iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** V **Month of publication:** May 2023

DOI: https://doi.org/10.22214/ijraset.2023.52070

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

Frequency Analysis of Rainfall Data

Nikunj Pawar[1], Dr. Nina R. Dhamge[1], Om Kharkar[2], Vedanti Yeole[4], Utkarsh Siddham[5], Nikhil Meshram[6]
[1, 3, 4, 5, 6]Student, [2]Professor, Department of Civil Engineering, KDK College of Engineering, (Nagpur)

Abstract: The intensity of precipitation plays a crucial role in hydraulic engineering when creating hydraulic structures and evaluating their impact. To design hydraulic projects such as drainage networks, bridges, and road culverts, Intensity-Duration-Frequency (IDF) curves are used. These curves are heavily influenced by project planning, design, and operation in the field of water resources engineering and are also used in various flood structures. In this study, the frequency of rainfall intensity duration was determined in Kamti, Khari, catchments, Nagpur Maharashtra, from 1996 to 2019. Three different frequency analysis methods, including log-normal, Gumbel, and log-Pearson type III distribution, were applied to fit the corresponding curves. Analyzing the precipitation data is of great importance in this field..
Keywords: IDF curves, probability distributions, good-ness of fit test, return period, IDM reduction formula .

I. INTRODUCTION

Analyzing precipitation data is an important area of research for meteorologists, hydrologists, and climatologists. Precipitation data are typically collected over long periods of time and are often used to predict future weather patterns and plan water resource management strategies. Probability distributions such as the lognormal distribution and the logPearson distribution are often used to analyze precipitation data. The mathematical relationship that relates the intensity, duration, and return period of rainfall is known as the intensity-duration-frequency (IDF) curve. Due to the increase in population and infrastructure development, many regions have become vulnerable to flooding risks. To design cost-effective flood control structures, IDF curves can be used. Over the past few decades this study focuses on the IDF relationship of rainfall, and various statistical distributions have been compared. The log Pearson type III (LPT III) and Gumbel was found to be the best method among the studied methods. The following study aimed to create IDF precipitation curves for the kamthi khari region under changing climatic conditions for the return period of 2, 3, 5, and 13 years duration. The lognormal distribution is a probability distribution commonly used to model natural phenomena that are expected to have a wide range of values. For precipitation data, the lognormal distribution is often used to model the frequency of precipitation events. This distribution assumes that the logarithm of the precipitation data follows a normal distribution. This is a useful assumption when dealing with variables with skewed distributions. The log-Pearson distribution is another probability distribution commonly used when analyzing precipitation data. This distribution is a variant of the Pearson distribution transformed to a logarithmic scale. The Gumbel distribution is commonly used to model the distribution of precipitation event magnitudes. This distribution assumes that the data follow a generalized extreme value distribution. This is a useful assumption when dealing with variables with heavy-tailed distributions. Overall, the use of probability distributions such as the lognormal and log Pearson distributions helps meteorologists, hydrologists, and climatologists better understand precipitation patterns and make more accurate predictions about future weather conditions.

II. AREA OF STUDY AND DATA COLLECTION

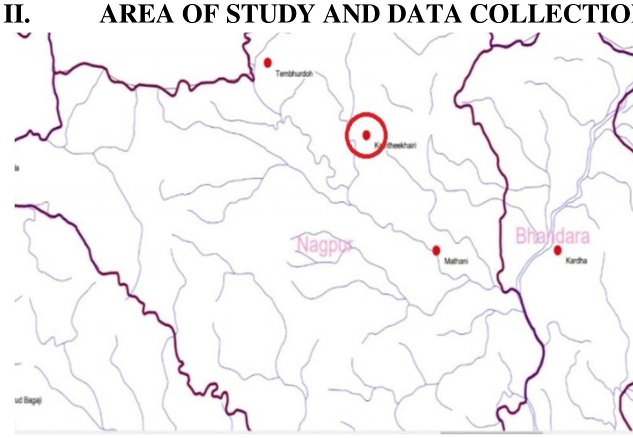

Fig 1 : Kamthi Khairi Catchment

Kamthi khairi Ctachment located at north east region of Nagpur district the catchment experiences tropical therefore the maximum rainfall experienced by the catchment 304 mm while the average rainfall 224 mm The station used to record the rainfall intensity was of FCS(Full climate station) type .

1) Station name: Kamthi Khairi Full climate station.
2) District: Nagpur
3) Tehsil/taluka : Nagpur
4) Major basin/zone: Pench river
5) Local river basin: kanhan river
6) Latitude: 20"31'49" N
7) Longitude: 78°22"26" E

III. ANNUAL RAINFALL DATA

This is the data which has been received for KAMTHI KHAIRI CATCHMENT from the hydrology department. We requested them to issue data from (1996-2019) which was 185269 rainfall intensity . But during the disaggregation of data, we found that values for some years like 2009 , 2011 to 2017 were missing so those years where not considered we and the study was conducted for data span of year 1996 -2019 **(16 years)** rainfall values for further calculation.The rainfall data available from gauging stations of Kamthi- Khairi Nagpur, Maharashtra are of 24hr duration, and the development of IDF Curves required rainfall of shorter durations 2hr, 4hr, 6hr,12hr, and 24 hr.

