. INTRODUCTION

Natural disaster like Flood is one of the most serious problems that are affecting countries worldwide. Many natural and Un-natural factors affects the flood, monitoring the factors responsible for flood is very important to minimize the impact and damage. Flood real time monitoring system and flood forecasting model is difficult to develop, due to a large amountof data required are difficult to obtain. In the present study, length of record for most of the stations is over 10 years (chosen from 1956 0nwards) and the data was procured from J.R.M Hosking for various project sites.

Many papers discuss the estimation of flood frequency using different distribution techniques.[3] used different methods were employed in applying probability distribution in hydrology.[4] develops flood Model to analyze the maximum water level and drainage density using Gumbel distribution. [5] analyzed the frequency ofNyanyadzi River floods using the Gumbel distribution.

Latest CWC guidelines, spelled out in the manual on estimation of design flood and USGS guidelines are used for analysis.In the present study many distribution techniques such as Gumbel distribution, Gumbel extreme value distribution, Log normal and Log Pearson distribution techniques were adopted to find the frequency of flood occurrence and best method is identified.

II. METHODOLOGY

The data has been checked for randomness and corrected for high and low outlier. Missing data is generated by correlation of flood peaks. Correction for trend has not been exercised because of absence of long term data for most of the stations. Each catchment is assumed to generate its own peak flood without intervention.

Gumbel Extreme value probability distribution by variates was adopted for frequency analysis at the initial stage assuming that most of the stations follow this type of distribution which eventually has proved otherwise. Excel spreadsheet simulation models and probability modeling tools are used for the analysis of the probability statistics.

Many probability distributions including Gumbel, lognormal, log-pearson type iii, General Extreme value have been tried to fit the data. The best fitting distribution works out to be General Extreme Value Distribution. Gumbels's distribution ranks poorly among different probability distributions. A trial version of probability software is used to evaluate the best fit distribution and parameters of distribution. The methods of moments and maximum likelihood are used for evaluation. The parameters obtained from easy fit for distributions like Gumbels (largest EV), lognormal(2p), General extreme value has been considered in working out the flood quintiles for various return periods like 2,2.33,10,25,50,100,200 and 500 years. The different distributions give closer value at small values of return periods and vary considerably at large return periods.

The station-wise, distribution-wise graphs depict the variations. Further, distribution- wise graphs for various

https://dx.doi.org/10.22161/ijaers.4.1.37 ISSN: 2349-6495(P) | 2456-1908(O)

stations also indicates the flood values at different return periods.

A) GUMBEL DISTRIBUTION

The annual flood peak of a catchment area forms the annual series. The data is arranged in the decreasing order of the magnitude and the probability peak periods of flood being equal or exceeding is estimated by the following equation.

P=m/N+1, where,

P=probability of exceedance,

m= order number.

N=total no of events.

The recurrence interval T also called return periods of frequency is calculated as,

$$T=1/P$$

Chow has shown that most frequency distributions applicable to hydrologic study can be expressed by the following general equation of hydrologic frequency analysis.

$X_t = \bar{x} + K\sigma$ where,

Xt = value of the variate of a random hydrologic series in the return period T.

$\bar{x}$ = the mean of the variate

σ = standard deviation of the variate

K = frequency factor which depend upon the return period.

B) LOG NORMAL DISTRIBUTION

Log trans series data to be used.

Z=norm lnV(F,0,1)

F=1-(1/T)

$X(f)=e^{(\log mean + \log stdev*Z)}$

C) LOG PEARSON DISTRIBUTION

F=1-(1/T)

Z=Norminv (F,0,1)

K=log skew/6

$Kt=Z+(Z^2-1)*K+(Z^3-6K)*K^2/3-(Z^2-1)*K^3+Z*K^4+K^5/3$

$X(f)=e^{(\log mean + \log stdev(Kt))}$

D) GENERALIZED EXTREME –VALUE DISTRIBUTION:

Parameters (3); ξ (location), α (scale),k (shape)

Range of x: $-\infty < x \le \xi + \alpha/k$ if k>0; $-\infty < x < \infty$ if k=0; $\xi + \alpha / k \le x < \infty$ if k<0

$f(x) = \alpha^{-1}*e^{-(1-k)y-e-y}$ $y=\{-k^{-1}*\log\{1-k(x-\xi)/\alpha\}$,k≠0

$y=\{(x-\xi)/\alpha.$ k=0

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

$x(F)=\{\xi + \alpha\{1-(-\log F)^k\}/k$ k≠0

$x(F)=\{\xi - \alpha*\log(-\log F)\}$ k=0

Special cases k=0 is the gumbel distribution, k=1 is a reverse exponential distribution on the interval $\xi \le x \le \xi + \alpha$.

