PROBABILITY DISTRIBUTION OF FLOOD FREQUENCY ANALYSIS

Vrushali Wagh[1], Purva Vaidya[1], Vaishnavi Akhare[1], Harsh Sureka[1], Prof. Dr. Shantini Bokil[2]

[1]Research students, Dept. of Civil Engineering, MITCOE, Pune
[2]Professor and Head, Department of Civil Engineering, MITCOE, Pune

---

***ABSTRACT**: The flood frequency assessment is a very important parameter for water resource management. Every pre-monsoon period the analytical evaluations help to streamline the water management in monsoon. So the selection of an appropriate probability distribution for describing flood frequency at a particular site is the governing criteria for assessment, highlighting the need for an optimum solution. This paper highlights the scope of a mathematical model or statistical tool and it's software applications, to give optimum results for the appropriate method of flood frequency. Highlighting these aspects will surely accelerate decision making and help the authority to achieve optimization on the site. The objective of the paper is to discuss the permutations and combinations of a particular parameter with one specific method as per the various field conditions and requirements of the authorities.*

Keywords: *flood frequency, probability distribution, optimization, parameter, a specific method*

1. INTRODUCTION: To describe the flood frequency at a particular site, the choice of an appropriate probability distribution and parameter estimation method is of immense importance. The appropriate selection of probability distribution and a parameter estimation method is important for at-site flood frequency analysis. Generalized extreme value, three-parameter log-normal, generalized logistic, Pearson type-III Distributions have been considered to describe the annual maximum steam flow at the site. The performance of these distributions is assessed based on goodness-of-fit tests and accuracy measures.

2.LITERATURE REVIEW: Selecting the best probability distribution for at‑site flood frequency analysis; a study of Torne River this research paper says about the generalized extreme value (GEV), Pearson type-III (P3) distribution, generalized logistic (GLO) distribution, Gumbel (GUM) distribution and three-parameter log-normal (LN3) distribution for the analysis of flood frequency at five gauging sites of the Torne River. The Comparison of GEV, Log-Pearson Type 3, and Gumbel Distributions in the Upper Thames River By Nick Millington Samiran Das And S. P. Simonovic(September 2011).In this study, historical data from the London International Airport station has been used, along with 11 different Atmosphere-Ocean Global Climate Models (AOGCMs), which are used to predict future climate variables. These models produced a total of 27 different data sets of annual maximum precipitation over 117 years, for storm durations of 1, 2, 6, 12, and 24 hours. The current Environment Canada recommended distribution is the Gumbel (EV1) distribution, and the current United States distribution is the Log-Pearson type 3 (LP3). This report investigates a third distribution, the Generalized Extreme Value (GEV) distribution, in the context of the Upper Thames River Watershed.

3. METHODOLOGY: To describe the flood frequency at a particular site, the selection of an appropriate probability distribution is always important. We have considered generalized extreme value (GEV), Pearson type-III (P3) distribution, and Gumbel (GUM) distribution for the analysis of flood frequency at gauging sites. The goodness-of-fit tests are used to verify that the observed data follow a particular distribution.

3.1 Statistical Distributions: The appropriateness of the distributions is investigated by the goodness of fit tests. For the goodness of fit tests, the Anderson-Darling (AD) and the Chi-Squared tests were used. The shape parameter of the GEV distribution was also analyzed, which provides more insight into the goodness of fit of the distribution.

3.1.1 The Generalized Extreme Value (GEV):
The GEV approach is widely applied to model extremes of hydrologic processes such as floods and rainfall. The GEV distribution is a family of continuous probability distributions combining the Gumbel (EV1), Frechet, and Weibull distributions. Location, scale, and shape are the three parameters used by GEV. The location parameter describes the shift of a distribution, the scale parameter describes how spread out the distribution is and the shape parameter, which strictly affects the shape of the distribution, and governs the tail of each distribution. The shape parameter is derived from skewness, as it represents where the majority of the data lies, which creates the tail(s) of the distribution. When the shape parameter (k) =0, this is the EV1 distribution. When k>0,

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 08 | Aug 2020 www.irjet.net p-ISSN: 2395-0072

this is EV2 (Frechet), and when k<0 is the EV3 (Weibull). A large problem in working with the Extreme Value distributions is determining whether to use Type 1, 2, or 3. In general, a distribution with a larger number of flexible parameters, for instance, GEV, will be able to model the input data more accurately than a distribution with a lesser number of parameters. EV1 is effective for small sample sizes, however, if the size is greater than 50, GEV shows a better overall performance. The 3 or 4 parameter distributions often have a negligible bias. As stated in the introduction, the shape parameter for GEV can greatly affect the results. A positive shape parameter will result in the distributions being upper-bounded whereas, negative shape parameter assures that the distribution is unbounded and that results in an increase in magnitudes, as the return period gets larger.

The CDF and PDF are defined in (Hosking, 1997) as:

$$F(x) = \exp\{-(1 - \frac{\kappa(x-\xi)}{\alpha})^{1/k} \quad (2.1)$$

$$f(x) = \alpha^{-1}\exp[-(1-\kappa)y - \exp(-y)] \quad (2.2)$$

where $y = -\kappa^{-1}\log\left[1 - \frac{\kappa(x-\xi)}{\alpha}\right]$, when k≠0

where, ξ is the location parameter, α is the scale parameter, and κ is the shape parameter.