Table 1: The rainfall depth $P(t)$ for t-hour duration (in mm)

year	Maximum Intensities for particular duration (mm).					
year	1 hr	2 hr	4 hr	6 hr	12 hr	24 hr
1996	70	70	70	70	70	85
1997	15	20	30	48	48.5	56.4
1998	54	79	109.7	109.7	110	134
1999	15.15	50	65	65	70	79
2000	54.5	200	218	218	218	230
2001	44.5	55	61.5	61.5	96	105
2002	45.5	59	69	69	137.5	140
2003	27.833	79.5	84	111	140	145
2004	22.5	37	37	48.5	67	73
2005	19.75	62	80.5	86.5	112.5	118
2006	30	38.5	55	58.5	64	68
2007	24.75	64	76	76	90	96
2008	47	68	124.5	146.5	156	165
2010	28.5	34.5	43	45.6	50	55
2018	15.8	23.7	34	45	48	50
2019	16.7	33.5	36	38	45	52

Results of Descriptive Statistics Of Rainfall Data Duration Series

Statistical Parameter	2 hr	4 hr	6 hr	12 hr	24 hr
std deviation	39.91	45.33	43.93	44.05	49.83
median	79.5	84	111	140	145
mode	110.25	98.41	161.97	216.93	228.57
Coefficient of variance	0.54	0.54	0.54	0.54	0.54
Coefficient of skewness	0.22	0.22	0.22	0.22	0.22

IV. GOODNESS OF FIT TEST

For this study, LPT III, lognormal, and Gumbel distributions were applied for all series of data duration. Software of EasyFit 5.5 was used to fit the probability distributions for every rainfall duration data. For assessing the stability of every distribution of probability, the goodness of fit test, cumulative distribution graph, the least sum of statistic model identification criterion (LSSMIC), and probability density function (PDF) graph were applied. The probability distribution with lowest value of statistics showed higher fitting distribution according to the tests of fit goodness. Figure 2 show the PDFs, cumulative distribution functions (CDFs), probability difference, and probability–probability (P–P) plot of three selected distribution for maximum depth through 16 years Figure 3 displays the probability density function (PDF) of three chosen distributions for a period of 1 hour. The PDF is a positive function that represents the probability distribution through integration, and its integral from negative infinity to positive infinity is always equal to one. It can be described as a continuous histogram that shows the relative frequencies within the ranges of the plotted results. The Gumbel distribution has a PDF value of less than 0.15, while the lognormal distribution is roughly 0.1, and the LPT III distribution is approximately 0.12, all of which are illustrated in the figure. Figure 5 depicts the cumulative distribution function (CDF), which is a function that gives the probability distribution of a random variable with a real number value. The CDF values for the three selected distributions for 1 hour were between 0.1 and 1, as shown in the figure. Figure 4 illustrates the probability difference plot, which shows the difference between theoretical and empirical CDF values.. The EasyFit 5.5 program generates different types of graphs to analyze data distribution. For continuous distributions, it displays a scatterplot or a continuous curve. For separated distributions, it shows a set of vertical lines at every integer x. The probability-probability (P-P) plot is also available, which compares the empirical and theoretical CDF values to assess the goodness of fit. The reference diagonal line is included to help users interpret the graph. Additionally, the descriptive statistics for each duration series are presented in Table 2. The data durations were analyzed, and it was found that their skewness values are very similar, with a value of 0.2268. This finding indicates that the distribution of each data duration is almost symmetrical. Additionally, the coefficient of variance is consistent for all periods, which confirms the accuracy of the program and the results. Specifically, the coefficient of variance is 0.54433 for all periods. To determine the best fit for each distribution, a Chi-square, Anderson–Darling, and Kolmogorov– Smirnov fit test was conducted for each period using EasyFit 5.5. Based on the results of the quality fit test and LSSMIC, it was determined that lognormal distribution is the most appropriate fit, as shown in Table 3.

LSSMIC = Abs (1-sum of statistic)

Table 3 : Best distribution for maximum intensity of rainfall throughout the data interval of 16 years.