The three parameters is given by,

$k= 7.8590c + 2.9554 c^2$

where, c= (2+ (3+t3))-(log 2 / log 3)

$\alpha= L2*k / \{(1-2^{-k})\ \beta\ (1+k)\}$

$\xi = L1 - \alpha\ \{1- \beta\ (1+k)\} / k$

III. RESULT AND DISCUSSION

Table.1: OUTLIER KnVALUES

Sample Size	Kn value
10	2.036
15	2.247
20	2.385
25	2.486
30	2.563
35	2.628
40	2.682
45	2.727
50	2.768
55	2.804
60	2.837
65	2.866
70	2.893

Kn =0.408 ln (n) + 1.158 (for 10 to 100)

Kn =0.352 ln (n) + 1.394 (for 60 to 100)

Kn =0.443 ln (n) + 1.040 (for 10 to 50)

Kn =0.364 ln (n) + 1.342 (for 50 to 100)

Rule a) High outlier must be greater than the maximum value.

b) Low outlier must be lesser than the minimum value.

Since it is difficult to show all the 44 stations, hence flood value for different return periods is shown only for four stations. We can also compare the flood values for different distributions.

https://dx.doi.org/10.22161/ijaers.4.1.37 ISSN: 2349-6495(P) | 2456-1908(O)

Table.2: FLOOD VALUE BY GUMBEL

T \ STATIONS	1	2	3	4
2	1537	1321	1431	1350
2.33	1570	1380	1552	1409
5	1712	1640	2076	1668
10	1827	1851	2503	1879
25	1973	2118	3043	2146
50	2081	2316	3443	2343
100	2189	2512	3840	2540
200	2296	2708	4236	2735
500	2437	2966	4758	2993

Table.3: FLOOD VALUE BY LOGNORMAL

T \ STATIONS	1	2	3	4
2	1558	1337	1425	1366
2.33	1594	1399	1535	1428
5	1733	1655	2027	1684
10	1832	1850	2438	1879
25	1944	2084	2967	2111
50	2020	2251	3369	2276
100	2090	2412	3777	2435
200	2157	2569	4193	2591
500	2189	2645	4339	2666

Table.4: FLOOD VALUE BY LOG PEARSON

T \ STATIONS	1	2	3	4
2	1560	1324	1322	1299
2.33	1596	1385	1427	1359
5	1734	1651	1982	1658
10	1831	1863	2574	1954
25	1939	2128	3554	2412
50	2012	2324	4496	2827
100	2079	2520	5665	3320
200	2142	2717	7121	3910
500	2221	2982	9623	4882

Table.5: FLOOD VALUE BY GUMBEL(EXTREME VALUE) DISTRIBUTION

T \ STATIONS	1	2	3	4
2	1563	1314	1299	1350
2.33	1601	1377	1394	1410
5	1746	1651	1894	1659
10	1842	1876	2427	1850
25	1943	2165	3315	2076
50	2005	2383	4174	2233
100	2056	2600	5245	2381
200	2100	2820	6582	2521
500	2148	3113	8878	2694

https://dx.doi.org/10.22161/ijaers.4.1.37 ISSN: 2349-6495(P) | 2456-1908(O)