3.1.2 Gumbel distribution

The EV1 distribution only uses 2 parameters, location (ξ), and scale (α). This is the current required method for all Precipitation Frequency Analysis in Canada.
The CDF and PDF as defined in (Hosking, 1997) are:

$$F(x) = \exp\left[-\exp\left(-\frac{x-\xi}{\alpha}\right)\right] \quad (2.3)$$

$$f(x) = \alpha^{-1}\exp\left(-\frac{x-\xi}{\alpha}\right)\exp\left[-exp\left(-\frac{x-\xi}{\alpha}\right)\right] \quad (2.4)$$

where, ξ is the location parameter, α is the scale parameter

3.1.3 Log Pearson Type 3 Distribution (LP3)

The LP3 distribution is a member of the family of Pearson Type 3 distributions and is also referred to as the Gamma distribution. This is the currently required method to be used for all Precipitation Frequency Analysis in the United States. The LP3 distribution is complicated, as it has 2 interacting shape parameters (Stedinger, 2007). Similar to GEV it uses 3 parameters, location (μ), scale (σ), and shape (γ). A problem arises with LP3 as it tends to give low upper bounds of the precipitation magnitudes, which is undesirable.

The CDF and PDF are defined in (Hosking, 1997) as:

The CDF and PDF are defined in (Hosking, 1997) as:

If $\gamma \neq 0$, let $\alpha = 4/\gamma^2$ and $\xi = \mu - 2\sigma/\gamma$

If $\gamma > 0$, then

$$F(x) = G(\alpha, \frac{x-\xi}{\beta})/\Gamma(\alpha) \quad (2.5)$$

$$f(x) = \frac{(x-\xi)^{\alpha-1}e^{-(x-\xi)/\beta}}{\beta^{\alpha}\Gamma(\alpha)} \quad (2.6)$$

if $\gamma = 0$ the distribution is Normal and

$$F(x) = \Phi(\frac{x-\mu}{\sigma}) \quad (2.7)$$

$$f(x) = \phi(\frac{x-\mu}{\sigma}) \quad (2.8)$$

if $\gamma < 0$, then

$$F(x) = 1 - G\,(\alpha, \frac{\xi-x}{\beta})/\Gamma(\alpha) \quad (2.9)$$

$$f(x) = \frac{(\xi-x)^{\alpha-1}e^{-(\xi-x)/\beta}}{\beta^{\alpha}\Gamma(\alpha)} \quad (2.10)$$

Return Period	Gumbel (Excel)	Gumbel (Hyfran)	Log-Pearson-III(Excel)	Log-Pearson-III(Hyfran)	GEV(Hyfran)
2	195.00	193.00	194.50	197.90	197.40
5	256.00	254.00	257.20	259.60	259.50
10	296.00	294.00	298.80	296.90	297.90
20	335.00	333.00		330.40	332.80

3.2 Goodness-of-fit

The goodness-of-fit tests are used to test that the observed data follow a particular distribution. We consider the Anderson–Darling (AD) test for the study. This test is often used in flood frequency analysis and has shown good performance in case of small

Σn i=1

2i−1 n _log (1 − F(yn−i+1)) + log(F(yi))

where F (yi) represents the cumulative distribution function (CDF) of the specified distribution.

Goodness of Fit - Summary

#	Distribution	Kolmogorov Smirnov Statistic	Kolmogorov Smirnov Rank	Anderson Darling Statistic	Anderson Darling Rank	Chi-Squared Statistic	Chi-Squared Rank
38	Log-Pearson 3	0.06678	9	0.15831	5	0.97482	15
25	Gumbel Max	0.07432	20	0.18547	18	1.5597	25
26	Gumbel Min	0.18444	50	2.8359	46	2.353	37
21	Gen. Extreme Value	0.05988	3	0.15602	2	0.86921	13
59	Weibull	0.09606	31	0.74594	33	1.0634	16

4. CONCLUSION

The selection of an appropriate probability distribution for describing flood frequency at a particular site needs to be giving an optimum solution. Highlighting these aspects gives us decision making authority or optimization tool whose objective is to enhance a particular parameter with one specific method. The selection of the method should be based on the need of the client and requirement as well. This paper will be referred to in the future with a scope of a mathematical model or statistical tool for the software making companies make users which give optimum results for an appropriate method of flood frequency.

REFERENCES

1. Advisory Circular (15th August 2013) US Department Of Transportation Federal Aviation Administration-AAA100
2. Ahmed Z., Rao DR., Reddy, and Raj E. (2012). Rainfall Intensity Variation for observed data and derived data - A Case Study of Imphal. ARPN Journal of Engineering and Applied Sciences, Vol. 7, No. 11, 1506-1513.
3. Bojana Mirkovic1 and Vojin Tosic (March 2014) Airport apron capacity: Estimation, representation, and flexibility
4. Chow V., Maidment DR., and Mays LW. (1988). Applied Hydrology. McGraw Hill Book Company, New York.
5. David S. Chang., (21 November 2016) Guidelines for Drainage Studies and Hydraulic Design.
6. Hydraulic Design Series No. 4 (2008) Introduction to Highway Hydraulics US Department Of Transportation Federal Aviation Administration
7. IRC: SP: 13 (2004) Guidelines for the design of small bridges and culverts Indian Roads Congress Special Publication, New Delhi.
8. IRC: SP: 42-2013. Guidelines for the design of small Bridges and Culverts on Road drainage. Indian Roads Congress Special Publication, New Delhi.

9. Nick Millington, Samiran Das and Slobodan P. Simonovic (September 2011) - The Comparison of GEV, Log-Pearson Type 3 and Gumbel Distributions in the Upper Thames River Watershed under Global Climate Models

10. Mahmood Ul Hassan, Omar Hayat, Zahra Noreen (29 October 2019) - Selecting the best probability distribution for at-site flood frequency analysis; a study of Torne River