Types of Probability Distribution	Kolmogorov Smirnov Statistic	Anderson Darling Statistic	Chi square Statistic	LSSMIC	Best fit
Lognormal	0.1623	0.48334	0.09736	0.2578	
Gumbel	0.17344	0.47077	0.22155	0.13424	Best
LPT III	0.16859	0.41715	0.20541	0.20885	

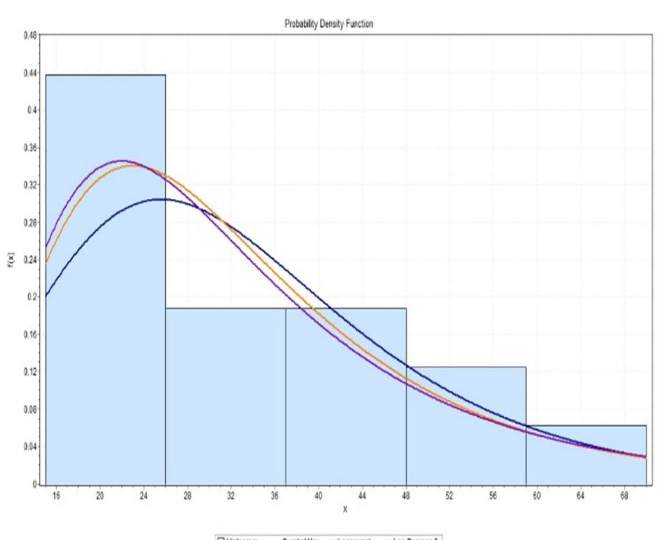

Fig 3 : PDFs of three selected distributions for 1 h

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

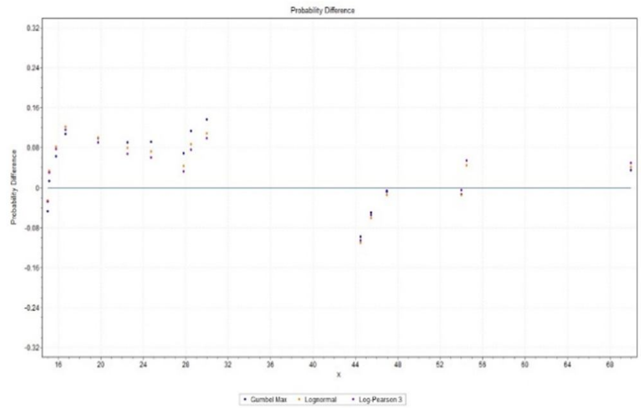

Fig 4 : Probability difference of three selected distributions for 1 h

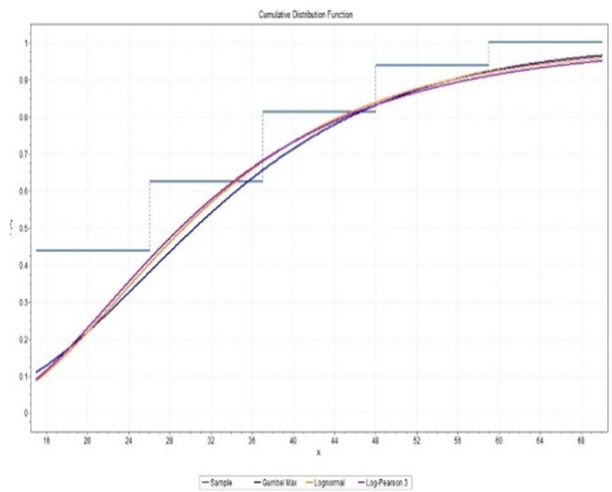

Fig 5:CDFs of three selected distributions for 1h

V. FITTING OF GUMBEL DISTRIBUTION

The Gumbel distribution method, also known as the generalized extreme distribution value Type-I, is a useful tool in probability and statistics. It can be utilized to model the distribution of the highest or lowest value from a set of various distributions. By using the following equation, one can determine the precipitation frequency (in mm) for any given duration and return period (T_R) in a year.

$$P_T = P_{avg} + K_T S \qquad \text{---(1)}$$

The given formula uses P_T to denote rainfall frequency measured in mm for a specific duration. S is used to represent the standard deviation of precipitation data. P_{avg} is the annual average of precipitation data. Lastly, K_T is the Gumbel frequency factor that can be calculated using the given equation.

$$K_T = \frac{\sqrt{6}}{\pi}\left\{0.5772 + \text{In}\left[ln\left(\frac{TR}{TR-1}\right)\right]\right\} \qquad \text{---(2)}$$

where, T_R is the return period (2,3,5,13) years. Then, the rainfall intensity I_T (in mm/h) for return period TR is obtained by equation.

$$I_T = \frac{PT}{d} \qquad \text{---(3)}$$

where d represents the time duration in hour.

The equation (2) is used to calculate the frequency factor. This factor helps determine the return periods of 2, 3, 5, and 13 years for all durations. Equation (1) can be used to find the rainfall frequency (Pt) for a certain duration and return period (T_R) in mm. Equation (3) can then be used to calculate the rainfall intensity (IT) in mm/h for the TR return
period. The findings in Table 4 indicate that rainfall intensity decreases as storm duration increases. Additionally, rainfall of a certain duration has a higher intensity when it has a high return period. The Gumbel method's IDF curves are illustrated in Figure 6, plotted on the normal scale for T_R values of 2, 3, 5, and 13 years.