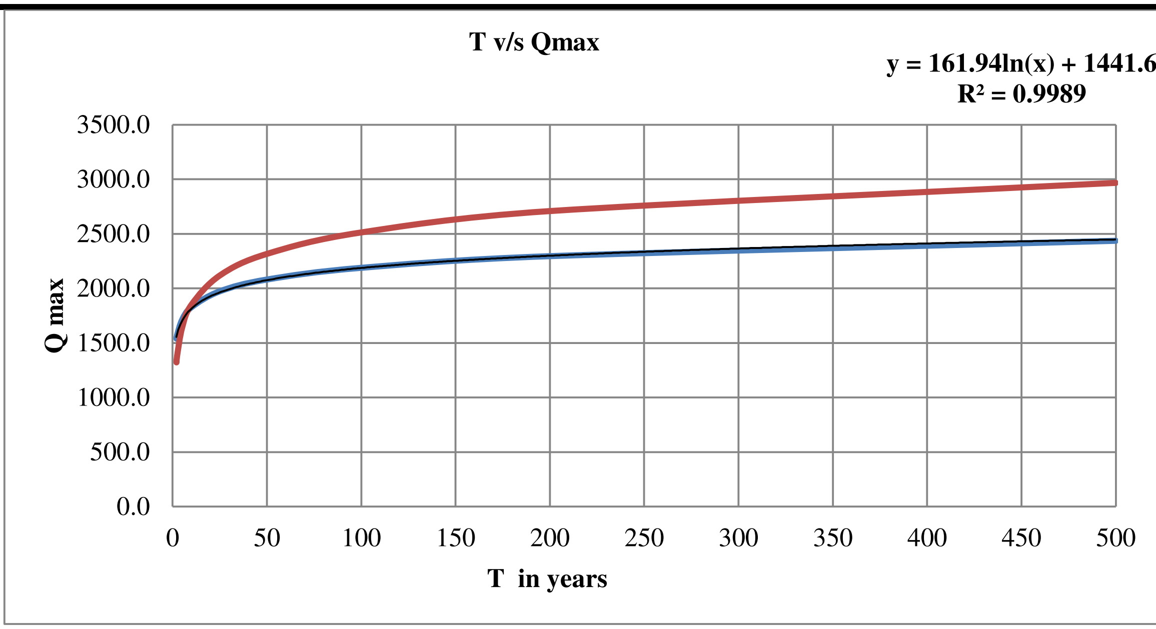

Fig.1: COMPARISON OF T v/s Qmax FOR GUMBEL AND LOGNORMAL DISTRIBUTION

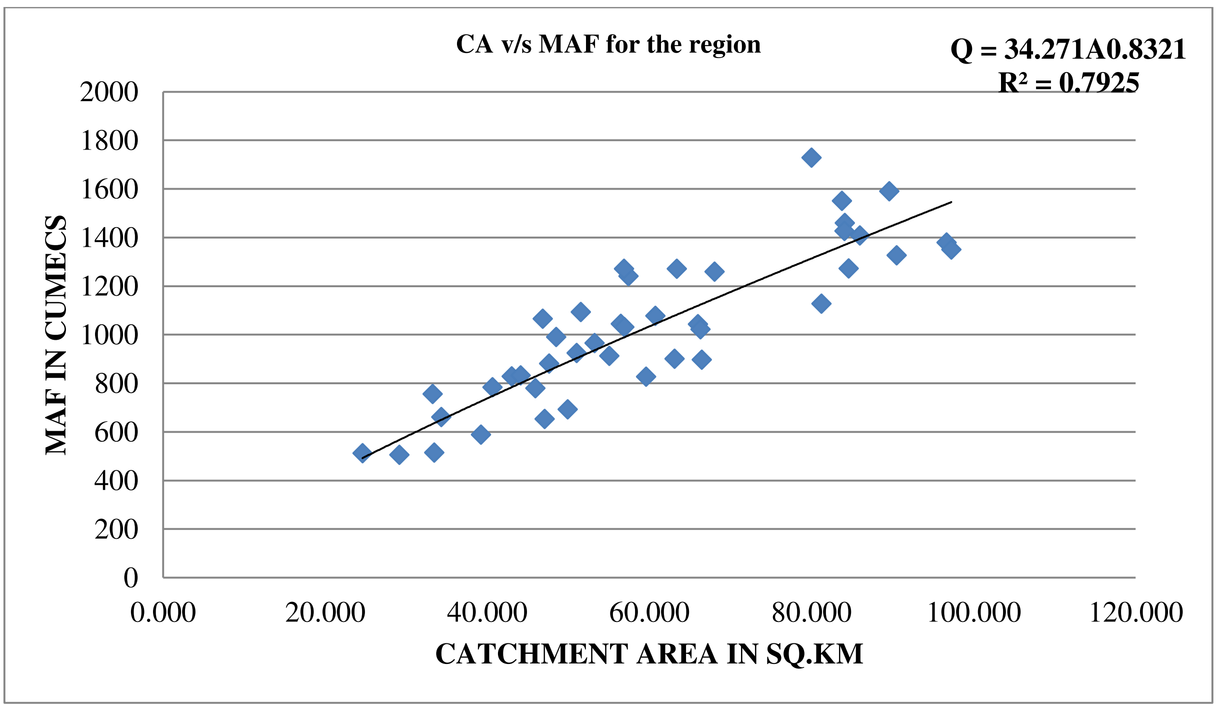

Fig.2: Regional Flood Equation arrived for the region

IV. CONCLUSIONS

a) Gumbel distribution does not too close as compare to lognormal and log Pearson distributions. Only GEV and LN3 distributions fit well for most of the sites.

b) Gumbel distribution used for modeling gives different magnitudes of flood quantiles at high return periods.

c) The regional flood equation arrived at for region is **$Q=34.271A^{0.8321}$ with $R^2=0.7925$**

d) Accuracy of the predicted flood values depends primarily on the accuracy of the data. This study being based on annual maximum stream flow data available in Hosking book.

e) This flood values obtained will be different for different methods. It is based upon the terrain, land

https://dx.doi.org/10.22161/ijaers.4.1.37 *ISSN: 2349-6495(P) | 2456-1908(O)*

cover and land use pattern. With the available data these methods will follow their own pattern to determine the flood value.

f) Discordancy test is conducted for all the stations, all stations are having the value less than three except one station which is having a value more than three, which will be discorded.

ACKNOWLEDGEMENT

Authors are grateful to Sri C.S.NAGENDRA Research officer and Sri. K.R.SATHYANARAYANA Assistant research officer, K.E.R.S, Krishnaraja Sagara, Government of Karnataka, for their valuable guidance encouragement and co-operation in all stages of work.

ABBREVIATIONS

CA: Catchment area, GEV: Generalized extreme value, m:order number, ξ:Location parameter,N:Number of observation, P or F:Probability of exceedance, PDF:Probability Density Function,2P:Two parameters,Q:Flood peak,R:Regression coefficient, σ: Standard deviation, T: Return period, USGS: United State Geological Survey, k: Shape parameters.

REFERENCES

[1] Developing a methodology for estimation of design flood in southern region of Karnataka, modified report research scheme applied to flood control project, sponsored by CBIP KERS, K.R.SAGAR.

[2] D van Bladeren.,P K Zawada and D Mahlangu. “Statistical Based Regional Flood Frequency Estimation Study for South Africa Using Systematic, Historical and Palaeoflood Data, Pilot Study – Catchment Management Area 15”. WRC Report No 1260/1/07, ISBN 078-1-77005-537-7,March 2007.

[3] Gh.Khosravi.,A.Majidi., A.Nohegar.” Determination of suitability probability distribution for annual mean and peak discharges Estimation (case study : Minab river-Barantin Gage, Iran). International Journal of Probability and statistics, 2012, 1(5): 160-163.

[4] Kantima M., Pattorawait P. “Mathematical modeling for flood forecasting in the chi river: Amphur Muang, Khon Kaen, Thailand case”. International conference on computer science and Information Technology, December 2011.

[5] Never Mujere. “Flood frequency analysis by Gumbel Distribution”. International Journal on computer science and Engineering, vol:3, no-7, july-2011.

[6] Regional frequency analysis by HOSKING and WALLIS.

[7] Satish Chandra, Director, Regional Flood Frequency Analysis, National Institute of Hydrology, Jal Vigyan Bhavan, Roorkee.

[8] Subramanya.K,(2006) “Engineering Hydrology” ii Edition, Tata McGraw Hill Publishing Company Ltd, New Delhi.

[9] Ven Te Chow “Applied hydrology McGraw hill Editions, New Delhi.