Table 3 : Computed Frequency factors using Gumbel's fromula.

Return period (T_R)	2	3	5	13
Frequency factors (K_T)	-0.1641	0.25396	0.77196	1.519

Table 4 : calculated intensity (IT) in (mm/h)(Gumbel's method)

Duration (hr)	Intensity (mm/hr)			
	2 years	3 years	5 years	13 years
2	28.7876	37.13071	47.46829	62.37676
4	17.33876	22.07646	27.94674	36.41264
6	13.05055	16.11157	19.90436	25.37417
12	7.858541	9.393202	11.29473	14.03705
24	3.959765	4.827872	5.903505	7.454743

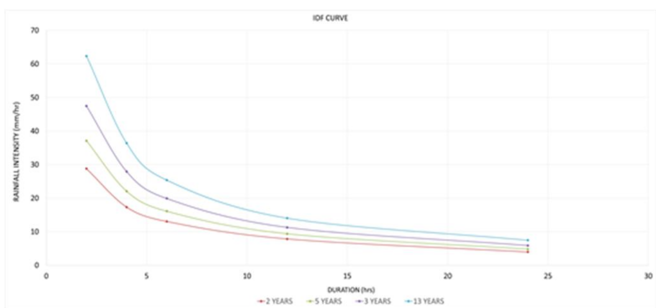

Fig 6 : IDF curve using Gumbel method.

VI. CONCLUSION

1) The study of 16 years (1996 -2019) of rainfall data for kamthi khariri catchment was carried out and frequency analysis was carried out
2) It was deduced that the highest intensity was observed in the 13-year return period, while the lowest intensity was recorded in the 2-year return period.
3) The Easy Fit 5.5 software was used to identify the best distribution of results. The LSSMIC statistical model was used to assess the fit of each probability distribution, and it was found that the distribution was the best option. The Gumbel distribution had the lowest LSSMIC value of 0.82895 and was selected as the ideal distribution for the data related to Nagpur city.
4) It is recommended to use the newly produced IDF curves to estimate rainfall intensities for different return periods. By utilizing the revised IDF curve, it is possible to construct and maintain urban water management systems, such as culverts and bridges, among other things.
5) The results obtained from the current study on precipitation data of Kamthi Khairi indicate that there is no significant difference in rainfall analysis results of IDF curves between the applied methods. This can be attributed to the sub-tropical climate and flat terrain dominant in Nagpur, which cause only slight differences in rainfall values.

REFERENCES

[1] Kumar Abhishek, Abhay Kumar (2012). "A rainfall prediction model using artificial neural network." International Journal of Engineering and Advanced Technology (IJEAT), 1(5), 140-143.

[2] Sharma, N. K. and Kumar, A. (2018). Frequency analysis of extreme rainfall events of Dharamshala region. Journal of the Geological Society of India, 92(6), 715-723.
[3] Joshi, H., & Tyagi, D. (2019). Forecasting modeling Monthly Rainfall in Bengaluru, and Application of Time Series Models. Journal of Earth Science & Climatic Change, 10(2), 476.
[4] Nanda, S. K., & Tripathy, D. P. (2015). Prediction of rainfall in India using artificial neural network (ANN) models. Journal of King Saud University-Engineering Sciences, 27(3), 312-324.
[5] Asim, M., & Nath, S. (2018). Study on Rainfall Probability Analysis at Allahabad District of Uttar Pradesh. International Journal of Engineering Technology Science and Research, 5(8), 284-288.
[6] Kumar, A., Varshney, A., Arya, A., & Bhatia, M. K. (2019). Survey on rainfall prediction using artificial neural network. International Journal of Emerging Trends in Engineering Research, 7(12), 118-124.
[7] Kumar, P., Kumar, S., Kumar, P., & Bhushan, M. (2022). Rainfall prediction using artificial neural network and sequential modeling. Journal of Water &Land Development, 50(1), 27-38. doi: 10.24425/jwld.2022.137654
[8] Thakur, N., Karmakar, S., & Soni, S. (2018). Rainfall Forecasting Using Various Artificial Neural Network Techniques – A Review. International Journal of Emerging Technology and Advanced Engineering, 8(10), 57-64.
[9] Ghosh Dastidar, A., & Ghosh, S. (2017). Statistical analysis of monsoon rainfall distribution over West Bengal, India. Theoretical and Applied Climatology, 128(3-4), 907-917.
[10] Zakwan, M., & Ara, Z. (2017). Statistical analysis of rainfall in Bihar. International Journal of Engineering T Technology Science and Research, 4(4), 527-532.